Nordic Journal of Systematic Reviews in Education
Volume 2 | 2024 | pp. 207–228

Protocol

A Systematic Review of the Effectiveness of Algebraic Teaching Interventions: A Meta-Analysis and Meta-Synthesis

Sheree T. Sharpe*, Matthew C. Mauntel, Rebecca Butler, Bitna Choi, Maryam Aswad & Jerome Amedu
University of New Hampshire, United States

ABSTRACT
Algebra has long been recognized as a fundamental component of mathematics education for K-12 students and has been identified as a subject with which students continually struggle. Researchers have utilized various methods across contexts and conditions at the classroom level to improve algebra learning. This systematic review and meta-research (i.e., meta-analysis and meta-synthesis) aims to elucidate which of these efforts are effective, along with the conditions and populations for which they are most effective. In this article, we present our framing for the study under a modified version of the conceptual framework for learning progression, justify selected moderators, and detail our anticipated research process. Conducting meta-research on this topic is essential for providing policymakers, instructors, and researchers with an adequate understanding of the historical landscape of effective practices in algebra instruction.

Keywords: *grades K-12; Algebraic teaching interventions; systematic review; meta-analysis; meta-synthesis*

Introduction

Research has demonstrated that a foundational understanding of algebraic content is the gateway to pursuing careers in the science, technology, engineering, and mathematics (STEM) fields (Hughes et al., 2014). More specifically, an understanding of algebraic content is crucial for excelling in higher-level mathematics and STEM-based courses, completing high school, and gaining admission into college (National Mathematics Advisory Panel, 2008; Vogel, 2008). However, many students lack the algebraic skills needed to be successful in STEM courses and subsequent careers (Booth et al., 2014). Given the importance of algebra to students' success with

*Correspondence: Sheree T. Sharpe, e-mail: Sheree.Sharpe@unh.edu

Citation: Sharpe, S. T., Mauntel, M. C., Butler, R., Choi, B., Aswad, M. & Amedu, J. (2024). A Systematic Review of the Effectiveness of Algebraic Teaching Interventions: A Meta-Analysis and Meta-Synthesis. Nordic Journal of Systematic Reviews in Education, 2, 207–228. http://doi.org/10.23865/njsre.v2.6355

mathematics content, this article describes a research protocol intended to identify which algebraic teaching interventions are effective, for whom, and under what conditions. We aim to do this by conducting a systematic review and meta-research (i.e., meta-analysis and meta-synthesis) of research across 40 years of algebra literature spanning Grades K-12. We define *Algebraic Teaching Interventions* as processes by which the teacher/researcher implements strategies in the classroom designed with the goal of improving the teaching and learning of algebraic content.

Algebra Difficulties Across the K-12 Spectrum

While arithmetic typically precedes algebra in traditional mathematics instruction, this approach has not been successful (Blanton et al., 2018; Filloy & Rojano, 1989; Sharpe, 2019), and there is evidence that students continue to have difficulty with algebra using this approach (Bednarz, 2001; Booth, 1984; Kieran, 2007; MacGregor, 1996; National Center for Educational Statistics, 2022). Students' difficulties with algebra and algebraic thinking have been well documented and are rooted in misunderstandings of the following algebraic skills: (a) Using the equal sign as a balancing point (Booth, 1984; Kieran, 1981, 1985; Vergnaud, 1985, 1988); (b) Using mathematical symbols to show the relationships between quantities (Bednarz, 2001; Bednarz & Janvier, 1996; Vergnaud, 1985; Wagner, 1981); (c) Applying the commutative and distributive properties (Boulton-Lewis et al., 2001; Demana & Leitzel, 1988; MacGregor, 1996); (d) Using letters as generalized numbers or as variables (Booth, 1984; Ely & Adams, 2012; Küchemann, 1981; Vergnaud, 1985); (e) Working with unknowns and applying equivalent transformations to both sides of an equation (Bednarz, 2001; Bednarz & Janvier, 1996; Filloy & Rojano, 1989; Kieran, 1985, 1989; Steinberg et al., 1990); and (f) Generating equations and solving algebra problems (Kieran, 2007).

One means by which these misunderstandings can be addressed is through the incorporation of algebraic skills into earlier mathematics instruction. Specifically, students who are introduced to algebraic thinking in elementary grades have a better understanding of algebraic content (Blanton et al., 2018; Carraher et al., 2008). Since the early incorporation of algebraic skills can positively influence students' algebra learning, standards documents (e.g., Principles and Standards for School Mathematics [National Council of Teachers of Mathematics (NCTM), 2000], Common Core State Standards [National Governors Association Center for Best Practices & Council of Chief State School Officers (NGA & CCSSO), 2010]) now incorporate algebraic thinking and algebra standards across K-12 grades, instead of restricting algebra instruction to the secondary grades.

Given this integration of algebraic content into the K-12 curricula, it is important to understand the ways in which algebra has been incorporated into these classrooms and the impacts this content has had on students' understanding. Our study aims to provide such an understanding by qualitatively and quantitatively synthesizing the existing literature on this topic.

Algebraic Teaching Framework

In order to conceptualize the linkages between algebraic curriculum, algebraic instruction, and K-12 students' mathematical learning, we will implement a modified version of the conceptual framework for learning progression (Blanton et al., 2018; Fonger et al., 2018). This conceptual framework for learning progression was initially developed from empirical research coupled with national and state standards. The authors of the conceptual framework for learning progression developed and implemented an instructional sequence for the elementary grades in order to determine levels of sophistication in children's thinking. The authors address the following components as children progress through the early algebra curriculum: "(a) A curricular framework and progression of learning goals across big ideas and thinking practices; (b) An instructional sequence; (c) Assessments; and (d) Levels of sophistication in children's thinking" (Fonger et al., 2018, p. 35). Blanton et al. (2018) found that the algebraic big ideas for students in elementary grades are: (1) Equivalence, expressions, equations, and inequality (EEEI); (2) Functional thinking; and (3) Generalized arithmetic. Thinking practices included generalizing, representing, justifying, and reasoning with math structure and relationships (Blanton et al., 2018; Kaput, 2008). These algebraic big ideas and thinking practices categories will be integrated into our algebraic teaching framework.

The aim of the algebraic teaching framework is to articulate the links found in the conceptual framework for learning progression, specifically for algebraic curriculum, algebraic instruction, and students' mathematical learning, across a wide grade span (K-12). Our algebraic teaching framework will serve as an analytic guide to understand and represent the links between algebraic curriculum, algebraic instruction, and K-12 students' mathematical learning from 1980 to the present in the meta-synthesis. Furthermore, this research will use the algebraic big ideas and thinking practices from the conceptual framework for learning progression to classify and describe the different types of algebraic teaching interventions, learning goals, and impacts on student learning over the K-12 grades across 40 years of research. More specifically, we will focus on algebraic big ideas and thinking practices as a proxy for curriculum, algebraic teaching interventions and learning goals as a proxy for instruction, and student mathematical thinking as a proxy for student learning (see Figure 1).

Prior Systematic Reviews on Mathematics and Algebraic Interventions

Since 1980, there have been many systematic reviews that have examined interventions for improving mathematics achievement at the K-12 level. Here we provide an overview of systematic reviews that have addressed mathematics interventions in general, highlighting literature focused on algebra interventions, and review their impact on mathematical achievement. We do this to contextualize our intended meta-research project within a broader body of work that draws inferences for mathematics learning across contexts. In what follows, we present these previous systematic reviews by the type(s) of interventions they analyzed and the populations they targeted. We also

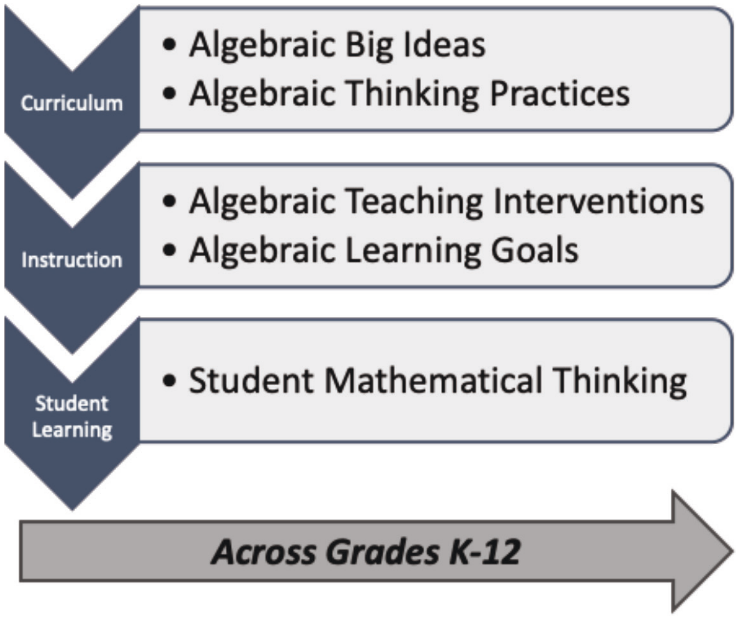

Figure 1. Algebraic Teaching Framework

provide relevant effect size data for the meta-analyses that are closely aligned with the goals of our study (algebraic teaching interventions spanning Grades K-12 and their impact on mathematics achievement).

Several meta-analyses have addressed specific types of math interventions based on their content. Ran et al. (2021, 2022) found statistically significant positive effects for computer and technology interventions, respectively. Sercenia and Prudente (2023) looked at metacognitive strategies for improving mathematics and found a large positive effect (g = 1.358). This indicates that content-targeted mathematics interventions can have large impacts on mathematics achievement.

Meta-analyses that focus on students in special education or those with learning difficulties in mathematics span a variety of topics and age groups. Lein et al. (2020) and Myers et al. (2022) analyzed studies with word problem-solving interventions for K-12 students in special education (g = .56) and mathematical difficulties (g = .71), respectively. Gersten et al. (2009) analyzed studies pertaining to students with learning disabilities and identified peer-assisted learning, explicit instruction with ongoing feedback, and opportunities for deliberate practice as effective practices. In contrast, computer-assisted instruction and general teaching practices were not effective (Gersten et al., 2009). Kroesbergen and Van Luit (2003) focused on elementary education students in special education and found basic skills to be the most effective intervention. Overall, interventions targeted at students with special education needs and students with mathematical difficulties can have large impacts on their mathematical achievement.

There are several meta-analyses that have analyzed general mathematical knowledge. Pellegrini et al. (2021) analyzed elementary school programs targeted at increasing mathematical achievement and found that computer-assisted instruction and tutoring were the most promising (g = .2). Williams et al. (2022) conducted a meta-analysis reviewing randomized experimental studies with mathematics interventions from 1991 to 2017. They found a small positive (g = .24) average effect size for algebra interventions during this period. Notably, the authors considered pre-K-12 mathematics interventions that were labeled algebra and pre-algebra interventions, and did not classify any of the interventions as early algebra interventions.

A number of previous meta-analyses have focused on categorizing and analyzing interventions or strategies targeting students' algebra achievement. For example, Haas (2005) identified six instructional intervention categories: Direct instruction, cooperative learning, communication and study skills, multiple representations, problem-based learning, and technology. Rakes et al. (2010) suggested five new categories of algebra instructional improvement strategies: Technology curricula, non-technology curricula, instructional strategies, manipulatives, and technology tools. Hughes et al. (2014) focused specifically on interventions aimed at improving algebra achievement targeted at students with disabilities or at risk of having disabilities. Some of these included guided practice, mastery learning, assessment, self-monitoring, and use of visual representations (Hughes et al., 2014). Other reviews (e.g., Impecoven-Lind & Foegen, 2010; Maccini et al., 1999) also identified specific interventions, strategies, or elements of instruction that increased the algebra achievement of students with specific learning disabilities.

While there have been many meta-analyses examining algebra and mathematics interventions, there have been a limited number of meta-syntheses on these topics. Those that exist are mostly focused on specialized topics or specific populations. Dibbs et al. (2020) conducted a meta-synthesis looking at how students combine like terms, a core topic in algebra. They found journaling, multiple representations, direct instruction, and group work to be effective methods for helping students learn how to combine like terms. Korkmaz and Morali (2022) investigated the impact of augmented reality (AR) on mathematics education from 2010 to 2021. They found that AR had a positive impact on students' understanding of mathematics. While these meta-syntheses identified practices, they did not provide an in-depth description of each practice and explore why those practices were effective.

Our work extends the previous meta-analysis and meta-synthesis studies in a variety of ways. First, we examine algebraic teaching interventions spanning Grades K-12 for a general population over a 40-year period. We consider not only algebraic teaching interventions that took place during an algebra class, but also early algebra and pre-algebra interventions. Additionally, we consider non-randomized experimental studies as well as randomized experimental studies, provide a categorization of interventions, and develop a rich description of effective interventions through a meta-synthesis (qualitative analysis). The scope of this study in terms of time, population, and research type distinguishes us from previous studies like Williams et al. (2022).

The field of mathematics education has produced many qualitative and quantitative studies over the last 40 years, yet there are currently no studies that include both a meta-analysis and meta-synthesis on the same research topic. This study includes both aspects and applies them to a wide swath of literature across several decades and settings. Meta-analysis and meta-synthesis each address different kinds of research questions, extract different types of data from studies, employ different methods of analysis, and deliver different characteristics in results. Our rationale for conducting a mixed-methods synthesis is to provide a richer understanding of algebraic teaching interventions by producing a holistic picture that can inform future research, policy, and practice in the teaching and learning of school algebra. The aim of the meta-analysis is to quantitatively analyze the effectiveness of algebraic teaching interventions, while the meta-synthesis complements this by providing a detailed picture of the characteristics of effective algebraic teaching interventions. The meta-analysis results will establish intervention efficacy, whereas the meta-synthesis will describe the relationships between algebraic curriculum, algebraic instruction, and student mathematical learning in K-12 school algebra.

Research on Potential Moderators

As we explore the effects of algebraic teaching interventions on student learning in the meta-analysis, we will also analyze the moderators that may explain the different effects of these algebraic teaching interventions.

Demographic Moderators: We plan to consider race/ethnicity, gender, socioeconomic status (SES), location (within the U.S.), and student grade level as demographic moderators since each of these have been previously identified as significant moderators in analyses of students' mathematics outcomes. The National Center for Educational Statistics (NCES) issued a report evidencing significant differences in mathematical achievement in relation to race/ethnicity, gender, and SES as part of its National Report Card (National Center for Educational Statistics, 2022). Sharpe and Marsh (2022) conducted a systematic review of algebra teaching and learning and found that race/ethnicity, SES, and gender each play a role in students' achievements. Voyer and Voyer (2014) conducted a meta-analysis looking at school achievement including mathematics and found that race and gender were significant moderators. Ran et al. (2021) and Cheung and Slavin (2013) conducted meta-analyses examining technology's impact on mathematics achievement and found that grade level was a significant modifier. Grade level is an important moderator in the current study, since changes like presenting algebra in 8th grade tend to impact on mathematical success differently when compared with the traditional introduction of algebra in 9th grade (Domina, 2014; Rickles, 2013).

Algebraic Teaching Intervention Moderators: We will investigate the impact of algebraic teaching interventions as categorized by the length of the intervention, who delivered the intervention (technology, teacher, researcher), the amount of teacher training used to implement the intervention (not reported, none, one-time training, multiple training/ongoing training), and the type of intervention. We consider these factors because Williams et al. (2022) found that intervention delivery was a significant moderator in their meta-analysis of mathematics interventions for grades pre-K-12. Additionally, a meta-analysis conducted by Blank and de las Alas (2009) found that teacher professional development/training had a significant effect on student mathematical achievement.

Algebraic Content Moderators: In concert with our algebraic teaching framework, we will be considering the algebraic big ideas (i.e., EEEI, functional thinking, generalized arithmetic) as potential moderators. We have already highlighted the difficulties that students face in each of these areas in our Algebra Difficulties section. By conducting a moderator analysis on the big ideas we hope to identify which ones contributed most to the heterogeneity of algebraic teaching intervention effects. This will contextualize and inform the meta-synthesis by providing a focus for analysis on the intervention types and algebra topics that explain most of the heterogeneity.

Study Characteristic Moderators: Williams et al. (2022) found that the type of outcome measure recorded (researcher-developed versus standardized achievement) and the publication decade were significant moderators. Since our study spans 40 years, we will look at both as potential moderators, in addition to publication type.

Purpose and Research Questions

In light of the previously discussed nature of algebra as being simultaneously essential to mathematics education and difficult for students to understand, the purpose of this research is to identify what types of algebraic teaching interventions are effective, for whom, and under what conditions.

Methodologically, this purpose is addressed through systematic review and meta-research of empirical studies examining algebraic teaching interventions in the US in Grades K-12 between 1980 and the present. We define meta-research as a mixed methods study with quantitative (meta-analysis) and qualitative (meta-synthesis) analyses of primary studies. The time frame for research inclusion was determined by the systemic changes in the field that included the release of several seminal pieces, including *An Agenda for Action* in 1980 (NCTM, 1980), *Principles and Standards for School Mathematics* in 2000 (NCTM, 2000), and *Common Core State Standards for Mathematics* (NGA & CCSSO, 2010) in 2010. These systemic changes in mathematics pedagogy warrant not only an overview of the research around algebraic teaching interventions, but also a synthesis that analyzes the effect of these changes on the field. The goal of the meta-analysis is to provide quantitative results and inform the focus of the meta-synthesis. We will conduct this systematic review and meta-research with the aim of answering the following research questions in regard to US elementary and secondary school students from 1980 to the present.

Meta-Analysis Research Question:

1. (a) What is the effectiveness of algebraic teaching interventions?
   (b) How heterogeneous are algebraic teaching intervention effects for US elementary and secondary school students since 1980?
   (c) What moderators (e.g., grade level/algebra level, algebraic big ideas, gender, socioeconomic status, race/ethnicity, location in US, intervention type, intervention delivery, intervention length, implementation training, publication year, publication type, assessment type) contribute to the heterogeneity of algebraic teaching intervention effects?

Meta-Synthesis Research Questions:

2. What are the characteristics of algebraic teaching interventions that have facilitated US elementary and secondary school students' mathematics learning across algebraic big ideas and grade levels since 1980?
   (a) What are the characteristics of effective algebraic teaching interventions and their learning goals (instruction) that have facilitated students' mathematics learning across algebraic big ideas and thinking practices (curriculum) and grade levels?
   (b) What are the levels of sophistication in students' mathematical thinking across algebraic big ideas, thinking practices, and grade levels?
   (c) How does mathematics learning differ across algebraic big ideas, thinking practices, and grade levels for students with diverse backgrounds?
3. What recommendations for algebraic teaching interventions that facilitate student mathematics learning across grade levels are derived from empirical studies?

Mixed Methods Research Question:

4. To what extent and in what ways does the meta-synthesis of algebraic teaching intervention characteristics contribute to a more comprehensive and nuanced understanding of the effect sizes from meta-analyses?

Methods

In this section, we describe the inclusion/exclusion criteria for selecting studies, the research process, the search strategy, the multi-stage screening process, full-text retrieval, data extraction, data coding, data management, and handling of missing data. The PRISMA-P checklist was used for items included in this systematic review protocol (Moher et al., 2015).

Inclusion and Exclusion Criteria

A primary goal of our meta-synthesis is to inform policy and practice by synthesizing the recommendation sections from a broad set of primary studies. Given the differences in policy and practice across countries, we have restricted ourselves to US-based research and excluded studies using data from other countries. A secondary goal of this study is to examine the impact of US standards documents in mathematics education since 1980, further leading us to exclusively consider studies from the US. We will use the Methods, Units, Treatments, Outcomes, and Settings (MUTOS) framework (Williams et al., 2022) to establish the study eligibility criteria for the systematic review. The methods will be Randomized Control Trial (RCT), Quasi-Experimental Design (QED) with baseline equivalence, empirical qualitative research, or mixed methods studies that include RCT, QED, and/or empirical qualitative research. The units will be K-12 students in the US or its territories. The treatments will be Algebraic Teaching Intervention and a business-as-usual control group. The business-as-usual control group is a comparison group that did not receive the intervention. The outcome will be measures of mathematics learning, thinking, or knowledge. The settings will be published (books, journal articles, etc.) and unpublished studies (dissertations, conference papers, independent reports, etc.) written in English and issued between 1980 and the present, conducted within the US or its territories. Given that studies with statistically significant results have a higher probability of being published, we will include both published and unpublished studies to address publication bias.

The aim of this systematic review is to include empirical studies, hence the following type of studies will be excluded from this analysis: Reviews of literature or summaries of research, meta-analyses, meta-syntheses, policy documents, calls for research, book reviews, op-ed pieces, and pedagogical or practitioners' articles describing implementation of teaching, tools, and/or practice with learners.

We will search for studies based on the first eight systematic review eligibility criteria below. Eligibility criteria 9 to 11 will come into consideration during the screening.

Table 1. Systematic review eligibility criteria for included studies

Type of Criteria	Eligibility Criteria
1. Intervention	Included at least one algebraic teaching intervention that occurred in the classroom different from business-as-usual.
2. Location	Study occurred in the US or one of its territories.
3. Sample	Included a sample of students in elementary or secondary grades (Grades K-12 only).
4. Outcome	Evaluated at least one measure of mathematics learning, thinking, or knowledge.
5. Language	Written in English.

(*Continued*)

Table 1. (*Continued*)

Type of Criteria	Eligibility Criteria
6. Study Year	Published and unpublished studies issued between January 1980 and the present.
7. Study Type – Qual	Qualitative studies: Empirical qualitative research.
8. Study Type – Quan (RCT/QED)	Quantitative studies: Conducted either Quasi-Experimental Design (baseline equivalence) or Randomized Control Trials.
9. Study Type – Quan (Control Group)	Quantitative studies: Included a business-as-usual control group.
10. Study Type – Quan (Sufficient Info)	Quantitative studies: Study provided sufficient information to calculate an effect size and its variance.

Systematic Review and Meta-Research Process

The process for conducting this systematic review and meta-research is: (1) Data collection of initial studies (literature search); (2) Quality appraisal of the initial studies (screening); (3) Data extraction; and (4) Data analysis (meta-analysis and meta-synthesis) (Ong et al., 2020; Williams et al., 2022). Figure 2 is a visual representation of the project's systematic review and meta-research process. We will use the Covidence platform to store and screen the studies. Note that Covidence has two

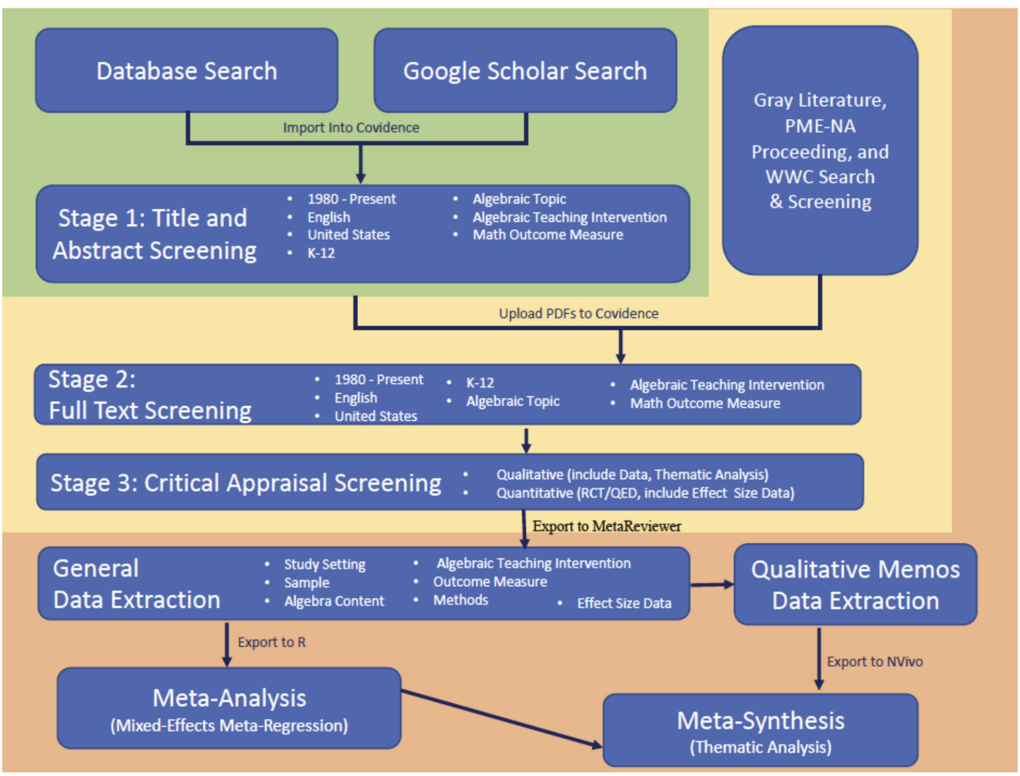

Figure 2. Systematic Review and Meta-Research Process

levels of screening: title/abstract and full-text. We will conduct the full-text and critical appraisal screening at the second level of Covidence screening. We will also use Covidence platform for data extraction, the R software for our meta-analysis and the NVivo software for our meta-synthesis.

Search Strategy

The systematic search strategy will include searches of electronic databases, Google Scholar, gray literature websites, What Works Clearinghouse (WWC) website, and the North American Chapter of the International Group for the Psychology of Mathematics Education (PME-NA) *Annual Conference Proceedings*.

The database and Google Scholar search results will be uploaded to the Zotero program and combined into one file. We will use Zotero's automated duplicate feature for duplicate detection and removal. A combined file will then be exported to the Covidence platform for title and abstract screening (Stage 1). We will then conduct a second round of duplication removal using Covidence. As shown in Figure 2, gray literature, WWC, and PME-NA searches and screening will be first conducted on the different websites (Stage 1 screening) and then the PDFs of the studies that pass the Stage 1 screening will be uploaded to the Covidence platform for full-text screening (Stage 2). After the upload, duplicates will again be removed using Covidence's automated duplicate removal feature.

The searches will utilize search terms that focus on the eligible criteria discussed above (Williams et al., 2022). For example, within these searches, we will use a combination of keywords that refer to (1) Algebra; (2) K-12 grades; (3) Outcomes; and (4) Methods. We will filter our searches based on date (1980–present), language of writing (English), and location of the studies (US or its territories). See Figure 3 below for an example of a detailed search string.

(Algebra* OR Equation* OR "Pre-algebra*" OR "functions" OR teaching algebra*) **AND**

("elementary education" OR elementary school OR elementary OR secondary OR "primary education" OR primary school OR middle school OR secondary school OR "secondary education" OR high school OR junior high OR children OR adolescen* OR teen* OR grade* OR classroom OR K12 OR K-12 OR school age OR school based OR school system OR school district OR school) **AND**

(ability* OR Achievement OR Acquisition OR Attainment OR Fluency OR Mastery OR Maintenance OR Outcome* OR Performance OR Proficiency OR Retention OR Skill* OR Understanding OR application* OR computation* OR word problem*) **AND**

(RCT OR experiment* OR randomize* OR QED OR "Quasi-Experiment*" OR "non-randomized" OR non-experimental OR effectiveness study OR efficacy study OR effectiveness trial OR efficacy trial* OR qualitative OR Qualitative Empirical Studies OR Empirical*)

Filters: 1980 – present; English Language; U.S.

Figure 3. Detailed Search String for APA PsycINFO on EBSCO

Database Searches: We will first conduct separate electronic database searches of ERIC, Education Source, PsycINFO, SocINDEX, Sociological Abstracts, Academic Search Ultimate, JSTOR, WorldCat FirstSearch, and ProQuest. We will also conduct a search of Google Scholar using a shorter version of the above search terms, since Google Scholar has a limit on the length of its search strings.

Gray Literature Searches: Similar to Williams et al. (2022), we will conduct search and screening of the gray (or unpublished) literature websites by searching the U.S. Department of Education websites and websites of research organizations known to fund or research mathematics interventions. We will search the following websites for citations or reports related to algebraic interventions: Abt Associates, American Institutes for Research, Bill & Melinda Gates Foundation, Campbell Collaboration, Center for Teaching Quality, Center on Great Teachers and Leaders, Charles A. Dana Center, Consortium for Policy Research in Education, Institute of Education Sciences (National Center for Education Research [NCER], and National Center for Education Evaluation and Regional Assistance [NCEE]), Mathematica Policy Research, MDRC, National Center for Analysis of Longitudinal Data in Education Research (CALDER), National Center on Intensive Intervention, National Research and Development Center on Cognition and Mathematics Instruction, National Research Council, National Science Foundation, Peabody Research Institute, Program in Mathematics Education, RAND Corporation, Randomized Controlled Trials Registry, Regional Educational Laboratories, Research in Cognition and Mathematics Education, RTI International, SRI International, University of Chicago Consortium on School Research, U.S. Department of Education's Investing in Innovation (i3), Westat, WestEd, and William and Ida Friday Institute for Educational Innovation.

What Works Clearinghouse (WWC) Search & Screening: The WWC is an investment by the U.S. Department of Education that reviews a collection of research to determine effective high-quality educational research. We will conduct a search and screening of this Clearinghouse for algebraic-related interventions.

The North American Chapter of the International Group for the Psychology of Mathematics Education (PME-NA) Search & Screening: Given that PME-NA is a North American conference representing a wide range of mathematics education research, with proceedings representing the last 40 years, we will conduct a search and screening of the PME-NA conference proceedings on the PME-NA website for algebraic-related interventions. The titles in the proceedings will be screened for relevant studies. A screener will then check the related abstract using the title and abstract screening criteria, and studies that pass will be extracted from the proceedings and uploaded into Covidence for full-text screening.

Multi-Stage Screening Strategy

Plans for screening potentially eligible studies: Screening will occur in three stages: (1) Title and abstract screening; (2) Full-text screening; and (3) Critical appraisal screening. The research team will be trained to conduct the screening at each stage. During training, 10 studies will be given to each team member to screen, then the team will meet to discuss their screening decisions and resolve discrepancies. This process will happen twice for each stage of screening. The principal investigator (PI) will also meet with all reviewers weekly to prevent screening drift over time. To continuously monitor screening, all studies will be dual and independently screened at each stage. Any discrepancies/conflicts will be discussed and resolved by the research team in their weekly meetings and interrater reliability will be evaluated via Cohen's kappa, in addition to recall, specificity, and balanced accuracy (Syriani et al., 2023). A separate screening form will be developed for all three stages of screening based upon the eligibility criteria. A "NO" response from the reviewers to any of the screening questions will exclude the study from further review. If the reviewer responds with a "YES" or "MAYBE" to all screening questions at Stage 1, the study will move to the next stage. In Stages 2 and 3, the reviewers will need to respond with a "YES" to all screening questions for the study to be included, that is, there will be no "MAYBE" option. Studies that make it through all three stages of screening will be included and coded. In summary, screening will be completed in the following three stages:

- Stage 1. Titles and abstracts will be screened (using the Covidence platform) for eligibility criteria. That is, *determine* whether the appropriate intervention, outcomes, and samples (studies issued between 1980–present, written in English, conducted in the US or its territories, involving students in Grades K-12) could be present in the study. The Gray Literature Stage 1 screening will be conducted on the different websites at this stage.
- Stage 2. Full-texts will be screened to *confirm* that the appropriate intervention, outcomes, and samples were included in the studies.
- Stage 3. Full-texts will be screened again for *high-quality* studies. We will conduct a critical appraisal of the methods sections for both qualitative and quantitative studies. The qualitative studies will be screened for the use of thematic qualitative analysis and the use of data to support their findings. The quantitative studies will be screened for the use of RCT or QED with equivalence baseline and sufficient data to compute effect sizes.

Full-Text Retrieval

Full-text retrieval of PDFs will be completed between Stages 1 and 2 of screening. The list of studies that pass title and abstract screening will be downloaded from Covidence and uploaded into Zotero. Zotero's automated full-text search feature will be used to search for texts on open databases and databases available from the university library. Full-texts not found using Zotero's automated search feature will be manually searched in the university's library databases and using Google Search. Remaining full-texts not located will be requested through an interlibrary loan system or purchased.

Data Extraction

Combination of Included Studies: Multiple studies that use the same dataset will NOT be counted as separate studies; however, we will use all available sources to extract the data from each study. For example, if a study includes a peer-reviewed article, report, and conference paper/abstract, we will use all three sources to extract the data and it will be counted as one study.

General Extraction (All Studies): The PI, in conjunction with the research team, will develop a data extraction template in Covidence. All variables will be extracted as reported in the study. All team members will be trained to extract data from the studies. During training, 10 studies will be given to each person to extract and the team will meet to discuss decisions or discrepancies. This process will be repeated twice to ensure agreement. The PI will meet with the team weekly to discuss any questions and prevent extraction drift over time. To continuously monitor data extraction, all of the studies will be extracted by two team members and checked by a third team member. Any extraction discrepancies will be discussed by the two team members and, if unresolved, the PI will make the final decision.

The goal for the analysis, and therefore the data extraction, is to identify the best practices from algebraic teaching interventions that work, for whom, and under what conditions. We will extract data involving study-level characteristics, sample characteristics, algebraic content, algebraic teaching intervention, outcome measure(s), and methods of the study. Studies that are classified with a quantitative or mixed methods design will also be extracted for additional data to be used in the meta-analysis.

Meta-Analysis Extraction (Quantitative/Mixed Methods Studies): Information to compute effect sizes will be extracted from each unique quantitative study. Any calculation of effect sizes will be done separately from the reported data. Our goal is to compute the standardized mean difference (SMD) effect size for all related math outcomes reported in each study. We will also compute effect size by race/ethnicity, SES, or gender (i.e., % female) if enough data is available (i.e., subgroup analyses). In addition, the following data will be extracted for the quantitative studies: Sample size, means, and standard deviations by group. If this data is not available, then we will extract the t-test, F-test, regression coefficients, or effect sizes in other metrics, that can be used to compute SMDs.

The extracted data will be downloaded from Covidence as a CSV file and the clean file will be uploaded to the R software for quantitative analysis.

Data Management: We will use a shared Google Drive to store all project documents. We will upload our initial list of citations to Zotero and then to Covidence for title and abstract screening. We will upload PDFs to Covidence for full-text screening, critical appraisal screening, and data extraction. All documents, citations, PDFs, and data files will be stored on our shared Google Drive, as well as on the PI's computer.

Meta-Synthesis Extraction (All Studies): All study PDFs will be downloaded from Covidence. A qualitative memo will be developed for each unique study. A qualitative memo includes the extracted data from a study relevant to the purpose of the research (Jaumot-Pascual et al., 2021). Using memos allows the same data to be analyzed across all studies and enables the removal of extraneous information (Finfgeld-Connett, 2018). For the purpose of this research, each study will be summarized as a memo, and each memo will be treated analogously as an informant for the meta-synthesis. A qualitative memo will include descriptive data and images from a study that are relevant to the meta-synthesis, including participants' descriptions, algebraic content description, description of algebraic teaching intervention, algebraic learning goals, main findings of the study, and any recommendations.

Handling of missing data: We expect that we will have missing data due to the nature of empirical research. There are three types of missing data: Missing studies, missing effect sizes, and missing moderators (Pigott, 2012). We will include both published and unpublished studies to address missing studies. If descriptive statistics to compute an effect size are missing, we will extract summary statistics in other metrics to address missing effect sizes. Finally, we will use multiple imputation to account for missing moderator data (Pigott, 2001, 2012). Pigott (2012) describes multiple imputation as:

> …a technique that generates multiple possible values for each missing observation in the data. Each of these values is used in turn to create a complete data set. The analyst uses standard statistical procedures to analyze each of these multiply imputed data sets, and then combines the results of these analyses for statistical inference (p. 99).

Analysis

In this section, we describe how we will assess risk of bias beyond our critical appraisal screening (i.e., methods screening), and the data analysis plans for both quantitative and qualitative analysis. That is, we describe the process for conducting the meta-analysis and meta-synthesis.

Assessment of Risk of Bias for Included Studies

To address study quality at the "front end" of the review, we will conduct a third stage of screening (i.e., the critical appraisal of the studies' methods). In the meta-synthesis, this means we will include qualitative studies that provide sufficient evidence to support the findings. In the meta-analysis, this means we will include RCT and QED (with baseline equivalence) studies with a control group. On the "back end" of the review for the meta-analysis, two members of the research team will independently assess the risk of bias using the Cochrane Risk of Bias Tool – RoB 2 for randomized trials and ROBINS-I tool for non-randomized studies of interventions (Sterne et al., 2019; Sterne et al., 2016). We will assess the following domains: (1) Bias arising from the randomization process; (2) Bias due to deviations from intended interventions; (3) Bias arising from baseline equivalence; (4) Bias in measurement of the outcome; and (5) Bias in selection of the reported result. Items will be rated for risk of bias as "low risk", "unclear", or "high risk" following the guidance in the Cochrane Risk of Bias Tool (Sterne et al., 2019). In regard to selective reporting bias, we will use funnel plots to assess publication bias and use meta-regression to assess small-study effects (Carter et al., 2019). We will also use publication type as a moderator variable.

Plans for data analysis and synthesis

The goal for the meta-analysis and the meta-synthesis is to identify the best practices from the algebraic teaching intervention that are effective, for whom, and under what conditions. A "meta-analysis uses a statistical procedure that aggregates and condenses a body of quantitative research studies to a common standard metric, such as a mean effect size" (Finlayson & Dixon, 2008; Thunder & Berry, 2016, p. 319). In contrast, a meta-synthesis uses a "deliberate process of selecting studies with an emphasis on synthesizing, analyzing, and interpreting findings across the selected studies" (Thunder & Berry, 2016, p. 319). All stages of the analysis will be discussed with the research team.

Meta-Analysis: The analyses will be completed by the PI using R software. Data cleaning will occur in the Excel spreadsheet before uploading to the R software.

Effect Sizes – The goal is to compute the standardized mean difference (SMD) effect size for all related math outcomes reported in each study. If descriptive statistics are not available, then reported summary statistics in other metrics will be used to compute the SMDs. The Hedges' g effect size will be used to adjust for small-sample bias and the PI will adjust the effect size variances for clustering when the level of random assignment is at the cluster level (i.e., teacher, school; Williams et al., 2022). Effect sizes will be generated from independent samples; if a single study contains multiple independent samples, effect size will be calculated for each sample. Given the inclusion of non-randomized studies, pretest-adjusted effect sizes will be calculated (Hedges et al., 2023).

The PI will use mixed-effects meta-regression models to investigate sources of effect size heterogeneity. The PI will conduct both multilevel model (hierarchical structure of data) and multivariate model (interdependencies of variables) analyses (Pustejovsky & Tipton, 2022; Williams et al., 2022). The PI will estimate models using a restricted maximum likelihood with the *metafor* package in the R software. To account for effect-size dependencies (i.e., multiple effects per study), the PI will use a robust variance estimation to adjust the standard errors and degrees of freedom for the regression coefficients, using the small-sample correction based on the Satterthwaite approximation and the *clubSandwich* R package (Tipton, 2015).

The goal of the meta-analysis is not only to explain the heterogeneity of the algebraic teaching intervention, but also to inform the focus of the meta-synthesis. That is, the meta-analysis will report which moderators significantly explain the variability across algebraic teaching interventions, and the meta-synthesis will provide a rich description of the differences in algebraic teaching interventions for these significant moderators.

Meta-Synthesis: The meta-synthesis will be completed by the research team using NVivo software. Qualitative memos will be written in a Microsoft Word document before being uploaded to the NVivo software.

Synthesize findings – A qualitative memo addressing participants' backgrounds, curriculum, instruction, student learning, and the study's recommendations will be developed for each unique study in a single Microsoft Word document to be uploaded to NVivo and coded. A thematic analysis (Braun & Clarke, 2006) will be used to code, categorize, and constantly compare data to develop a general theory of best practices from algebraic teaching interventions. The research team will be trained to independently open code the memos and then negotiate their independent open codes to reach a shared set of initial codes and definitions to be used consistently throughout the analysis of the data. The initial codes will be categorized and reconciled between at least two coders. The data will then be sorted by codes and reread, looking for themes within each code to see if there are dimensions that require the data to be further discriminated. Through this process, themes will be derived from the data through the shared inquiry of multiple researchers about the characteristics of algebraic teaching interventions that can facilitate US students' learning.

Summary

The purpose of the protocol is to identify which algebraic teaching interventions are effective, for whom, and under what conditions. Algebraic teaching interventions are categorized by length, type, and the role of the deliverer of the intervention. The first step of the protocol is a systematic search through databases, gray literature, WWC, and PME-NA proceedings. The next step is a multi-stage screening strategy using the MUTOS framework, including a title and abstract screening, a full-text screening, and a critical appraisal. The third step is extraction. Information from quantitative and mixed methods studies will be coded to compute effect sizes for a meta-analysis. Qualitative memos will be created for all studies for a meta-synthesis. The meta-analysis will explain heterogeneity in the algebraic teaching intervention's effect sizes and inform the focus of the meta-synthesis. Using the algebraic teaching framework, the meta-synthesis will work toward a general theory of best practices in teaching algebra, by describing the relationship between algebraic curriculum, algebraic instruction, and student mathematical learning in K-12 school algebra.

Significance of Research

The results of this research will inform future research, policy, and practice in the teaching and learning of algebra across the K-12 spectrum. The descriptive study will highlight gaps in algebra research literature related to algebraic teaching interventions and algebraic big ideas across grade levels and/or systemic changes in the field. The meta-analysis will explain the variability in algebraic teaching interventions across the selected studies as it pertains to our moderators. The meta-synthesis will explore the connections between students' backgrounds, algebraic curriculum, algebraic instruction, and mathematical learning.

The results of this research will also inform politicians who develop legislation concerning reforms in education, researchers who develop interventions, and teachers who establish best practices with regard to school algebra. In addition, results will connect research across multiple developmental areas – early algebra, pre-algebra, and algebra – across different demographic groups, and across grade level bands. Our work will provide teachers and educators with information to help students from all different backgrounds, and increase the information available to stakeholders interested in improving algebra skills for all K-12 students and educators.

Acknowledgments

This material is based upon work supported by the National Science Foundation DRK-12 under Grant No. 2142659. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Author biographies

Sheree T. Sharpe is an associate professor of mathematics education at the University of New Hampshire. Her research focuses on the teaching and learning of school algebra, spanning Kindergarten to Grade 12, and using systematic reviews, meta-analyses, and meta-syntheses methodologies to conduct research.

Matthew C. Mauntel is a postdoctoral mathematics education researcher at the University of New Hampshire. His research focuses on designing curriculum, integrating technology into the classroom, and using artificial intelligence to enhance K-12/undergraduate engagement and learning. He leverages meta-research and design-based methodologies to support curricular design and is passionate about his teaching.

Rebecca Butler is a mathematics education PhD candidate and Dissertation Year Fellow at the University of New Hampshire. She focuses her research on university STEM course transformation and mathematics teacher education within the undergraduate context. Across her work, she incorporates computational tools such as large language models, machine learning, and text mining into traditional discursive approaches.

Bitna Choi is a PhD student in mathematics education and a graduate research assistant at the University of New Hampshire. Her research interests include systematic review and meta-analysis methodologies, particularly in the context of educational inequality and algebra for K-12 students.

Maryam Aswad is a PhD student in mathematics education at the University of New Hampshire. She is interested in undergraduate teaching and learning, mathematics anxiety, and the integration of authentic situations into the learning experience.

Jerome Amedu is a postdoctoral mathematics education researcher at the University of New Hampshire. His research interests include teacher learning, educational technology, and systematic review of evidence. Jerome is a CADRE Fellow and was part of the inaugural cohort of the North Carolina State University Holmes Scholars Program.

References

Bednarz, N. (2001). A problem-solving approach to algebra: Accounting for the reasoning and notations developed by students. In H. Chick, K. Stacey, J. Vincent, & J. Vincent (Eds.), *The future of the teaching and learning of algebra: Proceedings of the 12th ICMI Study Conference* (Vol. 1, pp. 69–78). The University of Melbourne. https://doi.org/10.1007/1-4020-8131-6

Bednarz, N., & Janvier, B. (1996). Emergence and development of algebra as a problem-solving tool: Continuities and discontinuities with arithmetic. In N. Bednarz, C. Kieran, & L. Lee (Eds.), *Approaches to algebra. Perspectives for research and teaching* (pp. 115–136). Kluwer. https://doi.org/10.1007/978-94-009-1732-3_8

Blank, R., & de las Alas, N. (2009). *The effects of teacher professional development on gains in student achievement: How meta analysis provides scientific evidence useful to education leaders*. Council of Chief State School Officers.

Blanton, M., Brizuela, B., Stephens, A., Knuth, E., Isler, I., Gardiner, A., Stroud, R., Fonger, N., & Stylianou, D. (2018). Implementing a framework for early algebra. In C. Kieran (Ed.), *Teaching and learning algebraic thinking with 5- to 12-year-olds: The global evolution of an emerging field of research and practice*. (pp. 27–49). Springer International Publishing. https://doi.org/10.1080/14794802.2020.1725613

Booth, L. (1984). *Algebra: Children's strategies and errors. A report of the strategies and errors in secondary mathematics project*. NFER-Nelson.

Booth, J., Barbieri, C., Eyer., F., & Pare-Blagoev, J. (2014). Persistent and pernicious errors in algebraic problem solving. *Journal of Problem Solving, 7*, 10–23. http://dx.doi.org/10.7771/1932-6246.1161

Boulton-Lewis, G., Cooper, T., Atweh, B., Pillay, H., & Wilss, L. (2001). Readiness for algebra. In T. Nakahara & M. Koyama (Eds.), *Proceedings of the 24th International Conference for the Psychology of Mathematics Education* (Vol. 2, pp. 89–96). Retrieved July 28, 2023 from: https://files.eric.ed.gov/fulltext/ED452032.pdf#page=99

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Carraher, D., Schliemann, A., & Schwartz, J. (2008). Early algebra is not the same as algebra early. In J. Kaput, D. Carraher, & M. Blanton (Eds.), *Algebra in the early grades* (pp. 235–272). Routledge. https://doi.org/10.4324/9781315097435-12

Carter, E. C., Schonbrodt, F. D., Gervais, W. M., & Hilgard, J. (2019). Correcting for bias in psychology: A comparison of meta-analytic methods. *Advances in Methods and Practices in Psychological Science*, *2*(2), 115–144. https://doi.org/10.1177/2515245919847196

Cheung, A., & Slavin, R. (2013). The effectiveness of educational technology applications for enhancing mathematics achievement in K-12 classrooms: A meta-analysis. *Educational Research Review*, *9*, 88–113. https://doi.org/10.1016/j.edurev.2013.01.001

Demana, F., & Leitzel, J. (1988). Establishing fundamental concepts through numerical problem solving. In A. F. Coxford & A. P. Shulte (Eds.), *The ideas of algebra, K-12: 1988 Yearbook* (pp. 61–68). National Council of Teachers of Mathematics.

Dibbs, R. A., Hott, B. L., Martin, A., Raymond, L., & Kline, T. (2020). Combining like terms: A qualitative meta-synthesis of algebra I interventions in mathematics and special education. *International Journal of Education in Mathematics, Science and Technology*, *8*(3), 219–232. https://doi.org/10.46328/ijemst.v8i3.862

Domina, T. (2014). The link between middle school mathematics course placement and achievement. *Child Development*, *85*(5), 1948–1964. https://doi.org/10.1111/cdev.12255

Ely, R., & Adams, A. (2012). Unknown, placeholder, or variable: What is x? *Mathematics Education Research Journal*, *24*, 19–38. https://doi.org/10.1007/s13394-011-0029-9

Filloy, E., & Rojano, T. (1989). Solving equations: The transition from arithmetic to algebra. *For the Learning of Mathematics, 9*(2), 19–25.

Finfgeld-Connett, D. (2018). *A guide to qualitative meta-synthesis*. Routledge.

Finlayson, K. W., & Dixon, A. (2008). Qualitative meta-synthesis: A guide for the novice. *Nurse Researcher, 15*(2), 59–71. https://doi.org/10.7748/nr2008.01.15.2.59.c6330

Fonger, N., Stephens, A., Blanton, M., Isler, I., Knuth, E., & Gardiner, A. (2018). Developing a learning progression for curriculum, instruction, and student learning: An example from mathematics education. *Cognition and Instruction, 36*(1), 30–55. https://doi.org/10.1080/07370008.2017.1392965

Gersten, R., Chard, D. J., Jayanthi, M., Baker, S. K., Morphy, P., & Flojo, J. (2009). *A meta-analysis of mathematics instructional interventions for students with learning disabilities: Technical report*. Instructional Research Group. Retrieved July 28, 2023 from https://files.eric.ed.gov/fulltext/ED612790.pdf.

Haas, M. (2005). Teaching methods for secondary algebra: A meta-analysis of findings. *Nassp Bulletin, 89*(642), 24–46.

Hedges, L. V., Tipton, E., Zejnullahi, R., & Diaz, K. G. (2023). Effects sizes in ANCOVA and differences-in-differences designs. *British Journal of Mathematical and Statistical Psychology, 76*(2), 259–282. https://doi.org/10.1111/bmsp.12296

Hughes, E., Witzel, B., Riccomini, P., Fries, K., & Kanyongo, G. (2014). A meta-analysis of algebra interventions for learners with disabilities and struggling learners. *The Journal of the International Association of Special Education, 15*(1), 36–47. Retrieved July 28, 2023 from: https://uploads.documents.cimpress.io/v1/uploads/2c0e2671-73b3-47e9-a8cc-bc4943b563d1~110/original?tenant=vbu-digital

Impecoven-Lind, L. S., & Foegen, A. (2010). Teaching algebra to students with learning disabilities. *Intervention in School and Clinic, 46*(1), 31–37.

Jaumot-Pascual, N., Silva, C., Martinez-Gudapakkam, A., & Ong, M. (2021). Women of color in computing graduate education: Structural supports and navigation strategies for a hostile culture. *Research in Equity and Sustained Participation in Engineering, Computing, and Technology (RESPECT)*, 1–9. https://doi.org/10.1109/RESPECT51740.2021.9620675

Kaput, J. (2008). What is algebra? What is algebraic reasoning? In J. J. Kaput, D. W. Carraher & M. L. Blanton (Eds.), *Algebra in the early grades* (pp. 5–17). Lawrence Erlbaum Associates. https://doi.org/10.4324/9781315097435-2

Kieran, C. (1981). Concepts associated with the equality symbol. *Educational Studies in Mathematics, 12*(3), 317–326. https://doi.org/10.1007/bf00311062

Kieran, C. (1985). The equation-solving errors of novices and intermediate algebra students. In *Proceedings: IX International Conference Psychology of Mathematics Education*. Retrieved July 28, 2023 from: https://files.eric.ed.gov/fulltext/ED411130.pdf#page=153

Kieran, C. (1989). The early learning of algebra: A structural perspective. In S. Wagner & C. Kieran (Eds.), *Research agenda for mathematics education: Research issues in the learning and teaching of algebra* (Vol. 4, pp. 33–56). National Council of Teachers of Mathematics. https://doi.org/10.4324/9781315044378-4

Kieran, C. (2007). Learning and teaching algebra at the middle school through college levels: Building meaning for symbols and their manipulation. In F. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (Vol. II, pp. 707–762). Information Age Publishing.

Korkmaz, E., & Morali, H. S. (2022). A meta-synthesis of studies on the use of augmented reality in mathematics education. *International Electronic Journal of Mathematics Education*, *17*(4), em0701. https://doi.org/10.29333/iejme/12269

Kroesbergen, E. H., & Van Luit, J. E. (2003). Mathematics interventions for children with special educational needs: A meta-analysis. *Remedial and Special Education*, *24*(2), 97–114. https://doi.org/10.1177/07419325030240020501

Küchemann, D. (1981). Algebra. In K. Hart (Ed.), *Children's understanding of mathematics* (pp. 102–119). Murray.

Lein, A. E., Jitendra, A. K., & Harwell, M. R. (2020). Effectiveness of mathematical word problem solving interventions for students with learning disabilities and/or mathematics difficulties: A meta-analysis. *Journal of Educational Psychology*, *112*(7), 1388. https://doi.org/10.1037/edu0000453

Maccini, P., McNaughton, D., & Ruhl, K. L. (1999). Algebra instruction for students with learning disabilities: Implications from a research review. *Learning Disability Quarterly, 22*(2), 113–126. https://doi.org/10.2307/1511270

MacGregor, M. (1996). Curricular aspects of arithmetic and algebra. In R. C. L. J. Gimenez & B. Gomez (Eds.), *Arithmetics and algebra education: Searching for the future* (pp. 50–54). Universitat Rovira I Virgili.

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

Myers, J. A., Hughes, E. M., Witzel, B. S., Anderson, R. D., & Owens, J. (2022). A meta-analysis of mathematical interventions for increasing the word problem solving performance of upper elementary and secondary students with mathematics difficulties. *Journal of Research on Educational Effectiveness, 16*(1), 1–35. https://doi.org/10.1080/19345747.2022.2080131

National Center for Educational Statistics. (2022). *NAEP long-term trend assessment results: Reading and mathematics.* The National Center for Educational Statistics. Retrieved July 28, 2023 from https://www.nationsreportcard.gov/highlights/ltt/2023/

National Mathematics Advisory Panel. (2008). *Foundations for success: The final report of the National Mathematics Advisory Panel*. U.S. Department of Education.

National Council of Teachers of Mathematics (NCTM). (1980). *An agenda for action: Recommendations for school mathematics of the 1980s*. Retrieved July 28, 2023 from: https://www.nctm.org/Standards-and-Positions/More-NCTM-Standards/An-Agenda-for-Action-(1980s)/

National Council of Teachers of Mathematics (NCTM). (2000). *Principles and standards for school mathematics*. Retrieved July 28, 2023 from: https://www.nctm.org/Standards-and-Positions/Principles-and-Standards/

National Governors Association Center for Best Practices & Council of Chief State School Officers (NGA & CCSSO). (2010). *Common core state standards for mathematics.* Retrieved July 28, 2023 from: https://learning.ccsso.org/wp-content/uploads/2022/11/ADA-Compliant-Math-Standards.pdf

Ong, M., Jaumot-Pascual, N., & Ko, L. (2020). Research literature on women of color in undergraduate engineering education: A systematic thematic synthesis. *Journal of Engineering Education, 109*, 581–615. https://doi.org/10.1002/jee.20345

Pellegrini, M., Lake, C., Neitzel, A., & Slavin, R. E. (2021). Effective programs in elementary mathematics: A meta-analysis. *AERA Open, 7*(1), 1–29. https://doi.org/10.1177/2332858420986211

Pigott, T. (2001). Missing predictors in models of effect size. *Evaluation & the Health Professions, 24*(3), 277–307. https://doi.org/10.1177/01632780122034920

Pigott, T. (2012). Missing data in meta-analysis: Strategies and approaches. In *Advances in meta-analysis* (pp. 79–107). Springer. https://doi.org/10.1007/978-1-4614-2278-5_7

Pustejovsky, J. E., & Tipton, E. (2022). Meta-analysis with robust variance estimation: Expanding the range of working models. *Prevention Science, 23*, 425–438. https://doi.org/10.1007/s11121-021-01246-3

Rakes, C. R., Valentine, J. C., McGatha, M. B., & Ronau, R. N. (2010). Methods of instructional improvement in algebra: A systematic review and meta-analysis. *Review of Educational Research, 80*(3), 372–400. https://doi.org/10.3102/0034654310374880

Ran, H., Kasli, M., & Secada, W. G. (2021). A meta-analysis on computer technology intervention effects on mathematics achievement for low-performing students in K-12 classrooms. *Journal of Educational Computing Research, 59*(1), 119–153. https://doi.org/10.1177/0735633120952063

Ran, H., Kim, N., & Secada, W. (2022). A meta-analysis on the effects of technology's functions and roles on students' mathematics achievement in K-12 classrooms. *Journal of Computer Assisted Learning*, *38*(1), 258–284. https://doi.org/10.1111/jcal.12611

Rickles, J. (2013). Examining heterogeneity in the effect of taking algebra in eighth grade. *The Journal of Educational Research*, *106*(4), 251–268. https://doi.org/10.1080/00220671.2012.692731

Sercenia, J. C., & Prudente, M. S. (2023). Effectiveness of the metacognitive-based pedagogical intervention on mathematics achievement: A meta-analysis. *International Journal of Instruction, 16*(4), 561–578. https://doi.org/10.29333/iji.2023.16432a

Sharpe, S. (2019). An algebraic translation task solved by Grade 7–9 students. *Mathematical Thinking and Learning, 21*(1), 78–84. https://doi.org/10.1080/10986065.2019.1564970

Sharpe, S., & Marsh, D. (2022). A systematic review of factors associated with high schoolers' algebra achievement according to HSLS: 09 results. *Educational Studies in Mathematics*, 1–24. https://doi.org/10.1007/s10649-021-10130-4

Steinberg, R., Sleeman, D., & Ktorza, D. (1990). Algebra students' knowledge of equivalence of equations. *Journal for Research in Mathematics Education, 22*(2), 112–121. https://doi.org/10.5951/jresematheduc.22.2.0112

Sterne, J. A., Hernán, M. A., Reeves, B. C., Savović, J., Berkman, N. D., Viswanathan, M., Henry, D., Altman, D. G., Ansari, M. T., Boutron, I., Carpenter, J. R., Chan, A., Churchill, R., Deeks, J. J., Hróbjartsson, A., Kirkham, J., Jüni, P., Loke, Y. K., Pigott, T. D., …, & Higgins, J. P. (2016). ROBINS-I: A tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*, *355*, Article i4919. https://doi.org/10.1136/bmj.i4919

Sterne, J. A., Savović, J., Page, M. J., Elbers, R. G., Blencowe, N. S., Boutron, I., Cates, C. J., Cheng, H., Corbett, M. S., Eldridge, S. M., Hernán, M. A., Hopewell, S., Hróbjartsson, A., Junqueira, D. R., Jüni, P., Kirkham, J. J., Lasserson, T., Li, T., McAleenan, A., Reeves, B. C., …, & Higgins, J. P. (2019). RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ, 366*, Article 14898. https://doi.org/10.1136/bmj.l4898

Syriani, E., David, I., & Kumar, G. (2023). *Assessing the ability of chatgpt to screen articles for systematic reviews*. arXiv. https://doi.org/10.48550/arXiv.2307.06464

Thunder, K., & Berry, R. (2016). The promise of qualitative metasynthesis for mathematics education. *Journal for Research in Mathematics Education, 47*(4), 318–337. https://doi.org/10.5951/jresematheduc.47.4.0318

Tipton, E. (2015). Small sample adjustments for robust variance estimation with meta-regression. *Psychological Methods, 20*(3), 375–393. https://doi.org/10.1037/met0000011

Vergnaud, G. (1985). Understanding mathematics at the secondary-school level. In A. Bell, B. Low, & J. Kilpatrick (Eds.), *Theory, research & practice in mathematical education* (pp. 27–45). Shell Center for Mathematical Education.

Vergnaud, G. (1988). Long terme et court terme dans l'apprentissage de l'algèbre [Long term and short term in learning algebra]. In C. Laborde (Ed.), *Actes du premier colloque franco-allemand de didactique des mathematiques et de l'informatique* (pp. 189–199). La Pensée Sauvage.

Vogel, C. (2008). Algebra: Changing the equation. *District Administration, 44*, 34–40.

Voyer, D., & Voyer, S. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin*, *140*(4), 1174–1204. https://doi.org/10.1037/a0036620

Wagner, S. (1981). Conservation of equation and function under transformations of variable. *Journal for Research in Mathematics Education, 12*(2), 107–118. https://doi.org/10.5951/jresematheduc.12.2.0107

Williams, R., Citkowicz, M., Miller, D., Lindsay, J., & Walters, K. (2022). Heterogeneity in mathematics intervention effects: Evidence from a meta-analysis of 191 randomized experiments. *Journal of Research on Educational Effectiveness*, *15*(3), 584–634. https://doi.org/10.1080/19345747.2021.2009072