



## OPEN ACCESS

## EDITED BY

Mariëtte Huizinga,  
VU Amsterdam, Netherlands

## REVIEWED BY

Sabine Doebel,  
George Mason University, United States  
Hanna Mulder,  
Utrecht University, Netherlands

## \*CORRESPONDENCE

Sabrina Ann Kenny  
✉ sakenny@buffalo.edu;  
✉ sabrina.kenny@oregonstate.edu

RECEIVED 14 December 2022

ACCEPTED 26 May 2023

PUBLISHED 20 June 2023

## CITATION

Kenny SA, Cameron CE, Karing JT, Ahmadi A, Braithwaite PN and McClelland MM (2023) A meta-analysis of the validity of the Head-Toes-Knees-Shoulders task in predicting young children's academic performance. *Front. Psychol.* 14:1124235. doi: 10.3389/fpsyg.2023.1124235

## COPYRIGHT

© 2023 Kenny, Cameron, Karing, Ahmadi, Braithwaite and McClelland. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# A meta-analysis of the validity of the Head-Toes-Knees-Shoulders task in predicting young children's academic performance

Sabrina Ann Kenny<sup>1\*</sup>, Claire E. Cameron<sup>2</sup>, Jasmine Tua Karing<sup>1</sup>, Ahmad Ahmadi<sup>1</sup>, Paige Noelle Braithwaite<sup>1</sup> and Megan M. McClelland<sup>1</sup>

<sup>1</sup>Human Development and Family Sciences, Oregon State University, Corvallis, OR, United States,

<sup>2</sup>Department of Learning and Instruction, University at Buffalo, Buffalo, NY, United States

The present study represents the first meta-analytic synthesis of the utility of a widely used early-childhood self-regulation measure, the Head-Toes-Knees-Shoulders task, in predicting children's academic achievement. A systematic review of the literature yielded 69 studies accessed from peer reviewed journals representing 413 effect sizes and 19,917 children meeting the complete set of inclusion and exclusion criteria. Robust variance analysis demonstrated that the Head-Toes-Knees-Shoulders task was a consistent predictor of children's academic achievement across literacy, oral language, and mathematical outcomes. A moderator analysis indicated that in accordance with prior research, the Head-Toes-Knees-Shoulders task was more strongly associated with children's mathematics performance relative to their performance on language and literacy measures. The results of this meta-analysis suggest that the Head-Toes-Knees-Shoulders task demonstrated statistically significant, positive associations with children's overall academic performance. These associations remained stable across different participant and measurement factors and are comparable to meta-analyses examining the self-regulation and academic association with multiple measures of self-regulation and executive function.

## KEYWORDS

self-regulation, executive function, meta-analysis, academic achievement, early childhood assessment

## Introduction

Early childhood experts agree that children who can express their emotions and manage their behaviors to meet the demands of their classroom context begin school with an academic and developmental advantage (Korinek and deFur, 2016; McKown, 2017). Self-regulation, which is included under the self-management competency in the social and emotional learning literature (Collaborative for Academic, Social and Emotional Learning, n.d., para. 3), refers to the skills involved to exert control over one's own actions, thoughts, and feelings in accordance with the social expectations of the environment (Willoughby et al., 2012; Blair and Raver, 2015; Korinek and deFur, 2016; Nigg, 2017). Scholars suggest that the construct of self-regulation is multidimensional and involves both conscious (i.e., top-down) and unconscious (i.e., bottom-up) biopsychosocial processes (Best and Miller, 2010; McClelland and Cameron, 2012; Blair, 2016; Nigg, 2017). The conscious aspect of self-regulation depends greatly on three executive function (EF) components: working

memory (the ability to hold and manipulate information in short-term memory), attentional flexibility (the ability to shift one's thoughts and attention when necessary), and inhibitory control (the ability to suppress a habitual response in favor of a more socially acceptable one; Best and Miller, 2010; Nigg, 2017; Gonzales et al., 2021). In the current study, we focus on these cognitive processes as they relate to children's behavioral responses in school, such as paying attention, ignoring distractions, and following directions while completing academic tasks. We are particularly concerned with the extent to which children can apply and integrate their executive functions into observable motor actions. However, we acknowledge that the conceptualization of self-regulation, executive functioning, and other related constructs may have slight variations across disciplines with different scholarly traditions and measurement practices (Booth et al., 2018).

Given the widespread consensus that self-regulation is essential to children's successful school participation (Rimm-Kaufman et al., 2009; McClelland et al., 2014; Blair and Raver, 2015, 2016; Edossa et al., 2018; Hernández et al., 2018; Robson et al., 2020; Finders et al., 2021), there has been a rise in self-regulation interventions across research and practice settings (Reid et al., 2005; Raver et al., 2011; Carr et al., 2014; Flook et al., 2015; Schmitt et al., 2015; Pandey et al., 2018; McClelland et al., 2019; Hahn-Markowitz et al., 2020; Razza et al., 2020; Welsh et al., 2020). Nevertheless, the literature has continuously acknowledged the scarcity of high-quality, direct (i.e., performance-based) assessments of self-regulation suitable for young children and relevant to educational contexts (Raver et al., 2012; Halle and Darling-Churchill, 2016; Lipsey et al., 2017; McKown, 2017). A topic that has received much attention both in the U.S. and abroad has been on the validity of self-regulation measures in predicting outcomes deemed meaningful in school, most particularly, young children's academic achievement (e.g., Gestsdottir et al., 2014; McClelland et al., 2014, 2021; Schmitt et al., 2014, 2017; Lipsey et al., 2017; Hee et al., 2018).

One assessment tool used frequently in educational research is the Head-Toes-Knees-Shoulders task (HTKS; McClelland et al., 2007; Cameron Ponitz et al., 2008, 2009). The HTKS is a direct assessment of self-regulation that asks children to do the opposite of a series of gross motor actions (e.g., "touch your head when I say touch your toes"). If the child succeeds, more difficult instructions requiring rule-switching are introduced. A distinct advantage of the HTKS is that it emulates the real challenges of school, requiring children to regulate observable behavior by integrating all three aspects of EF simultaneously: (a) working memory because they listen to, remember, and respond to the multiple rules and commands; (b) inhibitory control because they must suppress their natural response to touch the body part they are being asked to touch; and (c) attentional flexibility because they focus on the multiple aspects of the task and transition between the old and new rules (McCabe et al., 2000, 2004; McClelland and Cameron, 2012; Gonzales et al., 2021). Thus, from a conceptual standpoint, the HTKS aligns well with contemporary, strength-based perspectives that emphasize the contextually and culturally situated nature of executive functioning (Doebel and Lillard, 2023).

The HTKS was developed in the U.S. and is now available in 28 languages. For over a decade, it has become a popular

choice for early childhood researchers examining self-regulation in populations with diverse racial, cultural, and socioeconomic backgrounds (Wanless et al., 2011a; Gestsdottir et al., 2014; Lonigan et al., 2017; Zhang and Rao, 2017; Hee et al., 2018; Hernández et al., 2018). To date, the HTKS has been used primarily by researchers, but several of its characteristics make it a promising prospect for educational practice; these include a straightforward administration/scoring procedure, portability, no technology required, developmental appropriateness, and testing items that relate to what children are frequently expected to do in real-life school settings. The HTKS has been viewed as an ecologically valid self-regulation measure because of its focus on assessing classroom behaviors that help children navigate early learning contexts. In addition, the gross motor component of the HTKS more closely approximates children's behavioral self-regulation compared to other measures (McClelland and Cameron, 2012).

Several researchers have tested the reliability of the HTKS, and the data show strong psychometric properties with little variability across studies, including coefficient alpha, test-retest reliability, and interrater reliability (Cameron et al., 2012; McClelland et al., 2014; Lipsey et al., 2017; Zhang and Rao, 2017; Hee et al., 2018; Liu et al., 2018; Spiegel and Lonigan, 2018; Cerino et al., 2019; Sezgin et al., 2019). Some studies have also found the HTKS predictive of academic outcomes and significantly related to children's working memory, attentional flexibility, and inhibitory control (Cameron Ponitz et al., 2008, 2009; Gestsdottir et al., 2014; McClelland et al., 2014, 2021; Schmitt et al., 2014, 2017; Fuhs et al., 2015; Lipsey et al., 2017; Hee et al., 2018; Gonzales et al., 2021). Recently, the most current version of the HTKS, the HTKS-R, showed a significant increase in its ability to capture variability in children with lower levels of self-regulation compared to earlier HTKS versions (Gonzales et al., 2021).

Considering the frequent use of the HTKS in school-based research across various cultural contexts, it is critical to understand all the available evidence on the instrument's measurement characteristics, most particularly, its relevance to educational outcomes. Statistically synthesizing the results of several studies will provide researchers and practitioners with a cohesive representation of this knowledge. Because the HTKS has grown in use since the first empirical studies in 2007–09, it would be challenging to evaluate such a large body of literature objectively without a quantitative approach. Borenstein et al. (2009) warned that attempting to synthesize the results of numerous studies without meta-analytic techniques often leads researchers to make inferences that are not well-informed because such an approach relies so heavily on one's subjective reasoning and judgment. Similarly, without a systematic and transparent approach to database searching and study selection, the chances of overlooking relevant publications with the HTKS is high (Borenstein et al., 2009).

To provide a more objective estimate of the HTKS and academic association, the primary purpose of this study was to systematically gather and quantitatively combine all the available evidence on the association between children's (ages 3–8 years of age) HTKS performance and their academic achievement in mathematics, oral language, and literacy outcomes with

meta-analytic techniques. Children's performance on mathematics, oral language, and literacy measures were the outcomes of focus in this study because these are the skills that are targeted most heavily in early childhood school settings. Other researchers have measured the behavioral regulation and academic association meta-analytically with multiple measures of self-regulation (Dent, 2013; Robson et al., 2020). These studies have provided valuable insights into self-regulation at the construct level, but are less valuable for making inferences related to self-regulation measurement. To our knowledge, this study represents the first attempt to examine the self-regulation and academic association meta-analytically with the HTKS only. Focusing on the measurement properties of the HTKS is significant because of how frequently researchers rely on the tool as an outcome measure.

## Variability in the HTKS and academic performance association

Meta-analysis is a quantitative design that uses statistical techniques to combine the results of existing studies related to a particular research topic or question (Borenstein et al., 2009; Wilson, 2019). In addition to providing a potentially more objective estimate of the association between children's performance on the HTKS and their academic achievement, a meta-analysis can uncover trends and patterns in the data that would otherwise be difficult to identify. For example, meta-analytic techniques can estimate the degree to which variability in effect size estimates stem from real effect size variation and identify factors that might be contributing to this variation (Borenstein et al., 2009). Unlike a traditional narrative review, a meta-analysis can, therefore, shed light on the circumstances and contexts where the HTKS performs best in predicting children's academic performance.

Variations in studies' findings can occur by chance, but several measurable factors can also potentially alter the magnitude or direction of the association between children's self-regulation and their academic performance. As mentioned earlier, researchers had meta-analyzed the association between children's academic performance and self-regulation on the construct level. The direction of the association remained positive and significant in these meta-analyses; however, some scholars have found the magnitude of the effect size to vary as a function of different participant and measurement factors (Dent, 2013; Robson et al., 2020). In the present study, we expected that the association between children's HTKS performance and their academic achievement would be positive and significant across studies but might vary as a function of the following participant, contextual, and measurement characteristics: average age of the sample, the country where the study took place, the academic subject domain, and the testing occasion. Understanding the degree to which the HTKS correlates with children's academic performance across different samples under various circumstances holds critical implications for educational researchers and practitioners seeking self-regulation measures that are consistently predictive of meaningful educational outcomes.

## Variation by age

There is some evidence that age may moderate the association between children's self-regulation skills and academic performance. For example, in a meta-analysis spanning a wide age range of children and adolescents, Dent (2013) found the average association between children's self-regulation skills and their academic achievement to be significantly stronger in elementary school children than preschool children. Although the exact mechanisms by which children's age influences the association between their self-regulation skills and academic achievement are not fully understood, we hypothesized that age may moderate the HTKS and academic achievement association in the present study because the HTKS captures less variability in children with emergent self-regulation skills (Gonzales et al., 2021; McClelland et al., 2021). The HTKS was recently revised (i.e., HTKS-R) to address this concern, but the first publication of the HTKS-R measure was in 2019 (McClelland et al., 2019). Therefore, because most of the investigations to date of the HTKS relied on the original HTKS, we proposed that the strength of the HTKS and academic performance association might decrease as the average age of the samples decreases.

## Variation by country

Several researchers outside the U.S. have found the HTKS predictive of children's academic performance (Lan et al., 2011; Wanless et al., 2011b; von Suchodoletz et al., 2013; Gestsdottir et al., 2014; Cadima et al., 2015; Birgisdottir et al., 2020; Lenes et al., 2020). Nevertheless, because society plays such a vital role in shaping how children behave in school, how teachers expect children to behave, and the academic demands of an early learning environment, it is plausible to suspect that the association between children's HTKS performance and their academic performance might vary from country to country (Wanless et al., 2011a,b; ten Braak et al., 2019). Cultural differences intersecting with measurement factors may also contribute to variations in the strength of the self-regulation and academic association. For example, scholars have suggested that children in China may perform better on self-regulation measures than their peers from other countries because of the Asian societal emphasis on self-discipline and behavioral control (Zhang and Rao, 2017). Thus, the ceiling effect on the HTKS may be more prominent in children from Asian countries; if the HTKS captures less variability in Asian children, it may yield weaker academic associations when used with this population.

## Variation by academic domain

Cross-sectional and longitudinal research has shown that children with well-developed self-regulation, including its underlying executive functioning components, demonstrate more positive short and long-term outcomes in mathematics, literacy, and oral language (Dent, 2013; Fuhs et al., 2014; Cadima et al., 2015; Cantin et al., 2016; Lonigan et al., 2017; Purpura et al., 2017; Schmitt et al., 2017; Hernández et al., 2018; McClelland and Cameron, 2019; Valcan et al., 2020). The strongest relations have been found with mathematics (McClelland et al., 2014; Blair et al., 2015). One common explanation linking self-regulation to

academic achievement relates to automaticity, a theory stating that through practice and repetition, most people will eventually become capable of performing specific skills with little-to-no conscious effort and thought (Wulf et al., 2001; Floyer-Lea and Matthews, 2005).

The concept of automaticity can shed light on why the strength of the HTKS and academic association might vary across different academic skills. In other words, academic skills can become automated and, therefore, necessitate less executive control (Cameron, 2018).

To illustrate, scholars proposed that although children will depend greatly upon all three key EF for early literacy (e.g., alphabet knowledge, decoding, and phonological awareness) acquisition initially, with practice and repetition, they will eventually be capable of applying these skills with little conscious effort given the relatively small amount of information to be acquired (Paris, 2005; Spiegel et al., 2021). Similarly, children often learn much of their vocabulary skills through social interactions as opposed to direct instruction (Fuhs et al., 2015). Given this frequent exposure, much like early literacy skills, oral language (e.g., vocabulary) acquisition may require less mental effort in early childhood than academic domains that continue to grow in complexity and novelty, such as mathematics (Blair et al., 2008; Fuhs et al., 2015). We, therefore, suspected that the association between children's HTKS performance and children's academic achievement would be strongest in mathematics.

### Variation by testing occasion

Research on the self-regulation and academic achievement association includes cross-sectional studies and longitudinal studies with multiple waves of data collection. We hypothesized that cross-sectional studies might yield substantially stronger correlations between children's HTKS performance and their academic achievement due to common factors within testing situations that occur simultaneously or very close in time. In other words, test concurrency could inflate the association between children's HTKS performance and their academic performance.

### Research aims

The purpose of this study was to systematically gather a wide range of studies containing both HTKS and academic outcomes and to examine this data with meta-analytic methods. In so doing, we extended knowledge about one of the most frequently and globally used self-regulation direct assessments with the following research questions. Though these research questions were not pre-registered, we established them prior to beginning the meta-analysis.

1. What is the overall association between the HTKS and children's academic achievement and is there evidence of variability in this effect?

2. Do between-study differences in participant, contextual, and measurement characteristics influence the magnitude of the HTKS and academic achievement effect?
3. Do within-study differences in participant, contextual, and measurement characteristics influence the magnitude of the HTKS and academic achievement effect?

## Methods

The methods included in this meta-analysis, which are based on best practice guidelines (Borenstein et al., 2009; Kugley et al., 2017; Wilson, 2019; PRISMA, 2020), entailed (a) systematically searching for studies to address the pre-established research questions above; (b) screening the studies based on pre-established inclusion and exclusion criteria; (c) reviewing each study that met the criteria for the full text review to determine whether they met the full set of criteria for the quantitative analysis; (d) carefully reviewing and coding the studies to extract and understand the data; (e) calculating the individual studies' effect sizes and variances; (f) combining the studies to calculate a weighted mean (i.e., summary effect, average effect); (g) examining heterogeneity with a moderator analysis; (h) assessing study quality; and (i) testing for publication bias. These steps are presented in more detail below. To guide this meta-analysis, we followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) to the greatest extent possible: "PRISMA is an evidence-based minimum set of items for reporting in systematic reviews and meta-analyses" (PRISMA, 2020, para. 1).

### Search process

Following the Campbell Collaboration's information retrieval guidelines (Kugley et al., 2017), in the winter of 2021 and the spring of 2022, the first author conducted several searches using ERIC, PsycINFO, Psychology and the Behavioral Sciences, and Web of Science. The search was restricted to the years 2007 through 2022 because HTKS publications did not exist before this time. To increase the chances of obtaining all relevant research with the HTKS instrument, the search terms included a Boolean, all-text search with the "or" operator combining all of the possible nomenclatures of the outcome measure (e.g., HTKS **OR** Head-Toes-Knees-Shoulders task **OR** Head Toes Knees Shoulders **OR** Head-to-Toes-Task). After this search, the first author cross-referenced the findings with a running list of HTKS studies; one of the authors of the HTKS, updates this list regularly to keep track of who is using the instrument. The first author also searched the publication reference lists of meta-analyses examining the association between children's self-regulation (and related constructs) and their academic performance.

### Study selection

To be included in the meta-analysis, the research had to meet the following full set of eligibility criteria: (a) provide a

full text in English that is published in a peer-review journal; (b) report original data only (i.e., studies reporting coefficients from other studies such as meta-analyses or narrative reviews were excluded); (c) report quantitative data from the HTKS as an outcome measure; (d) report quantitative data from at least one math, literacy, or oral-language outcome; (e) provide one or more 0-order Pearson correlation coefficients between the HTKS and children's academic performance (either in the article or obtained from the researchers); (f) use a sample of children who were between 3 and 8 years old at the time of HTKS testing; (g) not use an unconventional HTKS administration approach, such as a gamified version; and (h) not exclusively target children with developmental delays, disabilities, or emotional and behavioral challenges. For academic outcome measures, both standardized assessments and teacher reports or ratings were considered acceptable. Regarding design, both cross-sectional and longitudinal studies were included. In addition to non-experimental/observational studies, studies examining the effects of interventions were included when the correlations at baseline (prior to intervention) or for the control group were available in the article or from the researchers.

The final comprehensive search process yielded 403 studies ( $k = 218$  from PsycINFO,  $k = 70$  from Psychology and the Behavioral Sciences,  $k = 76$  from Web of Science, and  $k = 39$  from ERIC). After the first author removed all duplicates (e.g., same studies appearing across multiple databases),  $k = 311$  studies remained. This total included an additional 19 studies that were not found via the database search; the first author accessed the 19 publications via the previously mentioned running list of HTKS studies and the reference lists of relevant meta-analyses. Using a pre-established abstract screening checklist (available upon request from the corresponding author), the first author screened all records independently; a random sample of ~15% of the studies were then selected so that a team of two screeners (trained by the first author but blinded to her screening decisions) could conduct a double screening using the same abstract screening protocol. All three screeners reached an agreement level of 98 percent over which studies qualified for the full-text review.

The abstract screening process resulted in the elimination of 120 studies for one or more of the following reasons per the abstract screening checklist: the article contained no HTKS outcome; the article contained no academic outcome; the sample included adults or adolescents only; the article was a literature review or meta-analysis; the full paper was not available in English; the paper was not published in a peer reviewed journal. Following the abstract screening, the first author read each of the remaining articles ( $k = 191$ ) with a pre-established, full-text eligibility checklist stating the eligibility criteria for the quantitative analysis (available upon request from the corresponding author). Like the abstract screening process, a team of two assessors trained by the first author but blinded to the first author's eligibility decisions assessed a random sample of 15% of these remaining articles using the same full-text eligibility protocol. All three screeners reached an agreement level of 93% about which studies qualified for the meta-analysis. A written record of

disagreements and resolutions is available upon request from the corresponding author.

After the full-text review process, including attempts to obtain relevant missing data from the corresponding authors of studies with no 0-order correlations, the next step was to ensure that each study's correlations were based on a distinct sample. Therefore, the first, fourth, and last author conducted a sample comparison across all studies to assure the independence of each sample. Prior to analysis, the first author removed all duplicate datasets and studies in which the correlations could not be obtained.

The full inclusion/exclusion process resulted in a total of  $k = 69$  studies eligible for the quantitative analysis. A reference list of the included studies is in [Appendix A](#) in the Supplementary material. The reference list of excluded studies with the reasons for exclusion is in [Appendix B](#) in the Supplementary material. A PRISMA flow diagram outlining the different phases of the review is presented in [Figure 1](#). In both the abstract screening and full-text inclusion/exclusion process, all assessors resolved 100% of their discrepancies via a discussion and review of the eligibility criteria.

## Data extraction

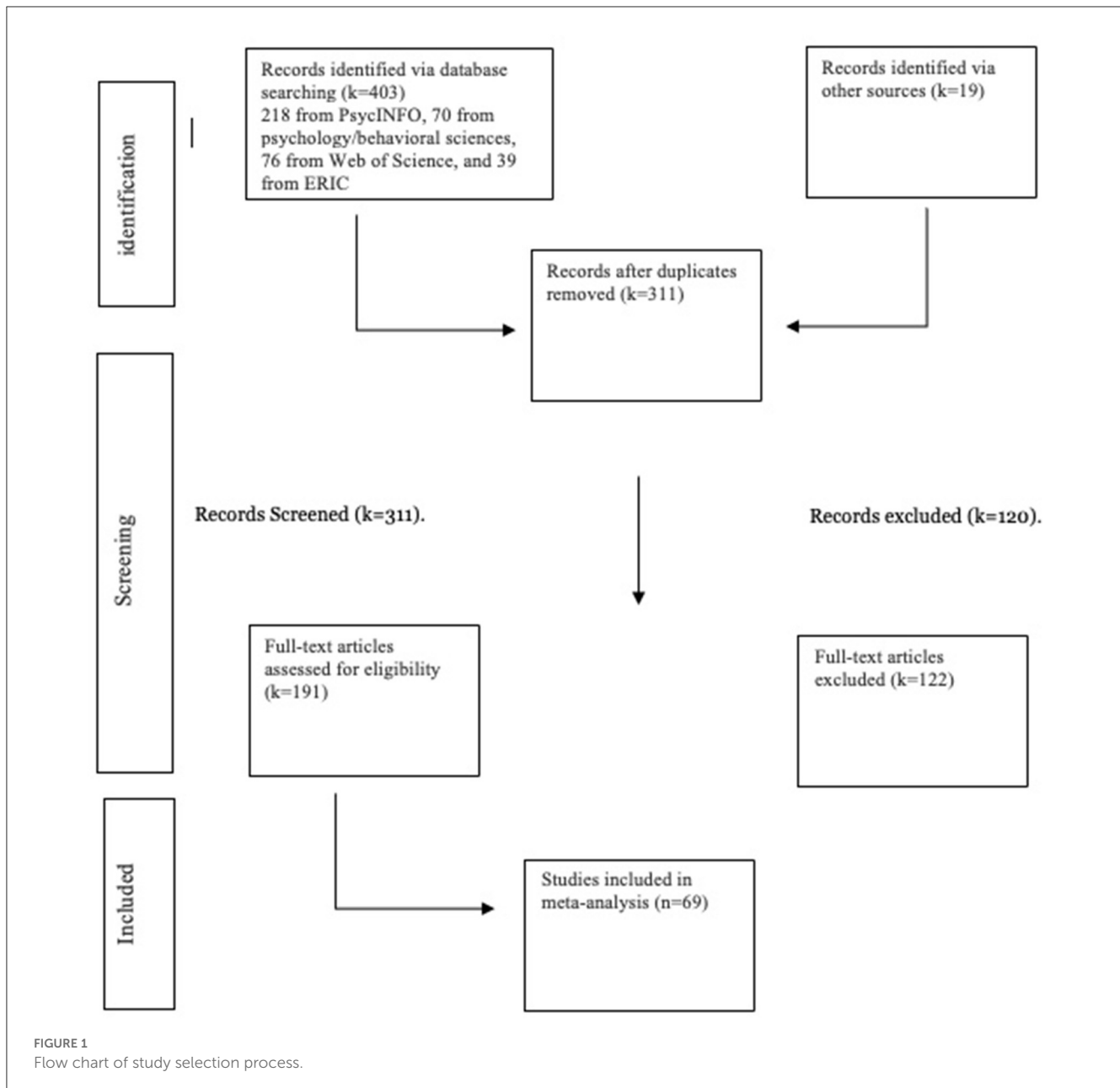
In addition to extracting each study's identifying information (e.g., authors, title, publication year, and journal) and the information needed to calculate the effect size (analytic sample size and 0-order correlations), the first author coded several study-level characteristics. Additionally, each study was given a score representing its overall study quality. All coding was conducted via a formal coding protocol developed by the first author (available upon request). All moderators and effect size coding were entered into Excel and checked for data entry errors prior to transferring it into Stata for the primary analysis. The first author also checked for data entry errors in Stata by comparing descriptive statistics calculated in both Stata and Excel.

## Moderator coding

In a meta-analysis, a moderator analysis uses statistical techniques like ordinary least squares regression (i.e., meta-regression) in the attempts to identify study-level factors that contribute to differences in the outcome of interest (i.e., the association between children's HTKS performance and their academic achievement). We selected the variables described below as effect-size moderators because (a) they pertained to the pre-established research questions; (b) were reported consistently and completely throughout the papers, and (c) demonstrated noticeable variation across papers.

### Age

We initially recorded the mean age of the participants in whatever way the researchers reported it (e.g., months or years). For the final analysis in Stata, age was represented by the average age of the sample in months at the time of HTKS testing. Per our inclusion/exclusion criteria, we only included studies and



timepoints where the average age of the sample was between 36 and 96 months old at the time of HTKS data collection.

### Country

During the initial coding process, we recorded the country where each study took place. However, for the final analysis, the country variable was collapsed into the following four categories: Asian countries, European countries, United States, and a category signifying countries outside of Asia, Europe, and the United States (i.e., other).

### Testing occasion

Because of our choice to include both cross-sectional and longitudinal data, an important measurement characteristic of the HTKS was the time it was administered to children relative to the

academic achievement testing. This meta-analysis contains effect sizes representing the HTKS and academic performance association where (a) both the HTKS and academic testing occurred at the same time point or where (b) the HTKS was administered to children prior to the academic testing (e.g., HTKS administered in the fall of preschool followed by kindergarten academic achievement testing). To capture this variation, initially, we recorded all the time points when the HTKS was administered and all the time points when the academic measures were administered. However, due to a low percentage of responses representing each of these categories, this variable was treated dichotomously for the moderator analysis representing either: (a) the HTKS administered at the same data collection wave as the academic testing and (b) the HTKS administered prior to academic testing (i.e., HTKS administered at time 1; academic measures administered at time 2).

## Academic domain

Informed by prior literature and how studies reported the correlations between the HTKS and the academic achievement measures, we initially created an academic domain variable consisting of the following categories: literacy skills (e.g., phonological awareness, print awareness, letter and word identification, letter-sound knowledge, and decoding); oral-language skills (e.g., expressive and receptive vocabulary, listening comprehension, grammar, and syntax); and mathematics (all skills related to numbers/cardinality, operations, measurement, geometry, and quantitative problem solving). These categories aligned with most of the academic achievement measures in this meta-analysis and other common taxonomies for classifying literacy, oral-language, and mathematics skills (National Governors Association Center for Best Practices Council of Chief State School Officers, 2010a,b; U.S. Department of Health, 2015; Lonigan and Milburn, 2017). The only exception was when researchers reported children's academic performance globally (e.g., a single academic score representing math, literacy, and oral-language). This type of outcome could not be classified into one of the subject-specific domain categories. Therefore, we classified such measures as academic global (i.e., math/language/literacy composite) and language arts global (i.e., language and literacy composite).

For the moderator analysis in Stata, the academic domain variable was transformed into a dichotomous variable representing content that was either math related or language arts related because there were not enough outcomes in each of the initial categories above. For the dichotomous variable, the term language arts refers to all literacy, oral language and language arts global measures collectively. For construct clarity purposes, the academic global measures were dropped from the moderator analysis, resulting in a marginal sample size reduction.

Using a detailed coding protocol, a team of two coders (trained by the first author but blinded to her coding decisions), double coded a random sample of ~15% of the studies and effect sizes representing each of the moderators above. Level of agreement among all three coders was 100% for each variable across all studies.

## Study quality assessment

After reviewing several rating scales and study quality indicators, we used a scale informed by the one that Zangaro and Soeken (2005) developed in their meta-analysis of the reliability and validity of a job satisfaction questionnaire. We revised this scale to include 10 items reflecting what was most critical for examining the validity of the HTKS in predicting children's academic achievement, namely how well the researchers described the academic and HTKS testing and scoring procedure (available in Supplementary material). The range of total quality points was 0–10.

To code for study quality, the first and third author divided the studies and rated each study using the previously mentioned quality rating scale. Following this process, the first and third author double-coded a random selection of ~15% of each other's studies. The average level of agreement across all 10 items was 93%. However, both authors were able to resolve 100% of these disagreements via a discussion and review of a quality rating scale codebook developed by the first author. All sources of disagreement and their resolutions are available upon request.

## Effect size coding

For this meta-analysis, the Fisher  $Z$ -transformation of the correlation coefficient served as the effect size variable. Each study has at least one effect size representing the association between children's HTKS performance and their academic achievement. To calculate the effect size, we extracted the 0-order Pearson's correlations and analytic sample sizes from the studies. The first author then used this information to transform each correlation to Fisher's  $z$  and calculated the variance for each of these estimates with an effect size calculator offered by the Campbell Collaboration. Using a detailed coding protocol (which included instructions for extracting the sample size, 0-order correlations, and converting the 0-order correlations to Fisher  $Z$ -transformation), the same coding team double-coded a random sample of ~15% of studies representing all effect size variables: sample size, Pearson's correlation, Fisher  $Z$ -transformation, and variance. Apart from sample size, level of agreement was 100% among the three coders. For sample size, level of agreement was 96% due to one coder's overlook regarding the analytic sample of the study.

## Data synthesis and statistical analysis

In the social and behavioral sciences, it is common for a single study to report multiple outcomes that may be of interest to the meta-analyst; yet, traditionally, most meta-analysis techniques have required each effect size to correspond to only one distinct sample. This guideline, although limiting, safeguards against the issue of assigning more weight to a study just because it has more effect sizes (Borenstein et al., 2009).

Meta-analyst scholars have recently determined that robust variance estimation (RVE) is one of few methodologically acceptable ways to combine multiple, statistically dependent outcomes, even when both the source and extent of the dependencies are not fully known (Hedges et al., 2010; Tanner-Smith and Tipton, 2013; Moeyaert et al., 2017; Cheung, 2019; Pustejovsky and Tipton, 2021). In addition to synthesizing more than one effect size in a single study, researchers using RVE can statistically separate and examine both between-study and within-study moderator effects. For example, most commonly, the average age of the sample will vary from study to study (i.e., between-studies) because each study represents different samples of children. However, the average age of the sample can also vary within the studies themselves (e.g., children's age will be different at fall and spring data collection); these two types of variation (i.e., between studies and within studies) could obscure one another's effects in a moderator analysis if they are not accounted for (Tanner-Smith and Tipton, 2013).

In this study, most effect-size dependencies stemmed from researchers collecting data over different time points with the same group of children, researchers reporting more than one academic outcome on the same group of participants, or a combination of these two types of dependencies. We used Robumeta packages offered by Stata 17 (small sample correction feature turned on) with a probability level set at 0.05, random model weights, and a correlated effects model. Robumeta has been endorsed by the Campbell Collaboration (Tanner-Smith and Tipton, 2013;

Pustejovsky and Tipton, 2021). Researchers have found the technique to be robust to normality violations and effective for handling both small and large sample sizes (Pustejovsky and Tipton, 2021).

After data cleaning, descriptive analysis, and outlier testing, the first author calculated the weighted summary effect for all the outcomes combined by estimating an intercept-only model. This analysis was followed by two additional intercept-only models that combined effect sizes representing math measures only and effect sizes representing language arts measures only. The tau-square statistic was used to examine heterogeneity among studies' effect sizes, and Edger's test was used to assess potential publication bias. Because the tau-square statistic suggested some heterogeneity, these analyses were followed with statistical significance testing (three multivariable meta-regression models with a  $p$ -value of 0.05) to examine whether the HTKS and academic achievement association varied as a function of the previously discussed moderators.

## Results

First, we discuss the characteristics of the 69 studies of this meta-analysis. Following this descriptive analysis (Table 1), we present the meta-analytic findings related to the three research questions (Tables 2, 3).

### General description of the papers included in the meta-analysis

We extracted 413 effect sizes from 69 studies representing a total of 19,917 children. These studies were all published in peer reviewed journals; thus, the following results may change if gray literature had been included in the analysis (e.g., dissertations, conference papers).

The number of effect sizes per study ranged between one and 30 with six effect sizes (on average) reported per study. Most papers (over 80%) reported more than one effect size representing the HTKS and academic performance association over different periods of data collection, the HTKS and academic performance association over different academic domains, the HTKS and academic performance association representing more than one distinct group of children, or a combination of these three types of dependencies.

### Sample characteristics

There was a wide range of sample sizes reported (17–2,406) with an average sample size of ~416 across all effect sizes. The cultural contexts of the samples were diverse: 237 effect sizes (57%) from 32 papers originated from children in the United States; 176 effect sizes (43%) from 38 papers came from children living in countries outside of the United States, most commonly European and Eastern Asian countries. Across effect sizes, the average age of the children was 63.22 months ( $SD = 11.25$ ) or a little over 5 years old at the time of HTKS testing. Most of the effect sizes (~95%) were extracted from studies

where the average age of the sample was at least 48 months (4 years old) at the time of HTKS testing, though children's age ranged from 37 to 95.87 months (~3–8 years old). The gender composition was split relatively equally between males and females in most studies.

### Measurement characteristics

Both the HTKS and academic measures were typically administered by the researchers in a quiet area of the school setting. Some researchers administered multiple measures of self-regulation, whereas others administered the HTKS only. Academic measures most frequently consisted of well-known, standardized assessments of achievement (e.g., Woodcock-Johnson-III Academic Achievement Test, Peabody Picture Vocabulary Test, Test of Preschool Early Literacy, and Bracken School Readiness Assessment) and modified versions of these measures. Three studies used teacher ratings of children's academic performance instead of direct assessments. These were combined with the directly assessed academic performance measures given the moderate to high associations between proxy ratings and direct measures of academic achievement (Hoge and Coladarci, 1989; Feinberg and Shapiro, 2003; Südkamp et al., 2012; Mack et al., 2023). The HTKS was administered at the same time as the academic measures across 62% of the effect sizes, whereas 38% of the effect sizes represented HTKS and academic performance associations where the HTKS was administered before academic measures. As evident in Table 1, most effect sizes (over 90%) represented academic outcome measures across either the math, literacy, or oral language domain, while a small portion of effect sizes represented children's academic performance more globally.

### Study quality

Study quality was rated on a scale of 0–10 (lowest quality to highest quality) with items deemed relevant to the focus of this meta-analysis. The papers in this meta-analysis had an overall score ranging from 6 to 10 ( $M = 8.75$ ,  $SD = 1.33$ ).

### HTKS and academic performance association: meta-analytic findings

Each of the effect sizes that we extracted for this meta-analysis represented the continuous association (Pearson's  $r$  converted to Fisher's  $z$ ) between children's HTKS performance and their academic achievement on an assessment of one of five possible content areas: mathematics ( $n = 138$ ), literacy ( $n = 145$ ), oral-language skills ( $n = 107$ ), language arts global ( $n = 16$ ), and academic global ( $n = 7$ ).

Borenstein et al. (2009) explains that combining different but conceptually related outcomes in meta-analysis (e.g., outcomes that represent different academic subjects, outcomes that provide information about different social and emotional domains) is permissible, as long as the findings relate to the research questions. Because we were interested in the association between



TABLE 1 Descriptive analysis.

Variables	<i>n</i>	<i>k</i>	<i>M</i> or %	SD	Range
<b>Country</b>					
US	237	32	57%		
Asia	56	9	14%		
Europe	73	15	18%		
Other	47	14	11%		
Male	381	65	50.06%	3.77%	38.33–63%
Age	413	69	63.22	11.25	37–95.87
Study quality	413	69	8.75	1.33	6–10
<b>Academic domain</b>					
Math	138	44	33.41%		
Literacy	145	44	35.11%		
Oral language	107	42	25.91%		
Language arts global	16	6	3.87%		
Academic global	7	6	1.69%		
<b>HTKS testing occasion</b>					
Same time as academic assessment	258	66	62.47%		
Before academic assessment	155	33	37.53%		
Sample size	413	69	415.69	465.34	17–2,406

Total number of effect sizes: ( $n = 413$ ). Total number of studies: ( $k = 69$ ).

Means (*M*), percentages (%), standard deviations (*SD*), and ranges are based on the total number of effect sizes extracted from studies used in this meta-analysis.

TABLE 2 Average weighted effects for the association between the HTKS and children's academic performance.

Variables	<i>k</i>	<i>N</i>	<i>z</i>	95% CI (LL)	95% CI (UL)	<i>R</i>	<i>p</i> -value	Tau-squared
All measures combined	69	413	0.4	0.37	0.44	0.38	<0.001	0.0197
By academic domain								
Math	44	138	0.49	0.45	0.53	0.46	<0.001	0.0143
Language arts	61	268	0.35	0.33	0.38	0.34	<0.001	0.0122

*k*, number of studies; *n*, number of effect sizes; CI, confidence interval; LL, lower limit; UL, upper limit; *z*, Fisher *z*.

Language arts = all oral language, language arts global, and literacy measures combined into a single category.

the HTKS and children's overall academic performance, we ran one intercept-only meta-regression model to examine the weighted summary effect for all the outcomes combined (overall effect synthesizing all 413 effect sizes). However, because we were also interested in the domain-specific effects of the HTKS on children's academic performance, we ran two additional intercept-only models that meta-analyzed effect sizes separately according to whether the academic outcome was math related ( $n = 138$ ) or language arts related ( $n = 268$ ). The findings are displayed in Table 2 and discussed under research question (RQ) 1.

To examine whether the HTKS and academic performance association varied as a function of the different effect size characteristics, we performed statistical significance testing with three multivariate meta-analytic regression models. These findings are displayed in Table 3 and discussed under RQ 2 and RQ 3.

## Relations between the HTKS and academic performance

The first research question focused on the overall association between the HTKS and children's academic performance, and evidence of variability in this association. Table 2 presents the weighted summary effect (Fisher *z*) and 95% confidence intervals for the association between the HTKS and children's academic performance across all academic measures and for the two additional intercept-only models (i.e., math only and language arts only). The weighted summary effect when effect sizes representing all measures were included ( $n = 413$ ,  $k = 69$ ) in the model was  $z = 0.4$  with a 95% confidence interval ranging between 0.37 and 0.44. Because the 95% confidence interval is relatively narrow for this effect and does not include 0, these findings can be interpreted as a medium summary effect that is positive, precise, and highly significant ( $p < 0.001$ ). As evident in Table 2, the average effect for the math only measures ( $n = 138$ ,  $k = 44$ ) was slightly higher

TABLE 3 Moderator analysis.

Variables	Model 1	Model 2	Model 3
Study quality	-0.0026 (0.0112)		
Age between		-0.0001 (0.0019)	
Academic domain between (math = ref coded as 0)		** -0.2089 (0.0722)	
<b>Country (US = ref category coded as 0)</b>			
Asia		-0.0700 (0.0434)	
Europe		-0.0149 (0.0442)	
Other		-0.0555 (0.0505)	
HTKS testing occasion between (same time as academics = ref coded as 0)		-0.0432 (0.0544)	
Age within			0.0039 (0.0026)
Academic domain within (math = ref coded as 0)			*** -0.1288 (0.0139)
HTKS testing occasion within			-0.0272 (0.0135)
Number of effect sizes	413	406	406
Number of studies	69	63	63

Standard error in parenthesis.

\*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ .

( $z = 0.49$ ). For the language arts only measures ( $n = 268$ ,  $k = 61$ ), the average effect was slightly lower ( $z = 0.35$ ) than the summary effect for both the math only measures and when effect sizes representing all measures were combined. Nevertheless, each of these separate analyses maintained the significance level ( $p < 0.001$ ) despite this sample size reduction. Although it is impossible to eliminate all publication bias in a meta-analysis, Edger's test showed no statistically significant small study effects among this particular pool of studies. As illustrated in Table 2, the tau-square value for the summary effects of all three analyses suggests the effect size variability in this meta-analysis stems from true variability (i.e., more than just error).

The forest plots in Figures 2, 3 provide an illustration of how each study's effect size relates to the overall summary effect. Figure 2 includes studies with effect sizes representing the HTKS and language arts association, whereas Figure 3 includes studies with effect sizes representing the HTKS and mathematics association. As illustrated in the plots, on average, effect sizes representing the HTKS and mathematics association were noticeably larger than effect sizes representing the HTKS and language arts association. For readability purposes, studies with multiple effect sizes are represented by the mean of those effect sizes in the forest plots. A forest plot with all 413 effect sizes (robumeta plot) is available in the Supplementary material.

### Sources of variability between and within studies

Research questions 2 and 3 focused on examining sources of variability in the effects between and within studies. Traditionally, the goal of a moderator analysis is to try to provide explanations for variations in effect sizes across the studies of a meta-analysis (Borenstein et al., 2009). However, effect sizes can vary both between and within the studies themselves. Thus, RQ 2 focused on examining between-study variation whereas RQ 3 focused on

examining within-study variation. In all of the following meta-regression models, the effect size representing the HTKS and academic performance association serves as the outcome variable, and the moderators serve as the predictor variables. We used study means to estimate between-study effects of moderators (RQ 2) and study mean-centered values of moderators to estimate the within-study effects (RQ 3). Prior to running these models, we regressed the effect size variable on study quality to ensure the quality of a study would not influence the results. Based on the non-significant  $p$ -value of 0.82 and small regression coefficient of  $-0.0026$ , we concluded that study quality was an unlikely moderating factor between studies. Thus, it was not included in the subsequent multivariate regression models.

### Variability in effects between studies

Research question 2 focused on whether between-study differences in participant and measurement characteristics influence the magnitude of the HTKS and academic performance effect. To explore heterogeneity in effects by between-study participant and measurement characteristics, we regressed the effect size variable (each study's Fisher  $z$ ) on country (US = Ref coded as 0), age (continuous variable represented by the average age of the participants in months), academic domain (math = Ref coded as 0) and testing occasion (same as academic testing = Ref coded as 0). As evident in Table 3, when controlling for children's age, country, and testing occasion, on average, academic measures that were math related had statistically significantly stronger effect sizes than those that were language arts related between studies ( $p = 0.0078$ ,  $b = -0.21$ ). All the other variables in the model yielded a  $p$ -value  $> 0.05$ .

### Variability in effects within studies

Research question 3 examined whether within-study differences in participant and measurement characteristics influence the magnitude of the HTKS and academic performance

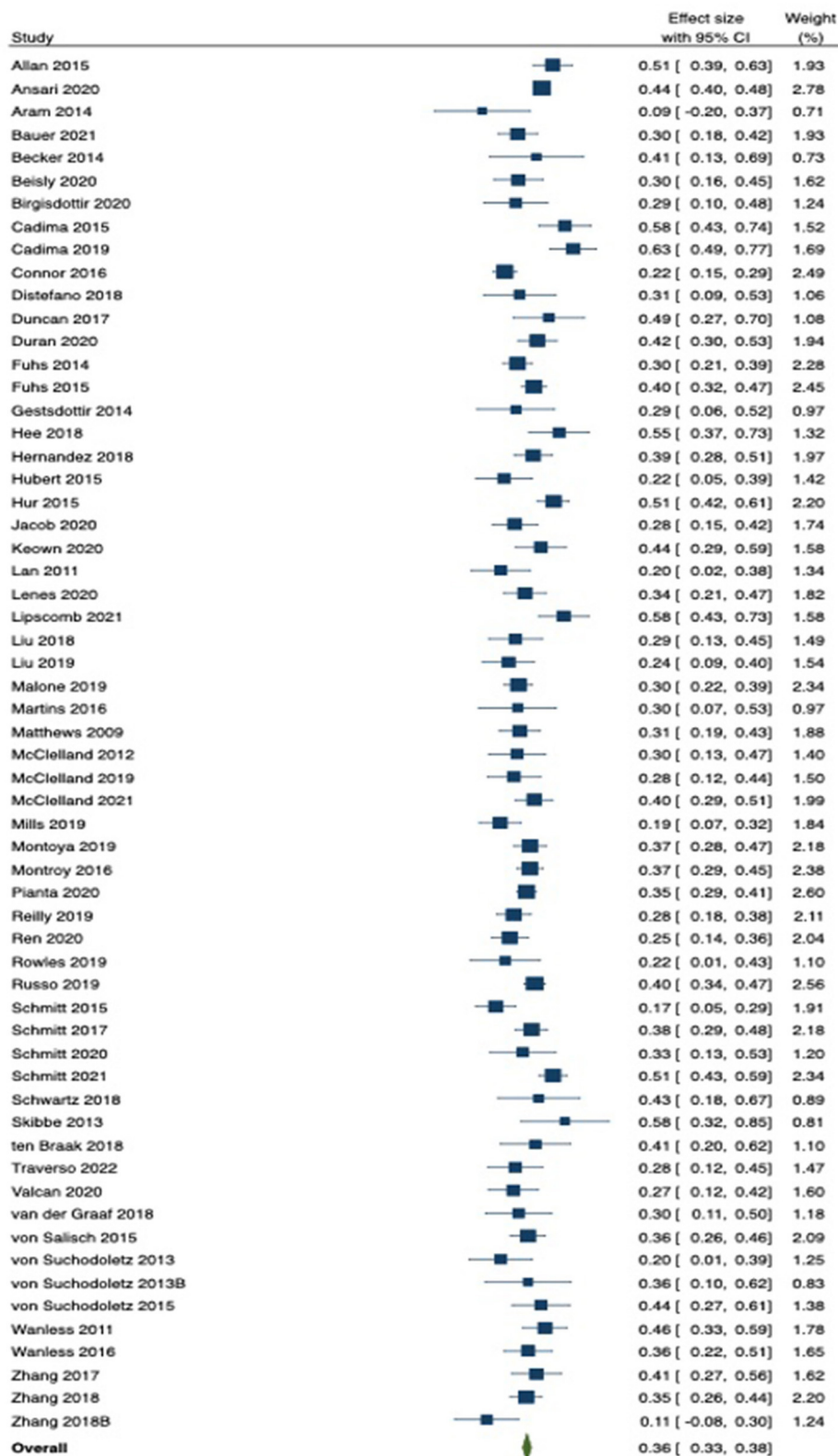


FIGURE 2 Forest plot for the HTKS and language arts association.

effect. To explore heterogeneity in effects by within-study participant and measurement characteristics, we regressed the effect size variable (each study's Fisher  $z$ ) on age (continuous

variable represented by the average age of the participants in months), academic domain (math = Ref coded as 0), and testing occasion (same as academic testing = Ref coded as

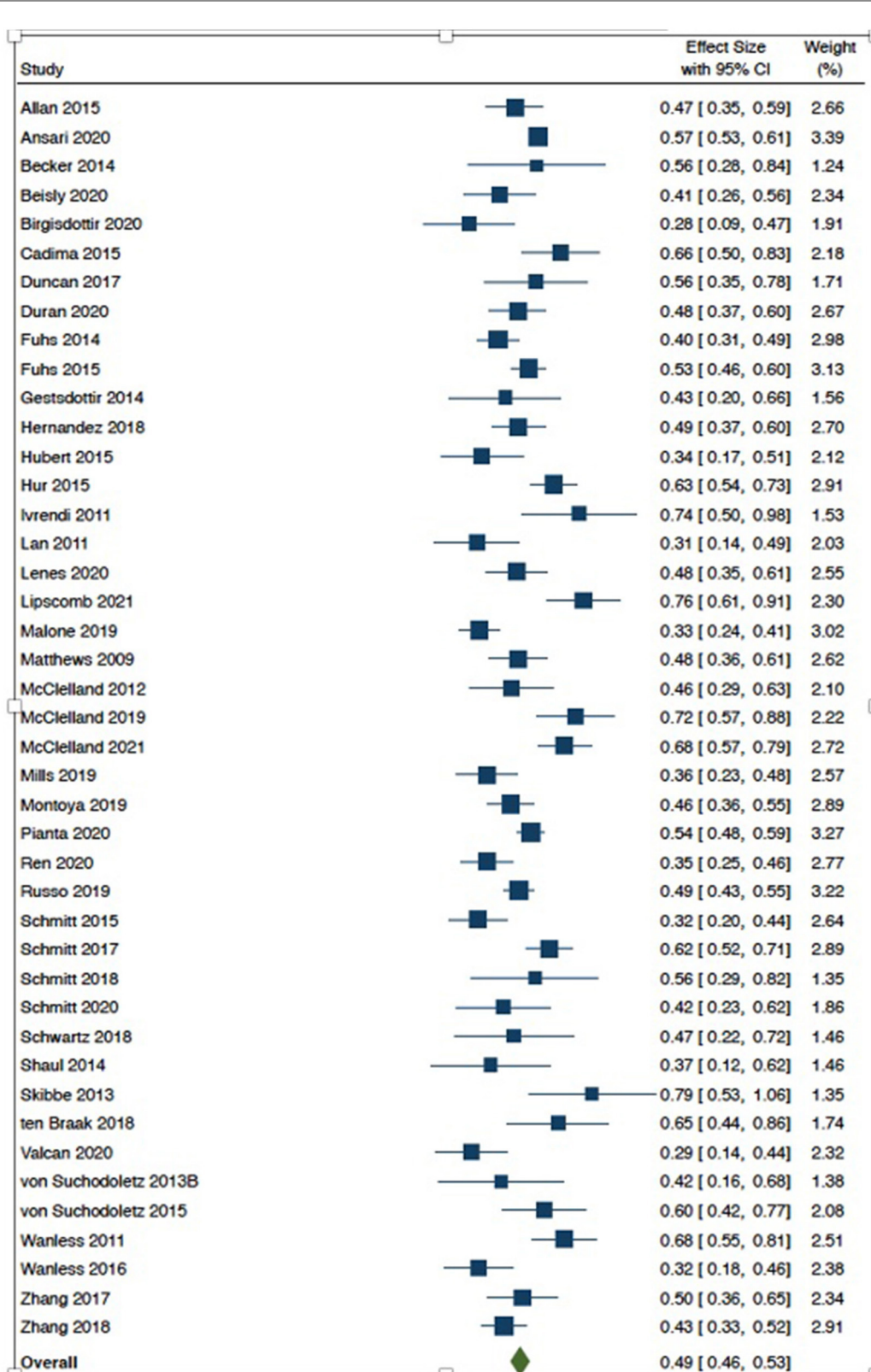


FIGURE 3 Forest plot for the HTKS and mathematics association.

0). Like RQ 2, when controlling for children’s age and testing occasion, on average, academic measures that were math-related had significantly stronger effect sizes than those that

were language arts related when compared within studies ( $p < 0.001$ ,  $b = -0.13$ ); the other variables in the model yielded a  $p$ -value  $> 0.05$ .

## Discussion

This study found evidence that the HTKS predicts academic outcomes targeted in early learning contexts. When we analyzed the correlations between children's HTKS performance and their academic achievement across 413 effect sizes from 69 studies, the HTKS demonstrated significant and relatively consistent positive associations with children's current and subsequent performance across math, literacy, and oral language outcomes. After conducting the primary meta-analysis with all the measures combined (413 effect sizes), we performed two separate exploratory meta-analyses across the math related and language arts related measures. Like the meta-analysis with all the academic measures combined, the summary effects of these different subject domains, although strongest in math, remained positive and statistically significant when they were meta-analyzed separately. Taken together, these findings confirm prior research suggesting that self-regulation, including the previously mentioned EF components, make both domain-general and domain-specific contributions to children's academic performance, with associations generally strongest for mathematics (Dent, 2013; Allan et al., 2014; Fuhs et al., 2014; McClelland et al., 2014; Schmitt et al., 2017; Cortés Pascual et al., 2019; Birgisdottir et al., 2020; Robson et al., 2020).

Our findings per the moderator analysis demonstrated that except for the subject domain variable, the HTKS and academic association was relatively stable across different participant, contextual, and measurement factors. In other words, although on average, the HTKS was most predictive of children's math performance, no statistically significant effect size differences emerged across children's age, country, and whether the HTKS was given before or concurrent with achievement measures. Considering the widespread international use of the HTKS in research, along with the potential it holds for educational practice, this study makes a valuable contribution to the existing self-regulation literature. We propose that the findings have two critical implications: (a) they provide additional insights into some of the predictive properties of a widely-used self-regulation measure and (b) they reinforce existing research demonstrating the importance of self-regulation for early academic achievement.

## Predictive properties of the HTKS

A chief concern in the measurement of self-regulation is the need to identify instruments feasible for school settings that predict outcomes deemed meaningful in school (Lipsey et al., 2017; McCoy, 2019). One outcome deemed meaningful in school is children's academic performance. The results of this meta-analysis found the HTKS to demonstrate associations with children's academic performance that were comparable to and in some cases, slightly stronger than, the associations determined by meta-analyses using multiple measures of self-regulation and related constructs (Dent, 2013; Allan et al., 2014; Jacob and Parkinson, 2015; Cortés Pascual et al., 2019; Robson et al., 2020; Spiegel et al., 2021). This finding offers a valuable extension of Lipsey et al.'s (2017) study that

identified the HTKS as one of the most robust performance-based, self-regulation measures for predicting children's academic achievement across the preschool and kindergarten years. One factor distinguishing the present study from Lipsey et al.'s work is that given the present study's meta-analytic methods, we were able to shed light on some of the conditions in which the HTKS performs best regarding its ability to predict children's academic performance across and within the studies via a moderator analysis. This knowledge is critical for researchers and practitioners seeking assessments with predictive properties that are reliable and valid in various contexts and populations.

The present study's moderator analysis suggested that the HTKS and academic association was relatively stable across different participant and measurement factors. In other words, although on average, the HTKS was more strongly associated with children's math performance, effect sizes remained positive and significant across the language and literacy domain as well. Furthermore, no statistically significant effect size differences emerged across children's age, country, and whether the HTKS was given before or concurrent with achievement measures. Below, we propose explanations related to the results of the moderator analysis with a focus on each of those four variables.

### Country

One characteristic distinguishing this meta-analysis from meta-analyses examining the self-regulation and academic performance effect across multiple measures is that we examined how differences in participants' country impacted the association. The fact that the HTKS is available in 28 languages and has been used cross-culturally made this possible. The present study did not find the HTKS and academic performance effect to be moderated by the different countries represented. This finding is understandable given that researchers have consistently found evidence for both the stability and variability of the association between children's HTKS performance and their academic achievement across cultures (Lan et al., 2011; Wanless et al., 2011b; von Suchodoletz et al., 2013; Gestsdottir et al., 2014). The present study extends that research by demonstrating that when these studies are synthesized and analyzed meta-analytically, there is more evidence for the stability in children's need for self-regulation to succeed academically cross-culturally. One explanation for our findings is that the HTKS does not depend on teachers' appraisal of children's behaviors, which may explain why it appeared to capture the self-regulation and academic association relatively the same, regardless of the children's country. The HTKS may have yielded similar associations with children's academic performance across countries because it may be tapping into an aspect of the construct that is less context- or reporter-specific and more dependent on child-level factors.

### Age

Although some research has found that age moderates the association between children's self-regulation and academic performance (Dent, 2013), our findings suggest that the HTKS predicts children's academic performance, regardless of the sample's

average age. However, it is important to note that this lack of statistical significance may be due to a lack of variability in age, rather than the absence of a moderating effect. For example, despite the wide age range represented (ages 3 through 8) in our inclusion criteria, most of the effect sizes (~95%) were extracted from studies where the average age of the sample was at least 48 months (4 years old) at the time of HTKS testing. Because some researchers have found the HTKS to capture less variability in children with nascent self-regulation skills (Gonzales et al., 2021; McClelland et al., 2021), it is plausible to suspect that the HTKS and academic association may have been weaker if we had more samples representative of younger children (e.g., average age younger than 4 years old).

### Testing occasion

While many researchers have found children's self-regulation to be associated with their academic performance, aspects of this relationship remain unclear even when controlling for various factors. One remaining question relates to whether the association is truly predictive, and if so, what assessments are most suitable for predicting children's academic performance (Lipsey et al., 2017). As recently mentioned, in a longitudinal study, Lipsey et al. (2017) ranked the HTKS as one of the most robust self-regulation measures for predicting children's academic performance across the preschool and kindergarten years. The fact that the present study found cross-sectional effect sizes, on average, not to statistically significantly differ from effect sizes in which the HTKS was administered prior to academic testing provides some corroborating support for Lipsey et al.'s conclusion. However, future research should investigate how unexplored factors may contribute to the stability of this effect. One possibility is that the time-lapse between the HTKS and academic data collection in the present meta-analysis was too narrow to represent effect size variability between cross-sectional and multi-waved studies. For example, most studies with multiple waves included fall and spring testing only. Thus, the results of this meta-analysis alone do not provide strong enough evidence to draw firm conclusions about whether the strength of HTKS and academic association varies by testing occasion, nor whether the association is truly predictive.

### Academic domain

The present study's results indicated that, on average, the association between children's HTKS performance and their mathematics achievement was statistically significantly stronger than the association between children's HTKS performance and their language arts achievement. Although the exact mechanisms behind this relationship are unclear, some researchers have proposed that the self-regulation and academic achievement association is strongest in mathematics because mathematical tasks are the least familiar and more difficult to automate for most children (Blair et al., 2008; Fuhs et al., 2014; McClelland and Cameron, 2019; Spiegel et al., 2021). However, other plausible explanations should be considered. For example, it is possible that the mathematics measures in this meta-analysis had more sensitivity than the language and literacy measures. We are unable

to provide a firm explanation as to why the HTKS and academic association was strongest for mathematics, as investigating the mechanisms behind this relationship was beyond our study's scope. Nevertheless, the finding that the self-regulation and academic achievement association was strongest in mathematics is consistent with prior research (Dent, 2013; Allan et al., 2014; Fuhs et al., 2014; McClelland et al., 2014; Schmitt et al., 2017; Cortés Pascual et al., 2019; Birgisdottir et al., 2020; Robson et al., 2020).

## The importance of self-regulation for early academic achievement

In addition to providing deeper insights into the predictive utility of the HTKS, this meta-analysis also sheds light on the self-regulation and academic association at the construct level. To our knowledge, no researchers have performed a meta-analysis focused on examining the self-regulation and academic association with a single measure. However, several researchers have synthesized this relationship on the construct level; these meta-analyses included multiple measures of self-regulation and related constructs, such as EF (Dent, 2013; Allan et al., 2014; Jacob and Parkinson, 2015; Cortés Pascual et al., 2019; Robson et al., 2020; Spiegel et al., 2021). The results of this meta-analysis found the HTKS to demonstrate associations with children's academic performance that were comparable to and in some cases, slightly stronger than these prior meta-analyses. Regardless of the particular measures, researchers are consistently finding evidence that self-regulation is a critical skill that children need for successful school participation. The significant, positive self-regulation and academic performance associations found across all the subject domains in the present study (coupled with the findings across these prior meta-analyses) confirm social and emotional learning advocacy efforts encouraging schools and classrooms to make teaching self-regulation to children a priority. In other words, if self-regulation is a consistent predictor of children's academic performance, to maximize all students' learning, schools may need to consider going beyond just teaching academics.

There is theoretical and empirical evidence showing that children's self-regulation and EF skills can improve over time in response to positive environmental supports and strategies (Blair, 2002; McClelland et al., 2010, 2015; Moffitt et al., 2011; Blair and Raver, 2015; Howard and Williams, 2018). Given the malleability of self-regulation (Blair and Raver, 2012; Morrison and Grammer, 2016) and the stability of the associations found in the present study, early childhood educators and practitioners should consider focusing more on promoting children's self-regulation and their related EF skills as one way to support their social-emotional growth and academic learning. This need has become increasingly pressing with school shutdowns and the stress and trauma children and families have been experiencing during the COVID-19 pandemic. A self-regulation assessment that is predictive of academic outcomes across differing cultural contexts can help researchers and practitioners monitor children's response to such interventions and help to inform school-based practices that can facilitate optimal school participation for diverse populations of children.

## Limitations and directions for future researchers

Several limitations bear mentioning. First, while 69 studies are considered sufficient for RVE, a larger sample would have allowed for a more expansive moderator analysis. Because we were limited to the number of moderators, we had to prioritize which variables made the most sense to examine, both methodologically and theoretically. For example, we decided not to include gender as a moderator because, during data cleaning, the descriptive analysis showed that the proportion of males to females across studies showed little variation across samples. In other words, the gender composition was split relatively equally between males and females in most studies. Like gender, we were unable to examine race/ethnicity or socioeconomic status because not enough studies reported this information. There is some evidence to suspect these factors to influence the association between self-regulation and academic performance (Blair and Raver, 2012; Cadima et al., 2015; Lawson et al., 2018).

An unexpected finding of this meta-analysis is that the HTKS has been utilized mostly with typically developing samples of children. The few studies that did target clinical samples of children included those with learning, language, or behavioral challenges (e.g., Graziano et al., 2016). Despite the significance of this topic, combining clinical samples with samples of children that are predominantly typically developing could produce too much unexplainable heterogeneity in the findings (Wilson, 2019); therefore, we excluded all clinical samples from this meta-analysis. A future approach could be to conduct a meta-analysis examining the self-regulation and academic performance association in studies focused exclusively on children with learning or behavioral challenges.

An additional limitation of this study is that it was not possible to establish causation. In other words, because most of the research is cross-sectional and non-experimental, we cannot confirm that the HTKS is truly predictive of children's academic achievement. For instance, the HTKS and academic performance association may exist because a third variable, such as processing speed or access to developmentally enriching experiences, is either enhancing or compromising children's behavioral regulation and academic skills concomitantly. Therefore, more research is needed to fully understand the validity and utility of the HTKS in practice settings. On a similar note, it is important to consider how the coding of the variables could have played a role in the results. For example, like most meta-analyses, the current study's age variable was represented by the mean age of the sample. One short-coming of this approach is that it doesn't take into account age range variation of the selected studies. Future research should examine whether age range variation both within and between studies influences the association between the HTKS and children's academic performance. Several of the studies took place in preschool settings where children's ages ranged from ~36 to 60 months. Considering the rapid development of children's self-regulation skills during the preschool years, this relatively wide range could have inflated the HTKS and academic association in those particular studies.

Finally, another limitation of this study is that we did not include the gray literature, a decision influenced much by constraints related to time and resources. This practice could increase the risk of publication bias in the findings of this meta-analysis. However, there are advantages associated with excluding gray literature that should also be considered. First, searching and retrieving unpublished studies is more resource intensive and involves methods that are often less reliable and reproducible than traditional database searching (Adams et al., 2016; Hartling et al., 2017). In addition, though encouraged by some meta-analysis guidelines and standards, there is limited empirical support that including unpublished studies and dissertations in a meta-analysis alters the results significantly (Hartling et al., 2017).

We conclude that despite the limitations above, this study has several strengths. The work represents the first meta-analysis of the predictive properties of the HTKS. Thus, it extends knowledge about the measurement characteristics of an instrument that researchers rely on frequently to make claims about children's self-regulation. The findings also provide further evidence that when self-regulation is assessed with a valid and reliable measure, children's self-regulation skills contribute significantly and consistently to educational outcomes. Finally, a notable strength of this work lies in the transparency of the methods. Future researchers could, therefore, easily extend or replicate the research to address some of the shortcomings mentioned.

## Data availability statement

The original contributions presented in the study are included in the article/[Supplementary material](#). Further inquiries can be directed to the corresponding author.

## Author contributions

SK performed the statistical analysis, developed the codebooks, and wrote the first draft of the manuscript. SK and CC contributed to the conception of the study and wrote the second draft. SK, CC, and MM contributed to the design. SK and JK managed/organized the data and conducted the quality appraisal of the studies. AA contributed to the presentation of results. AA, JK, and PB helped with the abstract screen/full-text review and data extraction/coding. MM contributed to the literature review of the second draft. All authors contributed to manuscript revision, read, and approved the submitted version.

## Funding

This study received partial funding through an internal award to CC from the University at Buffalo.

## Acknowledgments

We thank Michael Hebert, David Wilson, John Strong, Tiffany Karalis Noel, Ryan Krone, and Julia Rose Kenny for sharing their time and expertise to help inform this study.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

## References

- Adams, J., Hillier-Brown, F. C., Moore, H. J., Lake, A. A., Araujo-Soares, V., White, M., et al. (2016). Searching and synthesising 'grey literature' and 'grey information' in public health: Critical reflections on three case studies. *Systemat. Rev.* 5, 1–11. doi: 10.1186/s13643-016-0337-y
- Allan, N. P., Hume, L. E., Allan, D. M., Farrington, A. L., and Lonigan, C. J. (2014). Relations between inhibitory control and the development of academic skills in preschool and kindergarten: A meta-analysis. *Dev. Psychol.* 50, 2368–2379. doi: 10.1037/a0037493
- Best, J. R., and Miller, P. H. (2010). A developmental perspective on executive function. *Child Dev.* 81, 1641–1660. doi: 10.1111/j.1467-8624.2010.01499.x
- Birgisdottir, F., Gestsdottir, S., and Geldhof, G. J. (2020). Early predictors of first and fourth grade reading and math: The role of self-regulation and early literacy skills. *Early Childh. Res. Quart.* 53, 507–519. doi: 10.1016/j.ecresq.2020.05.001
- Blair, C. (2002). School readiness. Integrating cognition and emotion in a neurobiological conceptualization of children's functioning at school entry. *Am. Psychol.* 57, 111–127. doi: 10.1037/0003-066X.57.2.111
- Blair, C. (2016). "The development of executive functions and self-regulation: A bidirectional psychobiological model," in *Handbook of Self-Regulation, 3rd Edn*, eds V. D. Vohs and R. F. Baumeister (New York City, NY: The Guilford Press), 417–439.
- Blair, C., Knipe, H., and Gamson, D. (2008). Is there a role for executive functions in the development of mathematics ability? *Mind Brain Educ.* 2, 80–89. doi: 10.1111/j.1751-228X.2008.00036.x
- Blair, C., and Raver, C. C. (2012). Individual development and evolution: Experiential canalization of self-regulation. *Dev. Psychol.* 48, 647–657. doi: 10.1037/a0026472
- Blair, C., and Raver, C. C. (2015). School readiness and self-regulation: A developmental psychobiological approach. *Ann. Rev. Psychol.* 66, 711–731. doi: 10.1146/annurev-psych-010814-015221
- Blair, C., and Raver, C. C. (2016). Poverty, stress, and brain development: New directions for prevention and intervention. *Acad. Pediatr.* 16, S30–S36. doi: 10.1016/j.acap.2016.01.010
- Blair, C., Ursache, A., Greenberg, M., and Vernon-Feagans, L. (2015). Multiple aspects of self-regulation uniquely predict mathematics but not letter-word knowledge in the early elementary grades. *Develop. Psychol.* 51, 459–472. doi: 10.1037/a0038813
- Booth, A., Hennessy, E., and Doyle, O. (2018). Self-regulation: learning across disciplines. *J. Child Fam. Stud.* 27, 3767–3781. doi: 10.1007/s10826-018-1202-5
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., and Rothstein, H. R. (2009). *Introduction to Meta-analysis*. Hoboken, NJ: John Wiley & Sons. doi: 10.1002/9780470743386
- Cadima, J., Gamelas, A. M., McClelland, M., and Peixoto, C. (2015). Associations between early family risk, children's behavioral regulation, and academic achievement in Portugal. *Early Educ. Develop.* 26, 708–728. doi: 10.1080/10409289.2015.1005729
- Cameron Ponitz, C. E., McClelland, M. M., Jewkes, A. M., Connor, C. M., Farris, C. L., and Morrison, F. J. (2008). Touch your toes! Developing a direct measure of behavioral regulation in early childhood. *Early Childh. Res. Quart.* 23, 141–158. doi: 10.1016/j.ecresq.2007.01.004
- Cameron Ponitz, C. E., McClelland, M. M., Matthews, J. S., and Morrison, F. J. (2009). A structured observation of behavioral self-regulation and its contribution to kindergarten outcomes. *Develop. Psychol.* 45, 605–619. doi: 10.1037/a0015365
- Cameron, C. E. (2018). *Hands on, Minds on: How Executive Function, Motor, and Spatial Skills Foster School Readiness*. New York, NY: Teachers College Press.
- Cameron, C. E., Brock, L. L., Murrah, W. M., Bell, L. H., Worzalla, S. L., Grissmer, D., et al. (2012). Fine motor skills and executive function both contribute to kindergarten achievement. *Child Develop.* 83, 1229–1244. doi: 10.1111/j.1467-8624.2012.01768.x
- Cantin, R. H., Gnaedinger, E. K., Gallaway, K. C., Hesson-McInnis, M. S., and Hund, A. M. (2016). Executive functioning predicts reading, mathematics, and theory of mind during the elementary years. *J. Exp. Child Psychol.* 146, 66–78. doi: 10.1016/j.jecp.2016.01.014
- Carr, M. E., Moore, D. W., and Anderson, A. (2014). Self-management interventions on students with autism: A meta-analysis of single-subject research. *Except. Child.* 81, 28–44. doi: 10.1177/0014402914532235
- Cerino, E. S., Hooker, K., Stawski, R. S., and McClelland, M. (2019). A new brief measure of executive function: Adapting the head-toes-knees-shoulders task to older adults. *Gerontologist* 59, e258–e267. doi: 10.1093/geront/gny028
- Cheung, M. W. L. (2019). A guide to conducting a meta-analysis with non-independent effect sizes. *Neuropsychol. Rev.* 29, 387–396. doi: 10.1007/s11065-019-09415-6
- Collaborative for Academic, Social and Emotional Learning. *What Is the CASEL Framework?* (n.d.) Available online at: <https://www.casel.org/core-competencies/> (accessed December 15, 2021).
- Cortés Pascual, A., Moyano Muñoz, N., and Quílez Robres, A. (2019). The relationship between executive functions and academic performance in primary education: Review and Meta-analysis. *Front. Psychol.* 10, 1582. doi: 10.3389/fpsyg.2019.01582
- Dent, A. L. (2013). *The Relation Between Self-regulation and Academic Achievement: A Meta-analysis Exploring Variation in the Way Constructs Are Labeled, Defined, and Measured* (Doctoral dissertation), Duke University, Durham, NC, United States. DUKESPACE. Available online at: <https://dukespace.lib.duke.edu/dspace/handle/10161/7265> (accessed May 31, 2023).
- Doebel, S., and Lillard, A. S. (2023). How does play foster development? A new executive function perspective. *Dev. Rev.* 67, 101064. doi: 10.1016/j.dr.2022.101064
- Edossa, A. K., Schroeders, U., Weinert, S., and Artelt, C. (2018). The development of emotional and behavioral self-regulation and their effects on academic achievement in childhood. *Int. J. Behav. Dev.* 42, 192–202. doi: 10.1177/0165025416687412
- Feinberg, A. B., and Shapiro, E. S. (2003). Accuracy of teacher judgments in predicting oral reading fluency. *Sch. Psychol. Quart.* 18, 52–65. doi: 10.1521/scpq.18.1.52.20876
- Finders, J. K., McClelland, M. M., Geldhof, G. J., Rothwell, D. W., and Hatfield, B. E. (2021). Explaining achievement gaps in kindergarten and third grade: The role of self-regulation and executive function skills. *Early Childh. Res. Quart.* 54, 72–85. doi: 10.1016/j.ecresq.2020.07.008
- Flook, L., Goldberg, S. B., Pinger, L., and Davidson, R. J. (2015). Promoting prosocial behavior and self-regulatory skills in preschool children through a mindfulness-based kindness curriculum. *Dev. Psychol.* 51, 44–51. doi: 10.1037/a0038256
- Floyer-Lea, A., and Matthews, P. M. (2005). Distinguishable brain activation networks for short- and long-term motor skill learning. *J. Neurophysiol.* 94, 512–518. doi: 10.1152/jn.00717.2004
- Fuhs, M. W., Farran, D. C., and Nesbitt, K. T. (2015). Prekindergarten children's executive functioning skills and achievement gains: The utility of direct assessments and teacher ratings. *J. Educ. Psychol.* 107, 207–221. doi: 10.1037/a0037366
- Fuhs, M. W., Nesbitt, K. T., Farran, D. C., and Dong, N. (2014). Longitudinal associations between executive functioning and academic skills across content areas. *Dev. Psychol.* 50, 1698–1709. doi: 10.1037/a0036633
- Gestsdottir, S., von Suchodoletz, A., Wanless, S. B., Hubert, B., Guimard, P., Birgisdottir, F., et al. (2014). Early behavioral self-regulation, academic achievement,



- and gender: Longitudinal findings from France, Germany, and Iceland. *Appl. Dev. Sci.* 18, 90–109. doi: 10.1080/10888691.2014.894870
- Gonzales, C. R., Bowles, R., Geldhof, G. J., Cameron, C. E., Tracy, A., and McClelland, M. M. (2021). The Head-Toes-Knees-Shoulders Revised (HTKS-R): Development and psychometric properties of a revision to reduce floor effects. *Early Childh. Res. Quart.* 56, 320–332. doi: 10.1016/j.ecresq.2021.03.008
- Graziano, P. A., Garb, L. R., Ros, R., Hart, K., and Garcia, A. (2016). Executive functioning and school readiness among preschoolers with externalizing problems: the moderating role of the student–teacher relationship. *Early Educ. Dev.* 27, 573–589. doi: 10.1080/10409289.2016.1102019
- Hahn-Markowitz, J., Berger, I., Manor, I., and Maeir, A. (2020). Efficacy of Cognitive-Functional (Cog-Fun) occupational therapy intervention among children with ADHD: An RCT. *J. Attent. Disord.* 24, 655–666. doi: 10.1177/1087054716666955
- Halle, T. G., and Darling-Churchill, K. E. (2016). Review of measures of social and emotional development. *J. Appl. Dev. Psychol.* 45, 8–18. doi: 10.1016/j.appdev.2016.02.003
- Hartling, L., Featherstone, R., Nuspl, M., Shave, K., Dryden, D. M., and Vandermeer, B. (2017). Grey literature in systematic reviews: A cross-sectional study of the contribution of non-English reports, unpublished studies and dissertations to the results of meta-analyses in child-relevant reviews. *BMC Med. Res. Methodol.* 17, 1–11. doi: 10.1186/s12874-017-0347-z
- Hedges, L. V., Tipton, E., and Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Res. Synth. Methods* 1, 39–65. doi: 10.1002/jrsm.5
- Hee, P. J., Xu, Y., and Krieg, A. (2018). Validation of the head–toes–knees–shoulders task in Native Hawaiian and non-Hawaiian children. *Early Childh. Res. Quart.* 44, 192–205. doi: 10.1016/j.ecresq.2017.12.007
- Hernández, M. M., Eisenberg, N., Valiente, C., Spinrad, T. L., Johns, S. K., Berger, R. H., et al. (2018). Self-regulation and academic measures across the early elementary school grades: Examining longitudinal and bidirectional associations. *Early Educ. Dev.* 29, 914–938. doi: 10.1080/10409289.2018.1496722
- Hoge, R. D., and Coladarci, T. (1989). Teacher-based judgments of academic achievement: A review of literature. *Rev. Educ. Res.* 59, 297–313. doi: 10.3102/00346543059003297
- Howard, S. J., and Williams, K. E. (2018). Early self-regulation, early self-regulatory change, and their longitudinal relations to adolescents' academic, health, and mental well-being outcomes. *J. Dev. Behav. Pediatr.* 39, 489–496. doi: 10.1097/DBP.0000000000000578
- Jacob, R., and Parkinson, J. (2015). The potential for school-based interventions that target executive function to improve academic achievement: A review. *Rev. Educ. Res.* 85, 512–552. doi: 10.3102/0034654314561338
- Korinek, L., and deFur, S. H. (2016). Supporting student self-regulation to access the general education curriculum. *Teach. Except. Child.* 48, 232–242. doi: 10.1177/0040059915626134
- Kugley, S., Wade, A., Thomas, J., Mahood, Q., Jørgensen, A. K., Hammerstrøm, K., et al. (2017). Searching for studies: A guide to information retrieval for Campbell systematic reviews. *Campbell Systemat. Rev.* 13, 1–73. doi: 10.4073/cmg.2016.1
- Lan, X., Legare, C. H., Cameron Ponitz, C. E., Li, S., and Morrison, F. J. (2011). Investigating the links between the subcomponents of executive function and academic achievement: A cross-cultural analysis of Chinese and American preschoolers. *J. Exp. Child Psychol.* 108, 677–692. doi: 10.1016/j.jecp.2010.11.001
- Lawson, G. M., Hook, C. J., and Farah, M. J. (2018). A meta-analysis of the relationship between socioeconomic status and executive function performance among children. *Dev. Sci.* 21, e12529. doi: 10.1111/desc.12529
- Lenes, R., McClelland, M. M., ten Braak, D., Idsøe, T., and Størksen, I. (2020). Direct and indirect pathways from children's early self-regulation to academic achievement in fifth grade in Norway. *Early Childh. Res. Quart.* 53, 612–624. doi: 10.1016/j.ecresq.2020.07.005
- Lipsey, M. W., Nesbitt, K. T., Farran, D. C., Dong, N., Fuhs, M. W., and Wilson, S. J. (2017). Learning-related cognitive self-regulation measures for prekindergarten children: A comparative evaluation of the educational relevance of selected measures. *J. Educ. Psychol.* 109, 1084–1102. doi: 10.1037/edu0000203
- Liu, Y., Sun, H., Lin, D., Li, H., Yeung, S. S.-S., and Wong, T. T.-Y. (2018). The unique role of executive function skills in predicting Hong Kong kindergarteners' reading comprehension. *Br. J. Educ. Psychol.* 88, 628–644. doi: 10.1111/bjep.12207
- Lonigan, C. J., and Milburn, T. F. (2017). Identifying the dimensionality of oral language skills of children with typical development in preschool through fifth grade. *J. Speech Lang. Hear. Res.* 60, 2185–2198. doi: 10.1044/2017\_JSLHR-L-15-0402
- Lonigan, C. J., Spiegel, J. A., Goodrich, J. M., Morris, B. M., Osborne, C. M., Lerner, M. D., et al. (2017). Does preschool self-regulation predict later behavior problems in general or specific problem behaviors? *J. Abnorm. Child Psychol.* 45, 1491–1502. doi: 10.1007/s10802-016-0260-7
- Mack, E., Gnass, J., Voock, M., and Preckel, F. (2023). The domain-specificity of elementary school teachers' judgment accuracy. *Contempor. Educ. Psychol.* 72, 102142. doi: 10.1016/j.cedpsych.2022.102142
- McCabe, L. A., Cunnington, M., and Brooks-Gunn, J. (2004). "The development of self-regulation in young children: Individual characteristics and environmental contexts," in *Handbook of Self-regulation: Research, Theory, and Applications, Vol. 574*, eds R. F. Baumeister (New York, NY: The Guilford Press), 340–356.
- McCabe, L. A., Hernandez, M., Lara, S. L., and Brooks-Gunn, J. (2000). Assessing preschoolers' self-regulation in homes and classrooms: Lessons from the field. *Behav. Disord.* 26, 53–69. doi: 10.1177/019874290002600106
- McClelland, M. M., Cameron Ponitz, C. E., Messersmith, E. E., and Tominey, S. (2010). "Self-regulation: Integration of cognition and emotion," in *The handbook of life-span development, Vol. 1*, eds W. F. Overton and R. M. Lerner (Hoboken, NJ: John Wiley & Sons), 15. doi: 10.1002/9780470880166.hlsd001015
- McClelland, M. M., and Cameron, C. E. (2012). Self-regulation in early childhood: Improving conceptual clarity and developing ecologically valid measures. *Child Dev. Perspect.* 6, 136–142. doi: 10.1111/j.1550-8606.2011.00191.x
- McClelland, M. M., and Cameron, C. E. (2019). Developing together: The role of executive function and motor skills in children's early academic lives. *Early Childh. Res. Quart.* 46, 142–151. doi: 10.1016/j.ecresq.2018.03.014
- McClelland, M. M., Cameron, C. E., Connor, C. M., Farris, C. L., Jewkes, A. M., and Morrison, F. J. (2007). Links between behavioral regulation and preschoolers' literacy, vocabulary, and math skills. *Dev. Psychol.* 43, 947–959. doi: 10.1037/0012-1649.43.4.947
- McClelland, M. M., Cameron, C. E., Duncan, R., Bowles, R. P., Acock, A. C., Miao, A., et al. (2014). Predictors of early growth in academic achievement: The head-toes-knees-shoulders task. *Front. Psychol.* 5, 599. doi: 10.3389/fpsyg.2014.00599
- McClelland, M. M., Geldhof, J. G., Cameron, C. E., and Wanless, S. B. (2015). "Development and self-regulation," in *Handbook of Child Psychology and Developmental Science, 7th Edn*, eds M. E. Lamb and R. M. Lerner (Hoboken, NJ: John Wiley & Sons), 523–564. doi: 10.1002/9781118963418.childpsy114
- McClelland, M. M., Gonzales, C. R., Cameron, C. E., Geldhof, G. J., Bowles, R. P., Nancarrow, A. F., et al. (2021). The Head-Toes-Knees-Shoulders Revised: Links to academic outcomes and measures of EF in young children. *Front. Psychol.* 12, 721846. doi: 10.3389/fpsyg.2021.721846
- McClelland, M. M., Tominey, S. L., Schmitt, S. A., Hatfield, B. E., Purpura, D. J., Gonzales, C. R., et al. (2019). Red light, purple light! Results of an intervention to promote school readiness for children from low-income backgrounds. *Front. Psychol.* 10, 2365. doi: 10.3389/fpsyg.2019.02365
- McCoy, D. C. (2019). Measuring young children's executive function and self-regulation in classrooms and other real-world settings. *Clin. Child Fam. Psychol. Rev.* 22, 63–74. doi: 10.1007/s10567-019-00285-1
- McKown, C. (2017). *Social-Emotional Assessment, Performance, and Standards (EJ1144767)*. ERIC. Available online at: <https://files.eric.ed.gov/fulltext/EJ1144767.pdf> (accessed May 31, 2023).
- Moeyaert, M., Ugille, M., Natasha Beretvas, S., Ferron, J., Bunuan, R., and Van den Noortgate, W. (2017). Methods for dealing with multiple outcomes in meta-analysis: A comparison between averaging effect sizes, robust variance estimation and multilevel meta-analysis. *Int. J. Soc. Res. Methodol.* 20, 559–572. doi: 10.1080/13645579.2016.1252189
- Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., et al. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proc. Natl. Acad. Sci. U. S. A.* 108, 2693–2698. doi: 10.1073/pnas.1010076108
- Morrison, F. J., and Grammer, J. K. (2016). "Conceptual clutter and measurement mayhem: Proposals for cross-disciplinary integration in conceptualizing and measuring executive function," in *Executive Function in Preschool-Age Children: Integrating Measurement, Neurodevelopment, and Translational Research*, eds J. A. Griffin, P. McCauley, and L. S. Freund (American Psychological Association), 327–348. doi: 10.1037/14797-015
- National Governors Association Center for Best Practices and Council of Chief State School Officers (2010a). *Common Core State Standards for Mathematics*. Available online at: [https://learning.ccsso.org/wp-content/uploads/2022/11/Math\\_Standards1.pdf](https://learning.ccsso.org/wp-content/uploads/2022/11/Math_Standards1.pdf) (accessed May 31, 2023).
- National Governors Association Center for Best Practices and Council of Chief State School Officers (2010b). *Common Core State Standards for English Language Arts and Literacy in History/Social Studies, Science, and Technical Subjects*. Available online at: [https://learning.ccsso.org/wp-content/uploads/2022/11/ELA\\_Standards1.pdf](https://learning.ccsso.org/wp-content/uploads/2022/11/ELA_Standards1.pdf) (accessed May 31, 2023).
- Nigg, J. T. (2017). Annual research review: On the relations among self-regulation, self-control, executive functioning, effortful control, cognitive control, impulsivity, risk-taking, and inhibition for developmental psychopathology. *J. Child Psychol. Psychiatr. Allied Discipl.* 58, 361–383. doi: 10.1111/jcpp.12675
- Pandey, A., Hale, D., Das, S., Goddings, A.-L., Blakemore, S.-J., and Viner, R. M. (2018). Effectiveness of universal self-regulation-based interventions in children and adolescents: A systematic review and meta-analysis. *J. Am. Med. Assoc. Pediatr.* 172, 566–575. doi: 10.1001/jamapediatrics.2018.0232
- Paris, S. G. (2005). Reinterpreting the development of reading skills. *Read. Res. Quart.* 40, 184–202. doi: 10.1598/RRQ.40.2.3

- PRISMA. *Homepage*. (2020). Available online at: <http://www.prisma-statement.org/> (accessed December 16, 2021).
- Purpura, D. J., Schmitt, S. A., and Ganley, C. M. (2017). Foundations of mathematics and literacy: The role of executive functioning components. *J. Exp. Child Psychol.* 153, 15–34. doi: 10.1016/j.jecp.2016.08.010
- Pustejovsky, J. E., and Tipton, E. (2021). Meta-analysis with robust variance estimation: Expanding the range of working models. *Prev. Sci.* 2021, vyfj. doi: 10.31222/osf.io/vyfj
- Raver, C. C., Carter, J. S., McCoy, D. C., Roy, A., Ursache, A., and Friedman, A. (2012). “Testing models of children’s self-regulation within educational contexts: Implications for measurement,” in *Advances in Child Development and Behavior*, eds J. B. Benson (Cambridge, MA: Elsevier Academic Press), 245–270.
- Raver, C. C., Jones, S. M., Li-Grining, C., Zhai, F., Bub, K., and Pressler, E. (2011). CSRPs impact on low-income preschoolers’ preacademic skills: Self-regulation as a mediating mechanism. *Child Dev.* 82, 362–378. doi: 10.1111/j.1467-8624.2010.01561.x
- Razza, R. A., Linsner, R. U., Bergen-Cico, D., Carlson, E., and Reid, S. (2020). The feasibility and effectiveness of mindful yoga for preschoolers exposed to high levels of trauma. *J. Child Fam. Stud.* 29, 82–93. doi: 10.1007/s10826-019-01582-7
- Reid, R., Trout, A. L., and Schartz, M. (2005). Self-regulation interventions for children with attention deficit/hyperactivity disorder. *Except. Child.* 71, 361–377. Available online at: <https://psycnet.apa.org/record/2005-06917-001>.
- Rimm-Kaufman, S. E., Curby, T. W., Grimm, K. J., Nathanson, L., and Brock, L. L. (2009). The contribution of children’s self-regulation and classroom quality to children’s adaptive behaviors in the kindergarten classroom. *Dev. Psychol.* 45, 958–972. doi: 10.1037/a0015861
- Robson, D. A., Allen, M. S., and Howard, S. J. (2020). Self-regulation in childhood as a predictor of future outcomes: A meta-analytic review. *Psychol. Bull.* 146, 324–354. doi: 10.1037/bul0000227
- Schmitt, S. A., Geldhof, G. J., Purpura, D. J., Duncan, R., and McClelland, M. M. (2017). Examining the relations between executive function, math, and literacy during the transition to kindergarten: A multi-analytic approach. *J. Educ. Psychol.* 109, 1120–1140. doi: 10.1037/edu0000193
- Schmitt, S. A., McClelland, M. M., Tominey, S. L., and Acock, A. C. (2015). Strengthening school readiness for Head Start children: Evaluation of a self-regulation intervention. *Early Childh. Res. Quart.* 30, 20–31. doi: 10.1016/j.ecresq.2014.08.001
- Schmitt, S. A., Pratt, M. E., and McClelland, M. M. (2014). Examining the validity of behavioral self-regulation tools in predicting preschoolers’ academic achievement. *Early Educ. Dev.* 25, 641–660. doi: 10.1080/10409289.2014.850397
- Sezgin, E. Y., Ulus, L., and Sahin, I. (2019). The examination of relationship between behavioral regulation with phonological awareness, matching images and pre-writing skills in preschool children. *Sakarya Univ. J. Educ.* 9, 107–128. doi: 10.19126/suje.463610
- Spiegel, J. A., Goodrich, J. M., Morris, B. M., Osborne, C. M., and Lonigan, C. J. (2021). Relations between executive functions and academic outcomes in elementary school children: A meta-analysis. *Psychol. Bull.* 147, 329–351. doi: 10.1037/bul0000322
- Spiegel, J. A., and Lonigan, C. J. (2018). A head-to-toes approach to computerized testing of executive functioning in young children. *Early Childh. Res. Quart.* 44, 15–23. doi: 10.1016/j.ecresq.2018.01.008
- Südkamp, A., Kaiser, J., and Möller, J. (2012). Accuracy of teachers’ judgments of students’ academic achievement: A meta-analysis. *J. Educ. Psychol.* 104, 743–762. doi: 10.1037/a0027627
- Tanner-Smith, E. E., and Tipton, E. (2013). Robust variance estimation with dependent effect sizes: Practical considerations including a software tutorial in Stata and SPSS. *Res. Synth. Methods* 5, 13–30. doi: 10.1002/jrsm.1091
- ten Braak, D., Storksen, I., Idsoe, T., and McClelland, M. (2019). Bidirectionality in self-regulation and academic skills in play-based early childhood education. *J. Appl. Dev. Psychol.* 65, 101064. doi: 10.1016/j.appdev.2019.101064
- U.S. Department of Health, and Human Services Administration for Children and Families (2015). *Head Start Early Learning Outcomes Framework: Ages Birth to Five*. Available online at: <https://eclkc.ohs.acf.hhs.gov/sites/default/files/pdf/elof-ohs-framework.pdf> (accessed May 31, 2023).
- Valcan, D. S., Davis, H. L., Pino-Pasternak, D., and Malpique, A. A. (2020). Executive functioning as a predictor of children’s mathematics, reading and writing. *J. Appl. Dev. Psychol.* 70, 101196. doi: 10.1016/j.appdev.2020.101196
- von Suchodoletz, A., Gestsdottir, S., Wanless, S. B., McClelland, M. M., Birgisdottir, F., Gunzenhauser, C., et al. (2013). Behavioral self-regulation and relations to emergent academic skills among children in Germany and Iceland. *Early Childh. Res. Quart.* 28, 62–73. doi: 10.1016/j.ecresq.2012.05.003
- Wanless, S. B., McClelland, M. M., Acock, A. C., Cameron Ponitz, C. E., Son, S.-H., Lan, X., et al. (2011a). Measuring behavioral regulation in four societies. *Psychol. Assess.* 23, 364–378. doi: 10.1037/a0021768
- Wanless, S. B., McClelland, M. M., Tominey, S. L., and Acock, A. C. (2011b). The influence of demographic risk factors on children’s behavioral regulation in prekindergarten and kindergarten. *Early Educ. Dev.* 22, 461–488. doi: 10.1080/10409289.2011.536132
- Welsh, J. A., Bierman, K. L., Nix, R. L., and Heinrichs, B. N. (2020). Sustained effects of a school readiness intervention: 5th grade outcomes of the Head Start REDI program. *Early Childh. Res. Quart.* 53, 151–160. doi: 10.1016/j.ecresq.2020.03.009
- Willoughby, M. T., Blair, C. B., Wirth, R. J., and Greenberg, M. (2012). The measurement of executive function at age 5: Psychometric properties and relationship to academic achievement. *Psychol. Assess.* 24, 226–239. doi: 10.1037/a0025361
- Wilson, D. (2019). *Meta-Analysis [Course Manual]*. Ardmore, PA: Statistical Horizons.
- Wulf, G., McNevin, N., and Shea, C. H. (2001). The automaticity of complex motor skill learning as a function of attentional focus. *Quart. J. Exp. Psychol.* 54, 1143–1154. doi: 10.1080/713756012
- Zangaro, G. A., and Soeken, K. L. (2005). Meta-analysis of the reliability and validity of Part B of the Index of Work Satisfaction across studies. *J. Nurs. Measur.* 13, 7–22. doi: 10.1891/jnum.2005.13.1.7
- Zhang, L., and Rao, N. (2017). Effortful control and academic achievement in rural China. *Early Educ. Dev.* 28, 541–558. doi: 10.1080/10409289.2016.1255080