Prate et al. *Research and Practice in Technology Enhanced Learning* (2026) 21:1
https://doi.org/10.58459/rptel.2026.21001

Research and Practice in
Technology Enhanced Learning

RESEARCH **Free and Open Access**

A meta-analysis of technology usage in word problem solving interventions for elementary students: an application of the SAMR model

Joelle Prate *, Jennifer Kong, Lamiya Hoque, Trisha Sugita and Tanya Belom

*Correspondence:
Jprate@Chapman.edu
Attallah College of Education,
Chapman University,
1 University Dr.,
Orange, CA 92866, United States
Full list of author information is available at the end of the article

Abstract

This meta-analysis assessed the effect of the usage of technology in math interventions on the math achievement of students in kindergarten to the fifth grade. The SAMR (Substitution, Augmentation, Modification, Redefinition) (Puentedura, 2006) model was applied to evaluate the degree of technology integration within the math intervention. Twelve group studies met the criteria for inclusion in this meta-analysis. Overall, technology interventions yielded a significant positive effect on the math achievement of elementary-aged students with learning disabilities (LD) (ES = 1.34). While no significant differences in effect sizes were found by function of the SAMR model, the studies that utilize technology as substitutions yielded the highest effect sizes. These findings support the need for continued study into the effects of technology-mediated strategies in word problem solving interventions.

Keywords: Math intervention, Technology, Elementary, SAMR model

Introduction

Math word problem solving proficiency in early grades has been recognized as an important skill for future academic success (Suseelan et al., 2022; Swanson et al., 2013). As such, national math standards include instructional focus on visual and verbal problem solving as early as kindergarten (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010). Math word problems involve processes more demanding than basic math skills in that they require students to use linguistic information to construct a problem model and solve (Fuchs et al., 2006). To successfully solve word problems, students need various skills such as reading, comprehension, and conceptual understanding (Boonen et al., 2016; Kim & Xin, 2022). These advanced processes lead many students to experience significant difficulty with word problems (Swanson, 2006). Unfortunately, only 36% of fourth grade students in the U.S. demonstrate math achievement at proficient or advanced levels on the National Assessment of Educational Progress (U.S. Department of Education, 2022). It is critical to identify effective strategies in math word problem instruction to address these challenges.

One widely used practice in math education is the integration of technology in instruction. In recent years, the use of technology within society and the classroom has been widespread and increasingly accessible. The National Council of Teachers of Mathematics includes the use of technology in teaching and learning mathematics as one of six major principles for high-quality math education (National Council of Teachers of Mathematics, 2000). The use of technology, generally, has been shown to be effective in math interventions (Kim & Xin, 2022; Kim et al., 2021; Kiru et al., 2017; Ran et al., 2020; Xin & Jitendra, 1999; Zhang & Xin, 2012). As an example, a recent meta-analysis on the effects of computer assisted instruction (CAI) on general math achievement of students from first grade to postsecondary education reported overall positive effects (ES = 0.552; Benavides-Varela et al., 2020). However, effect sizes ranged from –1.356 to 2.543, demonstrating mixed findings as to the impact of technology assisted math instruction. More specifically, to the authors' knowledge, no studies to date have examined the degree to which technology is integrated within these math interventions. Furthermore, much of the research on technology enhanced math instruction investigated the impact on math generally. That is, the literature on the impact of technology integration within word problem solving specifically is extremely limited (Kim & Xin, 2022).

Technology-mediated learning, once relegated to obscure application, has become ubiquitous in modern classrooms. However, as we seek to understand the best practices and potential value of these technologies, there is a growing divide between device functionality and the pedagogical implications of its use in practice. When evaluating the role of technology in learning we must consider it within the contexts of learning which are cognitive and social in nature (Salomon & Perkins, 1998). As such, when seeking to understand how technology may impact outcomes for students in math interventions, we must look beyond the devices themselves into the specific, individual practices and surrounding contexts to understand their impact.

The purpose of this study is to better understand how the degree to which technology is applied within math problem solving interventions may improve student outcomes in word problem solving and the conditions under which technology application may be most effective at meeting this goal.

The SAMR model

The Substitution, Augmentation, Modification, Redefinition (SAMR) Model (Puentedura, 2006) offers a framework to classify the degree to which technology is utilized within a lesson. This model organizes technology not by the device or service capabilities, but by how it is deployed in the classroom or intervention. It has been used in numerous studies (Bicalho et al., 2023; Blundell et al., 2022; Hamilton et al., 2016) to provide a framework for categorizing the application of technology which looks most closely at how students engage with devices rather than what the device itself is capable of. These four layers are also grouped into two broader categories with substitution and augmentation considered as enhancements to learning whereas redefinition and modification levels represent a transformation in learning.

Each level of the SAMR model has been described in application. Figure 1 provides a visual representation of the taxonomy of the SAMR model and its application to instruction, and Table 1 provides examples for each level's description, examples from word problem solving (WPS) studies, and rationale for clarity. As referenced in Figure 1, the substitution and augmentation levels fall within the broad category of enhancement, whereas the modification and reimagination levels fall within the transformation category. Enhancement of instruction refers to ways existing instructional strategies can be improved through technology whereas transformation of instruction refers to opportunities that were not possible without technology.

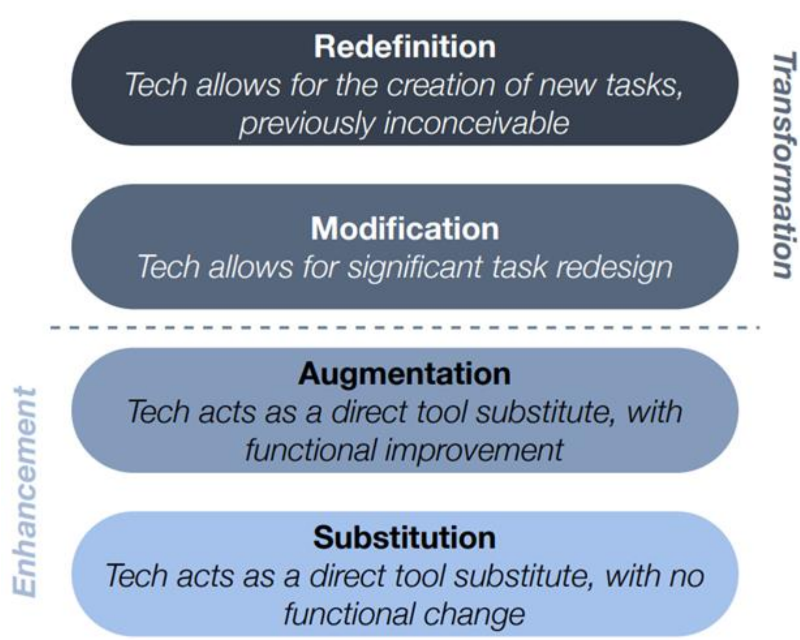

Fig. 1 The SAMR model (Puentedura, 2013)

Table 1 Technology applications in elementary word problem solving by SAMR descriptions

	Description	Study Examples	Rationale
Substitution	Technology uses were direct substitutes for their analog counterparts.	The use of a computerized version of word problem solving steps compared with a paper version (Chadli et al., 2018)	Digitizing resources and assessments which could otherwise be accessed via pencil and paper methods as a direct substitute.
Augmentation	Technology uses were functional improvements to instruction but still fell within traditional task design.	The use of adaptive questioning, which displays new problems leveled to students' individual performance and progress (Schoppek & Tullis, 2010)	The functional improvement of technology allows the students to receive dynamic questioning, leveled based on the students' answers to previous questions, without the assistance of a person.
Modification	Technology uses offered a significant task redesign but were not necessarily something previously inconceivable without technology.	The use of a gamified curriculum that included in-game rewards but also opportunities for students to communicate with peers and teachers in the virtual space (Yeh et al., 2019)	In-game rewards and communication were not previously inconceivable without technology, but the platform allowed for significant task redesign.
Redefinition	Technology uses permitted the inclusion of previously inconceivable tasks.	The use of a virtual reality environment to immerse students in a word problem solving game (Kim & Ke, 2016)	An immersive game environment such as those in virtual reality were not conceivable without the use of technology.

Enhancement

The substitution level is described in the original model as an enhancement where "tech[nology] acts as a direct tool substitute with no functional change" (Puentedura, 2006). From an instructional standpoint, examples could include asking students to use a word processor instead of hand writing (Setiyawati et al., 2023) or providing digital version of handouts (Hamilton et al., 2006). The augmentation level requires a functional improvement in the application of technology (Puentedura, 2006). For example, a teacher could use web resources to allow students access to reading selections or blogs not otherwise available in an analog format (Bicalho et al., 2023), using spell and grammar check functions (Nyayu et al., 2019; Setiyawati et al., 2023), or moving student discussions online (Kelsch & Wang, 2021).

Transformation

An important step in the SAMR taxonomy is the jump between augmentation and modification which constitutes not only a single step but also a move into the transformation category. A defining element of this difference is task redesign. For example, allowing students to access web resources for research (Setiyawati et al., 2023).

The expectation of greater depth of application, creating scenarios and learning opportunities of the SAMR model peaks at the redefinition category which is defined as a task which was previously inconceivable without the aid of technology (Puentedura, 2006). This includes a myriad of learning and creation opportunities including production of original multimedia content and collaboration across physical spaces (Hamilton et al., 2016; McClain & North, 2021).

Although the SAMR model is not inherently hierarchical, evidence from these studies suggests that there may be connections between these levels and student outcomes (McClain & North, 2021; Setiyawati et al., 2023). For example, in a quasi-experimental study focused on computer programming Setiyawati et al. (2023) found that it positively impacted students' critical thinking scores. Within the domain of mathematics education, McClain and North (2021) found statistically significant improvement of MAP Growth scores (NWEA, 2023) at each level of the SAMR framework.

Several reviews have analyzed the SAMR model in a variety of contexts including separating teacher and student actions with technology by instructional components (Blundell et al., 2022), teacher perception (Bicalho et al., 2023), approaches to teaching English (Nyayu et al., 2019) and social studies (Hilton, 2016), standardized measures of math (e.g., Measure of Academic Progress growth test; McClain and North, 2019), critical thinking (Andriani et al., 2022; Setiyawati et al., 2023) and mobile device applications (Crompton & Burke, 2020; Romrell et al., 2014). No studies, to the author's knowledge, have applied this model to word problem solving math interventions.

Prior meta-analyses/syntheses of technology interventions for word problem solving

Previous meta-analyses have investigated the effect of technology within an intervention to improve math achievement of students with math difficulties by way of moderator analysis (Kong et al., 2021; Lein et al., 2020; Xin & Jitendra, 1999; Zhang & Xin, 2012). The literature specifically focused on the use of technology within word problem solving interventions for students with math difficulties is especially limited (Kim & Xin, 2022). Word problem solving interventions have been found to be highly effective for students with learning or math disabilities in previous meta-analyses. This includes interventions which combined technology-based procedures with other strategies for word problem solving (Kong et al., 2021; Lein et al., 2020; Zhang & Xin, 2012). One synthesis focused specifically on computer-assisted math word problem solving instruction and it included 13 studies (10 group design and 3 single case studies) with participants kindergarten to middle school students (Kim & Xin, 2022). Computer-assisted instruction was defined as "the use of a computer to provide educational instruction," which includes the use of traditional computers and tablets. Kim and Xin categorized these interventions under four instructional categories: direct instruction/guided practice, cognitive/metacognitive strategy instruction, schema-based instruction, and mathematical model-based problem solving. Xin and Jitendra's study reported moderate effect size for technology interventions for students in group design studies (aggregated g = .77), with a wide range of effect sizes ranging from no effect to large effects, and large effect for single case studies (Tau-U = 0.99). Studies that utilized cognitive/metacognitive strategies alongside CAI yielded large effects (average g = 0.99, median Tau-U = 0.98).

The present study

The present study contributes to the literature base of the use of technology in word problem solving and extends upon the previous reviews in several ways. First, prior meta-analyses or syntheses in the area of technology to improve word problem solving provide effect sizes and instructional recommendations for a specific math practice (e.g., computer-assisted instruction or CAI; Kim & Xin, 2022). While this focus provides extensive knowledge in a targeted practice, this may not provide enough insight for teachers and researchers seeking recommendations for how best to integrate technology in math interventions including the group setting, effects by grade band, and the extent to which technology is applied.

Second, this study proposes to categorize studies not by the specific technological tool being utilized, but rather the degree to which technology is integrated within the lesson itself utilizing the SAMR model to differentiate between categories. A meta-analysis of math word problem solving is one approach to identifying valuable instructional practices. This study will address the following three research questions:

1. Are math interventions that incorporate technology effective for improving the math problem solving of kindergarten to grade 6 students with math difficulties? Effective outcomes will be based on the magnitude of the ESs.
2. Do effect sizes vary as a function of participant and intervention characteristics?
3. How does the level of technology integration as per the SAMR model affect the estimated ESs?

Literature search procedures

The PsycINFO, Science Direct, and ERIC online databases were systematically scanned for studies from 1990 to 2024 that met the inclusion criteria. Search terms describing word problem solving outcomes were combined with these keywords: word problem, elementary, technology, and intervention. Two rounds of initial searches generated 356 results. The reference lists of selected papers as well as prior literature reviews and meta-analyses (e.g., Gersten et al., 2009; Kim & Xin, 2022; Kong et al., 2021; Kroesbergen & Van Luit, 2003;

Myers et al., 2022; Xin & Jitendra, 1999; Zhang & Xin, 2012; Zheng et al., 2013) were also systematically scanned.

To be eligible for this analysis, each study had to meet the following criteria: (a) included students in elementary grades (K to 6); (b) tested an intervention to improve math integrating technology; (c) assessed students' math word problem solving outcomes (measure included normed or experimental/researcher developed measures); (d) involved an experimental design with randomization, quasi-experiment with pre- and post-test data, or a within-subjects design (i.e., all students participated in both the treatment and comparison conditions); (e) provided data to permit the calculation of effect sizes and average weighted ESs; and (f) was published in English. Studies investigating the effectiveness of instruction or improving only math calculations were not included. This procedure narrowed list to 12 studies which met inclusion criteria. Figure 2 depicts the process of this literature search and the number of studies excluded based on the above criteria. Some studies had more than one WPS intervention, so 21 different ESs were calculated.

Interrater agreement

Two graduate students independently coded all articles for inclusion criteria and coding accuracy. Interrater agreement was calculated as the number of agreements divided by the number of agreements plus disagreements multiplied by 100. The mean interrater agreement for article inclusion was above 95%. The mean interrater agreement for coding the SAMR framework and the twelve instructional components outlined below were also above 95%.

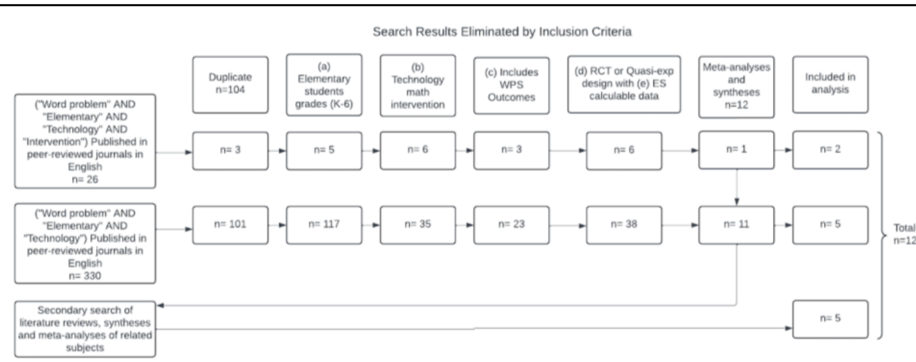

Fig. 2 Search results eliminated by inclusion criteria

Coding of study features

The general categories of coding for each study included: (a) year and country of publication, (b) sample characteristics (grade/age), and (c) intervention characteristics (SAMR categorization, number of sessions, number of minutes, group size, who delivered the instruction).

Categorization of the SAMR framework

Each study was coded on the degree to which technology was utilized in the intervention. We note that the SAMR framework is not conceptualized as a hierarchical model, and our coding was limited to the occurrence or nonoccurrence of these specific usages of technology based on the intervention descriptions provided. Table 1 includes a brief description, study examples, and rationale for each SAMR level.

Substitution. Studies that had descriptions in their interventions which included occurrences of technology uses which were direct substitutes for their analog counterparts were coded as “S”. Because a substitution may not include a functional change in the instruction, this can include things like pictures and graphics for counting, or static, prescribed content progression within the application.

Augmentation. Studies qualified as augmentation described functional improvement to instruction which still fell within traditional task design. For example, in mathematics instruction, an augmentation can look like a “hints” feature, adaptive structure, or engaging multimedia content. Studies which met these criteria were coded as “A”. This was found to be the most common type of technology-mediated intervention strategy.

Modification. The step from augmentation to modification is a greater step than the previous, because to reach the modification level, coded “M”, the study had to include use technology in a way which offered a significant task redesign, but was not necessarily something previous inconceivable, as it would be at the following level. Gamification was a common modification indicator, promoting student motivation using in-game rewards.

Redefinition. Redefinition in the area of mathematics instruction as a whole includes many unique instructional opportunities included augmented and virtual reality application. Although rare in intervention studies, the application of extended reality has many possibilities for the more applied and “real world” contexts.

Data analysis

Effect Size Calculation. Effect sizes (ESs) were calculated utilizing pretest and posttest means and standard deviations. Hedges’ *g* was the measure of ES for this study, calculated as the difference between pretest-posttest means for the treatment group and the pretest-posttest means for the comparison group. This difference score was then divided by the pooled within-group standard deviation of posttest scores. Hedges' *g* was calculated as

$$\frac{(X_{post_1} - X_{pre_1}) - (X_{post_2} - X_{pre_2})}{\sqrt{([n_1 - 1]s_1^2 + [n_2 - 1]s_2^2 / [n_1 + n_2 - 2])}}$$

where X_{pre1} and X_{pre2} were unadjusted pretest means, X_{post1} and X_{post2} were unadjusted posttest means, n_1 and n_2 were sample sizes, and s_1 were unadjusted standard deviations for the treatment and comparison groups, respectively. Planned tests to compare effect sizes as a function of the level of technology as per the SAMR model and intervention components utilizing a general linear model procedure were conducted (Borenstein et al., 2009). A summary of these findings is presented in Table 3.

Question 1: Are math interventions that incorporate technology effective for improving the math problem solving of kindergarten to grade 6 students with math difficulties?

To answer Research Question 1, a single weighted ES for all 12 studies was calculated. Additionally, individual ESs for each study were calculated. Table 2 provides a summary of the 12 studies included in this synthesis. The total *n* refers to the total number of students who were included in the studies. Table 2 also displays grade level, type of research design, and level of technology usage according to the SAMR model. All studies included in this synthesis were published in peer-reviewed journals, with publication dates ranging from 2010 to 2023. Participants' grade levels ranged from kindergarten to 5. Eight different countries of publication were represented, including the United Sates, Algeria, Netherlands, Spain, Sweden, Japan, Taiwan and Germany. Table 4 provides a summary of the objectives, approaches, findings, and implications of each study, described in detail below.

Table 2 Summary of study characteristics

	Study	Mean ES	Total n	Treat n	Grade/Age	Design	Duration	Tech Description	SAMR Category
1	Chadli et al. (2018)	5.34	52	26	2nd grade	RCT	10 sessions, 90 min	Computer-based worksheets	S
2	de Kock & Harskamp (2014)	0.08	390	280	5th grade	Quasi	50 session, 20 min	Computer-based intervention with hints	A
3	Fede et al. (2013)	0.60	32	16	5th grade	RCT	24 sessions, 45 min	Computer-based intervention with hints	A
4	Gonzalez-Castro et al. (2016)	3.09	216	216	Ages 6-9	RCT	45 sessions, 50 min	Computer-based worksheets	S
5	Hassler-Hallstedt et al. (2018)	0.17	261	131	2nd grade	RCT	100 sessions, 20 min	Tablet-based, adaptive intervention	A
6	Leh & Jitendra (2012)	-0.84	25	13	3rd grade	RCT	15 sessions, 50 min	Computer-based, adaptive intervention	A
7	Tajika et al. (2012)	0.55	139	71	5th grade	RCT	12 sessions, 30 min	Computer-based, adaptive intervention	A
8	Xin et al. (2017)	1.76	17	9	3rd-4th grade	RCT	36 sessions, 25 min	Computer-based adaptive intervention	A
9	Yeh et al. (2019)	0.42	334	209	2nd-3rd grade	RCT	287 tasks, 9 min per task	Computer-based adaptive, gamified intervention with peer/ teacher chat	M
10	Schoppek & Tulis (2010)	3.86	113	57	3rd grade	RCT	7 sessions, 60 min	Computer-based, adaptive intervention	A
11	Kim & Ke (2016)	0.31	132	66	4th grade	RCT	1 session, 30 min	VR-based immersive story and adaptive supports	R
12	Xin et al. (2023)	1.26	17	9	3rd grade	RCT	18 sessions, 20 min	Computer-based, adaptive intervention with virtual tutor	M

Note. Total *n* = total number of students who were included in the study; Treat *n* = number of students who received intervention; RCT = randomized control trials

Table 3 Mean effect sizes and confidence intervals as a function of moderator variables

Moderator Variable	K	Mean ES	SE	95% CI
Grade				
K-2	6	2.48	0.11	0.98 – 1.41
3-5	12	0.81	0.05	0.23 – 0.42
Duration of study				
1 session	1	0.31	0.18	-0.04 – 0.65
7 sessions	1	3.86	0.32	3.23 – 4.48
10 sessions	1	5.34	0.59	4.17 – 6.50
12 sessions	1	0.55	0.17	0.21 – 0.89
15 sessions	1	-0.84	0.42	-1.66 – -0.02
24 sessions	2	0.60	0.26	0.09 – 1.10
36 sessions	2	1.76	0.40	0.78 – 2.37
45 sessions	3	3.09	0.20	2.68 – 3.47
50 sessions	3	0.08	0.07	-0.05 – 0.21
64 sessions	2	0.17	0.13	-0.10 – 0.43
287 sessions	1	0.42	0.42	0.19 – 0.64
Deliverer of instruction				
Teacher	4	1.40	0.07	0.02 – 0.27
Media tool only	13	1.26	0.06	0.62 – 0.87
University or graduate student	2	1.76	0.40	0.78 – 2.37
Grouping of students				
Large group	3	0.08	0.07	-0.05 – 0.21
Small group	4	3.18	0.23	2.89 – 3.79
Individual	12	1.04	0.06	0.49 – 0.74
Type of measure				
Norm referenced	8	1.64	0.09	0.83 – 1.18
Researcher developed	11	1.12	0.05	0.19 – 0.39

K = number of effect sizes

All but one study utilized a randomized controlled trial design, while the remaining study was a quasi-experimental design. The duration of intervention sessions ranged from 1 (Kim & Ke, 2016) to 286 sessions/tasks (Yeh et al., 2019), with intervention times ranging from 8.86 (Yeh et al., 2019) to 90 minutes per session (Chadli et al., 2018). Of the 12 studies, one study was conducted in whole groups (de Kock & Harskamp, 2014), three in small groups (Chadli et al., 2018; Schoppek & Tullis, 2010; Xin et al., 2017), and eight individually (Fede et al., 2013; Gonzalez-Castro et al., 2016; Hassler-Hallstedt et al., 2018; Kim & Ke, 2016; Leh & Jitendra, 2013; Tajika et al., 2013; Xin et al., 2023; Yeh et al., 2019). A teacher delivered the instruction in two studies (Chadli et al., 2018; de Kock & Harskamp, 2014), a university or graduate student in one study (Xin et al., 2017), and a media tool only in the eight remaining studies (Fede et al., 2013; Gonzalez-Castro et al., 2016; Hassler-Hallstedt et al., 2018; Kim & Ke, 2016; Leh & Jitendra, 2013; Tajika et al., 2012; Yeh et al., 2019). Finally, eight studies used researcher-developed measures and five of the 12 studies used norm-referenced tests to assess word problem solving accuracy on pre- and post-test measures.

Table 4 Included Studies by Objectives, Approach, Findings and Implications, and Mean ES

Study	Objectives (RQs, Hypotheses, Goals)	Approach	Findings and Implications	Mean ES
Chadli et al. (2018)	RQs 1.Whether students benefit from computer-assisted stage-based mathematical word problem solving. 2.At which stage of the problem solving model do students encounter difficulties?	Computer-based worksheets	Students excelled in the initial stages of the CAI model but struggled with reviewing, indicating a need for interventions that enhance how students recognize different problems and improve problem solving performance.	5.34
de Kock & Harskamp (2014)	RQs 1. Are students able to make effective use of the metacognitive computer programs implemented by teachers in their mathematics lessons in terms of solving most of the problems correctly and using hints when they do not know how to proceed? 2. Does working with a metacognitive computer program have a positive effect on learning outcomes in terms of the analysis of word problems, solving word problems, and self-monitoring (as an aspect of metacognition)?	Computer-based intervention with hints	The metacognitive computer program improved students' problem solving skills better than traditional textbook approaches, indicating the importance of integrating technology to improve metacognition in classroom instruction.	0.08
Fede et al. (2013)	RQs 1. Do students who received the CA-SBI intervention show higher gains on a subset of MCAS items compared to students who received test prep review? 2. Do students who received the CA-SBI intervention show higher gains on the Process and Application subtest of the GMADE compared to students who received test prep? 3. Do students who received the CA-SBI intervention show stronger rates of growth on examiner-made probes compared to students who received test prep review?	Computer-based intervention with hints	The CA-SBI intervention significantly improved word problem solving outcomes for students compared to traditional test prep, indicating that schema-based instruction may be an effective strategy for enhancing math skills for this population.	0.60
Gonzalez-Castro et al. (2016)	RQs 1. Does the computerized tool provide better results in the mathematics skills than the typical instruction in students with ADHD, MLD, and ADHD+MLD? 2. Is the efficacy of the intervention modulated by the diagnosis?	Computer-based worksheets	The computerized intervention significantly improved math competencies for students with ADHD and MLD, particularly those with MLD, indicating that targeted, technology-based interventions can enhance outcomes for students with these disabilities.	3.09
Hassler-Hallstedt et al. (2018)	Hypotheses 1. Low performing children in second grade participating in math training improve mathematical skills compared with children in the control and placebo conditions. 2. The gained effects will be maintained during the follow-up period. 3. IQ and SES moderate effect of math training; those with lower IQ and lower SES may benefit more from the intervention. 4. WM training in combination with math training lead to a more superior outcome in terms of mathematical skills.	Tablet-based, adaptive intervention	Although the tablet intervention significantly improved arithmetic skills in low-performing students, the effects were reduced over time, indicating the need for longer interventions and adaptive technology to reduce the achievement gap.	0.17
Leh & Jitendra (2012)	Goal To evaluate the effectiveness of CMI and TMI on the word problem solving performance of third-grade students struggling in mathematics while controlling and balancing the key features (e.g., priming the problem structure, use of visual representations) deemed critical to successful word problem solving performance across conditions.	Computer-based adaptive intervention	No statistically significant difference was found between the CMI and TMI conditions, indicating that both methods can be used effectively to support at-risk students.	-0.84

Tajika et al. (2012)	Goal To examine the efficacy of self-explanation for helping elementary school students solve mathematical word problems through computer-based support over one year.	Computer-based adaptive intervention	Computer-based self-explanation support significantly improved students' WPS skills, indicating that integrating metacognitive strategies in educational technology can improve learning outcomes.	0.55
Xin et al. (2017)	RQ 1. What are the effects of the PGBM-COMPS intelligent tutor-assisted intervention program, in reference to traditional teacher-delivered intervention (TDI), on participating students' multiplicative problem solving performance measured by a researcher-developed criterion test and a norm-referenced standardized measure?	Computer-based adaptive intervention	The PGBM-COMPS tutoring intervention significantly improved students' multiplicative problem solving performance compared with the TDI, indicating that web-based tutoring can support students with LDs in math.	1.76
Yeh et al. (2019)	RQs 1. Did the Math Island system facilitate students' mathematics achievement in terms of conceptual understanding, calculating, and word problem solving? In particular, how was the mathematics achievement of the low-achieving students? 2. What was students' levels of interest in mathematics and the system, particularly that of low-achieving students?	Computer-based adaptive, gamified intervention with peer/ teacher chat	Math Island significantly improved math achievement and maintained high interest for high- and low-achieving students, suggesting that game-based learning can enhance engagement in math education.	0.42
Schoppek & Tulis (2010)	RQs 1. What do students gain from a small amount of additional individualized practice? 2. Do all students benefit from individualized practice with MMM in the same way? 3. How well does this version work?	Computer-based, adaptive intervention	The computer-based practice significantly improved students' arithmetic and WPS skills, indicating that personalized approaches in math instruction may lead to greater gains and improve engagement.	3.86
Kim & Ke (2016)	Hypotheses 1. The experimental group (GBL) will exhibit a higher learning achievement in a knowledge test that focuses on applying knowledge of fractions in real-life contexts than the control group (non-GBL) group. 2. The experimental group (GBL) will exhibit a higher score for perceived MQLA on the SIMMS than the control (non-GBL) group.	VR-based immersive story and adaptive supports	The VR-based treatment produced a significant effect on math knowledge test performance. However, no significant difference was found relating to the motivational quality of the activity.	0.31
Xin et al. (2023)	RQs 1. Did participants who received the MBPS intervention outperform the BAU group? Did the MBPS group maintain their performance after the termination of the intervention? 2. Did participants in the MBPS group improve their performance on a transfer measure that was designed to assess students' algebraic knowledge and skills? Did participants in the MBPS group improve their performance on solving problems taken from commercially published math textbooks? 3. Did participants in the MBPS group improve their performance on a distal measure, a standardized test?	Computer-based, adaptive intervention with virtual tutor	The MBPS tutor improved participants' performance beyond the BAU comparison group, indicating that the web-based MBPS tutor can be helpful for students with learning differences in intervention settings as well as inclusive classrooms.	1.26

Overall, word problem solving interventions had a positive effect on word problem solving accuracy across all studies, Hedges' g = 1.34 (K=19, 95% CI of .54 to 2.14). According to Cohen's (1988) criterion, this is a large effect size. A homogeneity statistic Q was computed to determine whether studies shared a common ES. The Q statistic's distribution is similar to that of Chi-square distribution with k-1 degrees of freedom, where k is the number of ESs. There was significant heterogeneity in the findings, Q (df=19) = 438.278, p = .00. Because the commonly reported Q statistic has been criticized, the I^2 statistic (Higgins & Thompson, 2002) was computed, using the following formula:

$$I^2 = \frac{Q - (k-1)}{Q}$$

I^2 indices of 75% are classified as high heterogeneity (e.g., Higgins & Thompson, 2002). The I^2 statistic was .96, which suggests an extremely high proportion of variation in study estimates across the majority of measures. Due to heterogeneity in ES, moderator variables were included in the analysis to determine the variability of ES as a function of variables of interest.

Question 2: Do effect sizes vary as a function of participant and intervention characteristics?

Table 3 displays the Hedges' g mean effect sizes and 95% CI for the following moderator variables: grade, duration of study, minutes per session, deliverer of instruction, grouping of students, and type of measure. When considering the sample characteristics (grade band K-2 and 3-5), there were significant differences in ESs as a function of grade band, $F(1,13) = 4.57$, $p = .05$, $R^2 = .25$. Studies with students in grade K-2 reported significantly higher effect sizes (g = 3.02) than studies in grade 3-5 (g = 1.36).

There were also significant differences in ESs as a function of intervention characteristics. As an example, there were statistical differences as a function of the number of intervention sessions, $F(9,5) = 21.98$, $p = .002$, $R^2 = .98$. Studies with 10 sessions yielded the highest ES (g = 5.34), while studies with 50 sessions yielded the lowest ES (g = 0.24).

Additionally, there were differences in ESs as a function of the number of minutes per intervention session, $F(7,7) = 37.41$, $p = <.0001$, $R^2 = .97$. The intervention with 90 minutes per session yielded the highest ES (g = 5.34). All intervention sessions with 50 minutes or more yielded effect sizes greater than 3. Additionally, there was a significant difference in ES as function of intervention group size, $F(2,12) = 6.28$, $p = .01$, $R^2 = .490$. Interventions delivered in small groups (g = 3.92) had higher effect sizes than individual (g = 1.70) and whole group instruction (g = 0.24). There were no significant differences in ES by who delivered the instruction, $F(2,12) = 0.03$, $p = .97$, $R^2 = .00$. Finally, there were also no significant differences in ESs as a function of type of measure, $F(1,13) = 0.02$, $p = .88$, $R^2 = .00$.

In summary, word problem solving interventions that integrate technology had a positive and large effect for students who are in grades K-5. Effect sizes for interventions that were delivered across 10 sessions and 90 minutes per session yielded the highest effect size. Additionally, interventions delivered in small group instruction yielded the highest effect sizes.

Question 3: How does the level of technology integration as per the SAMR model affect the estimated ESs?

To answer our second research question, we investigated effect sizes as a function of SAMR category. We coded for the occurrence or nonoccurrence of technology usage based on the descriptions of the instruction provided in the manuscripts. There was a significant difference in ESs as a function of the application of the SAMR model, $F(3,11) = 3.45$, $p = .05$, $R^2 = .46$. Studies that included technology usage as substitution yielded the highest ES ($g = 3.49$), followed by augmentation ($g = 1.59$, modification ($g = 0.52$), and finally redefinition ($g = 0.31$).

Based on the nature of our inquiry and categorization of studies by technology integration, we report our descriptions of each of the studies included in the analysis under the SAMR categories: (a) substitution, (b) augmentation, (c) modification, and (d) redefinition. Table 1 provides summarized descriptions, study examples, and rationale for the categorization of these studies, with each described in detail.

Enhancement

As referenced in Figure 1 substitution and augmentation fall under the broader category of *enhancement*. For the purposes of our study *enhancement* refers to an application of technology which does not constitute significant task re-design. In the enhancement category, technology may be used as a direct substitution for a non-digital activity or an augmentation which adds a layer of functional improvement.

In our search we discovered two such manuscripts which described substitutions and met our criteria. In one instance (Chadli et al., 2018) substitution was a function of the study itself, testing the digital modality against a paper and pencil model. The other (Gonzalez-Castro et al., 2016) included graphic representations of numeric values similar to what may be found on a traditional worksheet and followed a prescribed progression of math concepts.

Augmentation was the most common designation found in this study with seven of the eleven studies falling within this designation. Elements which defined this designation included the option of “hints” or guided instruction for students (de Kock & Harskamp, 2014; Leh & Jitendra, 2013; Tajika et al., 2012); the inclusion of audio/visual components

(Fede et al., 2013), and adaptive instruction (Hassler-Hallstedt et al., 2018; Schoppek & Tulis, 2010; Xin et al., 2017).

Chadli et al.'s 2018 work "An empirical investigation into student's mathematical word-based problem solving process: A computerized approach" divided 52 second graders from the Tiaret Province in west Algeria who scored at or below the 50th percentile in their first and second semester mathematics courses into two equal, randomly assigned groups. The baseline group attempted to solve word problems in a business as usual format while the experimental group was given a computer-based application to frame the steps of problem solving. Groups met for ten 90-minute sessions, and found that the experimental group produced statistically significantly improved scores between their pre and post-tests compared with the baseline group. This study was coded as a substitution because the computer-based scaffolding provided a static framework for students to guide their work.

The other substitution-categorized study included in this meta-analysis was Gonzalez-Castro et al.'s 2016 "Improvement of word problem solving and basic mathematics competencies in students with attention deficit/hyperactivity disorder and mathematical learning difficulties". In this study, 216 students aged 6-9 were placed into groups based on their learning differences. 72 students with ADHD, 82 with a math difficulty, and 62 who had both ADHD and a math disability were included in the study with each group divided into control and experimental group with statistically comparable demographics such as mean age and mean IQ. Experimental groups for all three categories were given a computer-based framework for addressing word problem solving where the control group utilized business as usual practices for 45 fifty-minute sessions. All three of the control groups were found to have statistically significant improvements greater than their control group peers in post intervention testing with the MLD-only group showing the greatest growth.

The augmentation category was the most prevalent with more than 63% of included studies employing it. De Kock and Harskamp (2014) studied the effectiveness of a computer-based intervention over the course of 10 weeks with fifth grade students in the Netherlands. 280 students received the intervention which supported the students with metacognitive hints in addition to their math textbooks. 110 students in the control condition were not offered the program, but did utilize the same textbook. Findings reported that intervention students used more hints initially, and out-performed their control-group peers in the areas of analyzing word problems, solving word problems, and self-monitoring on a post-test assessment.

Fede et al. (2013) examined the differences in computer-assisted, schema-based instruction and traditional classroom instruction for low-performing fifth grade students as identified as scoring below the 30th percentile in the process and application section of a standardized mathematics assessment. 32 students were randomly assigned into two groups of 16. The intervention employed a software called *GO Solve Word Problems* which "teaches students how to better understand word problems before solving them by illustrating the underlying mathematical models or situations represented in arithmetic word problems" (Fede et al., 2013, p. 13) and was administered over 50 twenty-minute sessions. Although the experimental group did out-perform the control in the pre and post-test analysis, those differences were found to not be statistically significant.

Hassler-Hallstedt et al. (2018) studied the effects of a tablet-based word problem solving intervention for low-performing second grade students in Sweden. Students were assigned to the control (n = 52), reading placebo (n = 78), math intervention (n = 76) or math plus working memory training (n = 77 groups). This study also used intelligence quotient (IQ) scores of participants as a moderating variable. The math intervention and intervention plus working memory training groups trained for 20 minutes per day on a program called Chasing Planets which uses animations, visual examples, and audio instructions to support student fluency. The working memory group was given an additional 30 minutes every other day using working memory training exercises. The placebo group spent 20 minutes each day on a tablet-based reading exercise. The two treatment conditions demonstrated medium effect size increases in student performance with fadeout effects in the 6- and 12-month follow up probes. IQ was found to have a significant modifier effect with lower IQ students benefitting more than those with higher IQs.

Leh and Jitendra (2012) compared computer and teacher-mediated word problem solving outcomes for American third grade students who scored at or below the 50th percentile in math on a standardized assessment. Both the computer-mediated group (n = 13) and teacher-mediated group (n = 12) utilized the same curriculum in 15 daily 50-minute instructional periods with the primary differences being the presentation modality and personalized word problems only presented in the computer-mediated treatment group. Researchers found no statistically significant differences in the outcomes between these two groups.

Tajika et al. (2012) sought to explore the role of computer-based support to extend word problem solving strategies to include self-explanation. 71 fifth grade students in Japan were placed into a treatment group whose intervention included the computer-mediated self-explanation step to word problem solving, while an additional 62 students served as a control without the self-explanation, computer-mediated intervention over the course of 12 sessions of thirty minutes each, delivered once per week. The researchers conducted assessments throughout the process and although there was not a difference at every testing increment, the computer-mediated group did produce statistically significantly better results by the end of the study and in the transfer assessment.

Xin et al. (2017) used a computer-based tutoring intervention to examine its impact for students identified as with or at-risk for math disabilities in third and fourth grade. Nine students were assigned to the treatment group using the Please Go Bring Me-Conceptual Model-Based Problem Solving (PGBM-COMPS) intervention, while the remaining eight students received a teacher-delivered intervention as a control. On a researcher-developed assessment, the two groups showed no statistically significant difference in outcome.

Schoppek and Tullis (2010) sought to investigate the effectiveness of individualized, adaptive, computer-assisted practice for word problem solving. 94 German students from four different third grade classes were divided into control (n = 56) and training (n = 54) groups for 7 hour-long sessions delivered in a once per week format. Researchers found that the training group had statistically significantly higher scores on a standardized assessment, and these improvements persisted on a follow-up assessment three months later. The authors attribute this difference to the individualization possible in the computer-assisted practice.

Transformation

While substitution and augmentation are classified together as *enhancements*, the remaining modification and redefinition categories are grouped together as *transformations* which are defined by their ability to provide significant task redesign. Both modification and redefinition applications are more interactive and allow for new curricular options, but to qualify as a redefinition, the application allows for something previously inconceivable or impossible without the technology such as collaboration across great distances or the application of VR technology to transport students to another space.

Two interventions from the selected literature were identified that qualified as modifications. In Yeh et al.'s 2019 study, students engaged in a gamified curriculum that included in-game rewards but also opportunities for students to communicate with peers and teachers in the virtual space. 215 second and third grade Taiwanese students participated in a treatment group, and another 125 second and third graders in a similar school which did not receive the intervention as a control. For two years the treatment group used a researcher-created, game-based learning environment called Math Island which gamified student progress and allowed them scaffolded instruction, support, and feedback in addition to motivating game elements and rewards made available to students as a supplement to take home and to use at school. At the end of the two-year study, post-test data suggests the Math Island intervention was most effective for students identified as low-achieving in math compared to the control school, and most significantly in the area of word problem solving for all students in the treatment group. Both high- and low-achieving students from the treatment group showed increased interest in math following the intervention.

Xin et al. (2023) investigated a model-based problem solving intervention for third graders who struggle in math. The remaining 8 students served as a control and received traditional teacher-led math instruction in an afterschool setting. The group that received virtual tutoring showed a greater increase in performance compared with the teacher-led group. However, only 56% of students improved on the standardized test. Although this was better than the teacher-led group, the authors suggest that further instruction could be necessary to help students problem solving skills.

Only one study which met the inclusion criteria applied technology in a way which could be considered a redefinition which employed a previously inconceivable task. In the study (Kim & Ke, 2016), 132 fourth grade students engaged in an intervention which immersed them in a virtual reality environment and with in-game challenges, stories, rewards, and math supports

Discussion and implications

The purpose of this meta-analysis was to determine if instruction that integrates technology within word problem solving interventions are effective for improving word problem solving accuracy in students in elementary grades and if so, determine if effect sizes vary as a function of participant and/or intervention characteristics. Three important findings emerged. First, math interventions that integrate technology had a positive effect on word problem solving accuracy overall. Second, there were no significant differences as a function of grade band. Interventions that were delivered across 10 sessions and 90 minutes yielded the highest effect sizes, though all interventions with 50 or more minutes per session yielded high effect sizes. Interventions delivered in small groups yielded the highest effect sizes. Finally, there were no significant differences in effect sizes as a function of the level of technology integration, though this statement is qualified with a low number of studies in the modification and redefinition categories.

We will now address the three questions that directed this study.

Question 1: Are math interventions that incorporate technology effective for improving the math problem solving of kindergarten to grade 6 students with math difficulties?

Although not all of these studies included specific outcomes for students with or at risk for math disabilities, those that did reported a positive effect on word problem solving accuracy overall. Previous work in this area has indicated mixed results for technology integration in math interventions for students with disabilities (Seo & Bryant, 2009), but more recent work has focused more specifically on the application of technology in the classroom (Crompton & Burke, 2020; McClain & North, 2019) particularly using models such as SAMR to assess the ways technology is integrated. Particularly for students with disabilities, this work is still slowly emerging, and the current body of literature is limited.

For this reason, we included studies that integrated technology in math and word problem solving instruction for students broadly including researcher-chosen inclusion criteria such as standardized test scores below the 30th (de Kock & Harskamp, 2014; Fede et al., 2013) or 50th (Chadli et al., 2018; Leh & Jitendra, 2012) percentile, identification of a math disability or difficulty (Gonzalez-Castro et al., 2016; Xin et al., 2017), and intelligence quotient scores (Hassler-Hallstedt et al., 2018).

Researchers and practitioners seeking to apply these findings in math interventions using technology for students should be advised of considerations for instructional design and the SAMR model. Although not every study included in this meta-analysis analyzed populations of students in kindergarten through grade 6 with math difficulties, those that did reported improvement for math problem solving using technology interventions, indicating that technology-based interventions may be beneficial to the improvement of WPS skills for these students.

Question 2: Do effect sizes vary as a function of participant and intervention characteristics?

Interventions that were delivered across 10 sessions and 90 minutes yielded the highest effect sizes, though all interventions with 50 or more minutes per session yielded high effect sizes. Finally, interventions delivered in small groups yielded the highest effect sizes. These findings suggest that a minimum of 50-minute sessions are needed to implement an intervention which integrates technology in a meaningful way. This presents an additional challenge to math educators who unfortunately face pressing time constraints in the elementary classroom with disproportionate attention given to English Language Arts (ELA). Although there were no significant differences as a function of grade level, interventions which had students working in small group and social arrangements also produced statistically significantly higher effect sizes than those who assigned students to work independently or as a large group.

In terms of implications for instructional design with technology, these findings suggest that a minimum of 50 minutes per session is necessary to produce results with 90-minute sessions reporting the largest effect sizes. Optimally these were delivered over 10 intervention sessions. Ideal grouping for these technology interventions occurred in small group settings compared with individual or large group delivery. Additionally, studies that included students in grades K-2 reported significantly higher ESs than studies in grades 3-5. This seems to imply that attention to effective word problem solving instruction that incorporates technology in early grades is crucial to begin to build a strong mathematical foundation.

Question 3: How does the level of technology integration as per the SAMR model affect the estimated ESs?

There was a significant difference in ESs as a function of the application of the SAMR model, $F(3,11) = 3.45$, $p = .05$, $R^2 = .46$. Studies that included technology usage as substitution yielded the highest ES ($g = 3.49$), followed by augmentation ($g = 1.59$), modification ($g = 0.52$), and finally redefinition ($g = 0.31$). This portion of the discussion is divided into the two broader categories of the SAMR model as well as a discussion of the framework itself.

Enhancement. Studies that included technology usage as substitution yielded the highest ES ($g = 3.49$), followed by augmentation ($g = 1.59$). With the limited number of studies included (9 total in enhancement), we recommend interpreting these differences with caution as detailed in the “SAMR Framework” section below.

Transformation. A limitation of comparison by SAMR in the case of elementary word problem solving interventions relates to the limited number of studies available. Only two studies were found which met the criteria for modification, and only one which was categorized as a redefinition. Although these studies reported effect sizes greater than their business-as-usual counterparts, they were found to have the lowest comparable effect sizes (modification ($g = 0.52$) and redefinition ($g = 0.31$)).

The SAMR Framework. Although more broad studies have reported evidence of varying effect sizes as a function of the SAMR framework (Setiyawati et al., 2023; McClain & North, 2019), limited available studies in the area of word problem solving for elementary students, did not allow for a thorough comparison- particularly at the modification and redefinition levels where only two studies met the criteria for modification, and only one redefinition. An additional challenge of categorizing and assessing the effect sizes at the SAMR level comes from the lack of instructional context inherent in the model. This critique is common in other studies (Bicalho et al., 2023; Hamilton et al., 2016) which call for an extension of the model to better reflect the appropriateness of technology applications within instruction. While it may be tempting to conceptualize the SAMR model as a continuum where a higher level of technology integration is “better”, effects may be linked more to utilizing technology *as appropriate* to the instruction.

In terms of the SAMR framework evaluation, it is important for all stakeholders and designers to consider more than which level of the SAMR framework a technology application would fall, and to instead attend most closely to the appropriateness of these choices for the instructional goals, situations, and student populations. Although the SAMR framework does provide a lens through which educators may reflect on their technology application, it should not be the only consideration. It is important for educators to consider not only the digital tool being used but also the function and/or application of this tool.

Conclusion

Limitations

Although this synthesis provided information about technology integration within math instruction in the elementary grades, the findings should be interpreted with caution. First, only group studies that presented adequate data for comparison, published in peer-reviewed articles were included in our study, which excludes unpublished work and single-subject designs. These more stringent criteria, coupled with the narrow criteria of only analyzing word problem solving interventions, may reduce the generalizability of these findings.

In terms of the SAMR model, studies that integrate technology to transform instruction (modification and redefinition) were extremely limited (3 total). The extent to which we could make meaningful comparisons across categories was limited. Within the literature, the SAMR model has come into criticism for its inability to consider the context such as situational, content, teacher preparation, and learning objectives (Blundell et al., 2022; Hamilton et al., 2016). Also, the descriptions of technology application within each study can contain nebulous or unclear language. In order to categorize these studies, some assumptions had to be made to place them on the SAMR framework and determine effect sizes. Suggestions for improvement of technology classification include a variety of ways to reconsider these factors such as measuring from situational established baselines (Blundell et al., 2022), increasing context sensitivity, and increased flexibility (Hamilton et al., 2016). Additionally, clarification on the type of technology (e.g., hardware, software) may be a potential source of bias in the present study, and analyzing studies by these measures may be helpful in the further understanding of how these tools are best used in math interventions.

Abbreviations

SAMR: Substitution, Augmentation, Modification, Redefinition; LD: Learning Disabilities; CAI: Computer Assisted Instruction; WPS: Word Problem Solving; ELA: English Language Arts; ES: Effect Size; IQ: Intelligence Quotient.

Authors' contributions

Joelle Prate conducted the literature search, contributed to coding of the studies, wrote the introduction, study description, results narrative study descriptions, discussion, and implications sections, developed Tables 1, 2, and 4, and Figure 2, and developed the SAMR Categorization framework

Jennifer Kong conducted the statistical analysis, wrote the methods and results sections, provided methodological and project support.

Lamiya Hoque conducted the literature search, contributed to coding of the studies, and provided writing support in the Discussion and Implications section of the manuscript.

Trisha Sugita provided writing support in the Discussion and Implications section of the manuscript and editing support throughout.

Tanya Belom provided writing support in the Discussion and Implications section of the manuscript and editing support throughout.

Authors' information

JP is a Ph.D. candidate in Teacher Education and former K-12 science teacher.

JK is an Assistant Professor in Special Education and a former special education teacher.

LH is a Ph.D. student in Education.

TS is an Associate Professor in Special Education and Director of Teacher Education programs.

TB is a current school psychologist and Ph.D. student in School Psychology.

Funding

This study was supported by the Attallah College of Educational Studies Solutions Grant, which supported the hiring of these graduate students.

Availability of data and materials

Not applicable.

Declarations

Competing interests

The authors declare that they have no competing interests.

Author details

All authors are affiliated with the Attallah College of Educational Studies at Chapman University, United States.

Received: 28 March 2024 Accepted: 16 October 2024

Published online: 1 January 2026 (Online First: 3 March 2025)

References

Andriani, T., Ulya, N. H. A., Alfiana, T. P., Solicha, S., Hafsari, S. B. A., & Ishartono, N. (2022). Improving student's critical thinking skill in mathematics through GeoGebra-based flipped learning during pandemic Covid-19: An experimental study. *Journal of Medives: Journal of Mathematics Education IKIP Veteran Semarang*, *6*(1), 49. https://doi.org/10.31331/medivesveteran.v6i1.1901

Benavides-Varela, S., Callegher, C. Z., Fagiolini, B., Leo, I., Altoè, G., & Lucangeli, D. (2020). Effectiveness of digital-based interventions for children with mathematical learning difficulties: A meta-analysis. *Computers & Education*, *157*, 103953. https://doi.org/10.1016/j.compedu.2020.103953

Bicalho, R., Coll, C., Engel, A., & Oliveira, M. (2023). Integration of ICTs in teaching practices: Propositions to the SAMR model. *Educational Technology Research and Development*, *71*, 563–578. https://doi.org/10.1007/s11423-022-10169-x

Blundell, C. N., Mukherjee, M., & Nykvist, S. (2022). A scoping review of the application of the SAMR model in research. *Computers and Education Open*, *3*, 100093. https://doi.org/10.1016/j.caeo.2022.100093

Boonen, A. J. H., De Koning, B. B., Jolles, J., & Van Der Schoot, M. (2016). Word problem solving in contemporary math education: A plea for reading comprehension skills training. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00191

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Wiley. https://doi.org/10.1002/9780470743386

Bumgarner, E., Martin, A., & Brooks-Gunn, J. (2013). Approaches to learning and Hispanic children's math scores: The moderating role of English proficiency. *Hispanic Journal of Behavioral Sciences*, *35*(2), 241–259. https://doi.org/10.1177/0739986312473580

Carlberg, C. G., & Kavale, K. A. (1980). The efficacy of special versus regular class placement for exceptional children: A meta-analysis. *Journal of Special Education*, *14*(3), 295–309. https://doi.org/10.1177/002246698001400304

Cawley, J., Parmar, R., Foley, T. E., Salmon, S., & Roy, S. (2001). Arithmetic performance of students: Implications for standards and programming. *Exceptional Children*, *67*(3), 311–328. https://doi.org/10.1177/001440290106700302

Chadli, A., Tranvouez, E., Dahmani, Y., Bendella, F., & Belmabrouk, K. (2018). An empirical investigation into student's mathematical word-based problem-solving process: A computerized approach. *Journal of Computer Assisted Learning*, *34*(6), 928–938. https://doi.org/10.1111/jcal.12301

Cheung, A., & Slavin, R. E. (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

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159. https://doi.org/10.1037/0033-2909.112.1.155

Crompton, H., & Burke, D. (2020). Mobile learning and pedagogical opportunities: A configurative systematic review of preK-12 research using the SAMR framework. *Computers & Education*, *156*, 103945. https://doi.org/10.1016/j.compedu.2020.103945

de Kock, W. D., & Harskamp, E. (2014). Can teachers in primary education implement a metacognitive computer programme for word problem solving in their mathematics classes? *Educational Research and Evaluation*, *20*(3), 231–250. https://doi.org/10.1080/13803611.2014.901921

Dennis, M. S., Sharp, E., Chovanes, J., Thomas, A., Burns, R. M., Custer, B., & Park, J. (2016). A meta-analysis of empirical research on teaching students with mathematics learning difficulties. *Learning Disabilities Research and Practice*, *31*(3), 156–168. https://doi.org/10.1111/ldrp.12107

Fede, J. L., Pierce, M. E., Matthews, W. J., & Wells, C. S. (2013). The effects of a computer-assisted, schema-based instruction intervention on word problem-solving skills of low-performing fifth grade students. *Journal of Special Education Technology*, *28*(1), 9–21. https://doi.org/10.1177/016264341302800102

Fletcher, J. M., Epsy, K. A., Francis, P. J., Davidson, K. C., Rourke, B. P., & Shaywitz, S. E. (1989). Comparison of cutoff and regression-based definitions of reading disabilities. *Journal of Learning Disabilities, 22*(6)*,* 334–338. https://doi.org/10.1177/002221948902200603

Fuchs, D., Mock, D., Morgan, P. L., & Young, C. L. (2003) Responsiveness-to-intervention: Definitions, evidence, and implications for the learning disabilities construct. *Learning Disabilities Research and Practice, 18*(3)*,*157–171. https://doi.org/10.1111/1540-5826.00072

Fuchs, L. S., & Fuchs, D. (2007) Mathematical problem solving. In D. B. Berch & M. M. Mazzocco (Eds.), *Why is math so hard for some children? The nature and origins of mathematical learning difficulties and disabilities* (pp. 397–414). Paul H. Brookes Publishing Co.

*Fuchs, L. S., Fuchs, D., & Prentice, K. (2004). Responsiveness to mathematical problem-solving instruction comparing students at risk of mathematics disability with and without risk of reading disability. *Journal of Learning Disabilities*, *37*(4), 293–306. https://doi.org/10.1177/00222194040370040201

Fuchs, L. S., Fuchs, D., Compton, D. L., Powell, S. R., Seethaler, P. M., Capizzi, A. M., Schatschneider, C., & Fletcher, J. M. (2006). The cognitive correlates of third-grade skill in arithmetic, algorithmic computation, and arithmetic word problems. *Journal of Educational Psychology*, *98*(1), 29–43. https://doi.org/10.1037/0022-0663.98.1.29

*Fuchs, L. S., Fuchs, D., Craddock, C., Hollenbeck, K. N., Hamlett, C. L., & Schatschneider, C. (2008a). Effects of small-group tutoring with and without validated classroom instruction on at-risk students' math problem solving: Are two tiers of prevention better than one? *Journal of Educational Psychology*, *100*(3), 491–509. https://doi.org/10.1037/0022-0663.100.3.491

*Fuchs, L. S., Fuchs, D., Finelli, R., Courey, S. J., & Hamlett, C. L. (2004a). Expanding schema-based transfer instruction to help third graders solve real-life mathematical problems. *American Educational Research Journal*, *41*(2), 419–445. https://doi.org/10.3102/00028312041002419

*Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Appleton, A. C. (2002). Explicitly teaching for transfer: Effects on the mathematical problem-solving performance of students with mathematics disabilities. *Learning Disabilities Research & Practice*, *17*(2), 90–106. https://doi.org/10.1111/1540-5826.00036

*Fuchs, L. S., Fuchs, D., Prentice, K., Burch, M., Hamlett, C. L., Owen, R., Hosp, M., & Jancek, D. (2003a). Explicitly teaching for transfer: Effects on third-grade students' mathematical problem solving. *Journal of Educational Psychology*, *95*(2), 293–305. https://doi.org/10.1037/0022-0663.95.2.293

*Fuchs, L. S., Fuchs, D., Prentice, K., Burch, M., Hamlett, C. L., Owen, R., & Schroeter, K. (2003b). Enhancing third-grade students' mathematical problem solving with self-regulated learning strategies. *Journal of Educational Psychology*, *95*(2), 306–315. https://doi.org/10.1037/0022-0663.95.2.306

*Fuchs, L. S., Fuchs, D., Prentice, K., Hamlett, C. L., Finelli, R., & Courey, S. J. (2004b). Enhancing mathematical problem solving among third-grade students with schema-based instruction. *Journal of Educational Psychology*, *96*(4), 635–647. https://doi.org/10.1037/0022-0663.96.4.635

*Fuchs, L. S., Seethaler, P. M., Powell, S. R., Fuchs, D., Hamlett, C. L., & Fletcher, J. M. (2008b). Effects of preventative tutoring on the mathematical problem solving of third-grade students with math and reading difficulties. *Exceptional Children*, *74*(2), 155–173. https://doi.org/10.1177/001440290807400202

Gersten, R., & Baker, S. (2000). What we know about effective instructional practices for English-language learners. *Exceptional Children*, *66*(4), 454–470. https://doi.org/10.1177/001440290006600402

Gersten, R., Chard, D. J., Jayanthi, M., Baker, S. K., Morphy, P., & Flojo, J. (2009). Mathematics instruction for students with learning disabilities: A meta-analysis of instructional components. *Review of Educational Research*, *79*(3), 1202–1242. https://doi.org/10.3102/0034654309334431

Glass, G. V. (1977). Integrating findings: The meta-analysis of research. *Review of Research in Education*, *5*(1), 351–379. https://doi.org/10.3102/0091732X005001351

Gonzalez-Castro, P., Cueli, M., Areces, D., Pérez, C. R., & Sideridis, G. D. (2016). Improvement of word problem solving and basic mathematics competencies in students with attention deficit/hyperactivity disorder and mathematical learning difficulties. *Learning Disabilities Research and Practice*, *31*(3), 142–155. https://doi.org/10.1111/ldrp.12106

*Griffin, C. C., & Jitendra, A. K. (2009). Word problem-solving instruction in inclusive third-grade mathematics classrooms. *The Journal of Educational Research*, *102*(3), 187–202. https://doi.org/10.3200/JOER.102.3.187-202

Griffin, S. A., Case, R., & Siegler, R. S. (1994). Rightstart: Providing the central conceptual prerequisites for first formal learning of arithmetic to students at risk for school failure. In K. McGilly (Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice* (pp. 24–49). MIT Press.

Hamilton, E. R., Rosenberg, J. M., & Akcaoğlu, M. (2016). The substitution augmentation modification redefinition (SAMR) model: A critical review and suggestions for its use. *TechTrends*, *60*(5), 433–441. https://doi.org/10.1007/s11528-016-0091-y

Hassler-Hallstedt, M. H., Klingberg, T., & Ghaderi, A. (2018). Short and long-term effects of a mathematics tablet intervention for low performing second graders. *Journal of Educational Psychology*, *110*(8), 1127–1148. https://doi.org/10.1037/edu0000264

Higgins, J. P., & Thomspon, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, *21*(11), 1539–1558. https://doi.org/10.1002/sim.1186

Hilton, J. T. (2016). A case study of the application of SAMR and TPACK for reflection on technology integration into two social studies classrooms. *The Social Studies*, *107*(2), 68–73. https://doi.org/10.1080/00377996.2015.1124376

Jitendra, A. K., & Hoff, K. (1996). The effects of schema-based instruction on the mathematical word-problem-solving performance of students with learning disabilities. *Journal of Learning Disabilities*, *29*(4), 422–431. https://doi.org/10.1177/002221949602900410

Jitendra, A. K., Dupuis, D. N., Rodriguez, M. C., Zaslofsky, A. F., Slater, S., Cozine-Corroy, K., & Church, C. (2013a). A randomized controlled trial of the impact of schema-based instruction on mathematical outcomes for third-grade students with mathematics difficulties. *The Elementary School Journal*, *114*(2), 252–276. https://doi.org/10.1086/673199

Jitendra, A. K., Griffin, C. C., Haria, P., Leh, J., Adams, A., & Kaduvettoor, A. (2007). A comparison of single and multiple strategy instruction on third-grade students' mathematical problem solving. *Journal of Educational Psychology*, *99*(1), 115–127. https://doi.org/10.1037/0022-0663.99.1.115

Jitendra, A. K., Griffin, C. C., McGoey, K., Gardill, M. C., Bhat, P., & Riley, T. (1998). Effects of mathematical word problem solving by students at risk or with mild disabilities. *The Journal of Educational Research*, *91*, 345–355. https://doi.org/10.1080/00220679809597564

Jitendra, A. K., Rodriguez, M., Kanive, R., Huang, J. P., Church, C., Corroy, K. A., & Zaslofsky, A. (2013b). Impact of small-group tutoring interventions on the mathematical problem solving and achievement of third-grade students with mathematics difficulties. *Learning Disability Quarterly*, *36*(1), 21–35. https://doi.org/10.1177/0731948712457561

Kelsch, J., & Wang, T. (2021). An action research of using SAMR to guide blended learning adoption during COVID-19. In W. Jia, Y. Tang, R. S. T. Lee, M. Herzog, H. Zhang, T. Hao & T. Wang (Eds.), *Emerging Technologies for Education. SETE 2021. Lecture Notes in Computer Science, vol 13089* (pp. 207–218). Springer, Cham. https://doi.org/10.1007/978-3-030-92836-0_18

Kim, H., & Ke, F. (2016). Effects of game-based learning in an OpenSim-supported virtual environment on mathematical performance. *Interactive Learning Environments*, *25*(4), 543–557. https://doi.org/10.1080/10494820.2016.1167744

Kim, J. S., Gilbert, J., Yu, Q., & Gale, C. (2021). Measures matter: A meta-analysis of the effects of educational apps on preschool to grade 3 children's literacy and math skills. *AERA Open*, *7*, 233285842110041. https://doi.org/10.1177/23328584211004183

Kim, S. J., & Xin, Y. P. (2022). A synthesis of computer-assisted mathematical word problem-solving instruction for students with learning disabilities or difficulties. *Learning Disabilities: A Contemporary Journal*, *20*(1), 27–45.

Kiru, E. W., Doabler, C. T., Sorrells, A. M., & Cooc, N. (2017). A synthesis of technology-mediated mathematics interventions for students with or at risk for mathematics learning disabilities. *Journal of Special Education Technology*, *33*(2), 111–123. https://doi.org/10.1177/0162643417745835

Kong, J. E., & Swanson, H. L. (2019). The effects of a paraphrasing intervention on word problem-solving accuracy of English learners at risk of mathematic disabilities. *Learning Disability Quarterly*, *42*(2), 92–104. https://doi.org/10.1177/0731948718806659

Kong, J. E., Yan, C., Serceki, A., & Swanson, H. L. (2021). Word-problem-solving interventions for elementary students with learning disabilities: A selective meta-analysis of the literature. *Learning Disability Quarterly*, *44*(4), 248–260. https://doi.org/10.1177/0731948721994843

Kroesbergen, E. H., & Van Luit, J. E. H. (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

Leh, J. M., & Jitendra, A. K. (2012). Effects of computer-mediated versus teacher-mediated instruction on the mathematical word problem-solving performance of third-grade students with mathematical difficulties. *Learning Disability Quarterly*, *36*(2), 68–79. https://doi.org/10.1177/0731948712461447

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), 1388–1408. https://doi.org/10.1037/edu0000453

McClain, A., & North, T. (2021). Effect of technology integration on middle school math proficiency: A multiple linear regression study. *International Journal of Education in Mathematics, Science, and Technology*, *9*(4), 557–570. https://doi.org/10.46328/ijemst.1456

McNeish, D. (2017). Small sample methods for multilevel modeling: A colloquial elucidation of REML and the Kenward-Roger correction. *Multivariate Behavioral Research*, *52*(5), 661–670. https://doi.org/10.1080/00273171.2017.1344538

Moran, A. S., Swanson, H. L., Gerber, M. M., & Fung, W. (2014). The effects of paraphrasing interventions on problem-solving accuracy for children at risk for math disabilities. *Learning Disabilities Research & Practice*, *29*(3), 97–105. https://doi.org/10.1111/ldrp.12035

Myers, J. A., Witzel, B. S., Powell, S. R., Liu, H., Pigott, T. D., Xin, Y. P., & Hughes, E. M. (2022). A meta-analysis of mathematics word-problem solving interventions for elementary students who evidence mathematics difficulties. *Review of Educational Research*, *92*(5), 695–742. https://doi.org/10.3102/00346543211070049

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. National Council of Teachers of Mathematics.

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards for Mathematics*. National Governors Association Center for Best Practices & Council of Chief State School Officers.

National Research Council. (2001). *Adding it up: Helping children learn mathematics*. The National Academies Press. https://doi.org/10.17226/9822

NWEA. (2023). *Quick references*. NWEA. https://teach.mapnwea.org/impl/maphelp/Content/AboutMAP/QuickReferences.htm

Nyayu, S. Y., Setiawan, H., & Sulaiman, M. (2019). The use of technology integration SAMR model in teaching English. *Ideas: Journal on English Language and Learning, Linguistics and Literature*, *7*(1). https://doi.org/10.24256/ideas.v7i1.720

Orosco, M. J., Swanson, H. L., O'Connor, R., & Lussier, C. (2013). The effects of dynamic strategic math on English language learners' word problem solving. *The Journal of Special Education*, *47*(2), 96–107.

*Owen, R. L., & Fuchs, L. S. (2002). Mathematical problem-solving strategy instruction for third-grade students with learning disabilities. *Remedial and Special Education*, *23*, 268–278. https://doi.org/10.1177/07419325020230050201

Powell, S. R., Berry, K. A., & Tran, L. M. (2020). Performance differences on a measure of mathematics vocabulary for English learners and non-English learners with and without mathematics difficulty. *Reading & Writing Quarterly*, *36*(2), 124–141. https://doi.org/10.1080/10573569.2019.1677538

Puentedura, R. (2006, November 26). *Transformation, technology, and education in the state of Maine* [Web log post]. http://www.hippasus.com/rrpweblog/archives/2006_11.html

Puentedura, R. (2013, May 29). *SAMR: Moving from enhancement to transformation* [Web log post]. http://www.hippasus.com/rrpweblog/archives/000095.html

Ran, H., Kasli, M., & Secada, W. G. (2020). 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

Romrell, D., Kidder, L. C., & Wood, E. R. (2014). The SAMR model as a framework for evaluating learning. *Online Learning*, *18*(2). https://doi.org/10.24059/olj.v18i2.435

Rosenthal, R. (1994). Parametric measures of effect size. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 231–244). Sage.

Salomon, G., & Perkins, D. N. (1998). Chapter 1: Individual and social aspects of learning. *Review of Research in Education*, *23*(1), 1–24. https://doi.org/10.3102/0091732x023001001

Schoppek, W., & Tulis, M. (2010). Enhancing arithmetic and word-problem solving skills efficiently by individualized computer-assisted practice. *Journal of Educational Research*, *103*(4), 239–252. https://doi.org/10.1080/00220670903382962

Seo, Y., & Bryant, D. P. (2009). Analysis of studies of the effects of computer-assisted instruction on the mathematics performance of students with learning disabilities. *Computers & Education*, *53*(3), 913–928. https://doi.org/10.1016/j.compedu.2009.05.002

Setiyawati, S., Basori, B., & Efendi, A. (2023). Substitution, augmentation, modification and redefinition (SAMR) model to improve student's critical thinking ability. *Indonesian Journal of Informatics Education*, *7*(1), 8. https://doi.org/10.20961/ijie.v7i1.60356

Siegel, L. S., & Ryan, E. B. (1989). The development of working memory in normally achieving and subtypes of learning disabled children. *Child Development*, *60*, 973–980. https://doi.org/10.2307/1131037

Slavin, R. E. (1986). Best-evidence synthesis: An alternative to meta-analytic and traditional reviews. *Educational Researcher*, *15*(9), 5–11. https://doi.org/10.3102/0013189x015009005

Suseelan, M., Meng, C. C., & Chin, H. (2022). Research on mathematics problem solving in elementary education conducted from 1969 to 2021: A bibliometric review. *International Journal of Education in Mathematics, Science and Technology*, *10*(4), 1003–1029. https://doi.org/10.46328/ijemst.2198

Swanson, H. L. (2006). Cross-sectional and incremental changes in working memory and mathematical problem solving. *Journal of Educational Psychology*, *98*(2), 265–281. https://doi.org/10.1037/0022-0663.98.2.265

Swanson, H. L., & Hoskyn, M. (1998). Experimental intervention research on students with learning disabilities: A meta-analysis of treatment outcomes. *Review of Educational Research*, *68*(3), 277–321. https://doi.org/10.3102/00346543068003277

Swanson, H. L., & Sachse-Lee, C. (2000). A meta-analysis of single-subject-design intervention research for students with LD. *Journal of Learning Disabilities*, *33*(2), 114–136. https://doi.org/10.1177/002221940003300201

Swanson, H. L., Kong, J. E., Moran, A. S., & Orosco, M. J. (2019). Paraphrasing interventions and problem-solving accuracy: Do generative procedures help English language learners with math difficulties? *Learning Disabilities Research & Practice*, *34*(2), 68–84. https://doi.org/10.1111/ldrp.12194

Swanson, H. L., Kong, J., & Petcu, S. D. (2019). Growth in math computation among monolingual and English language learners: Does the executive system have a role? *Developmental Neuropsychology*, *44*(8), 566–593. https://doi.org/10.1080/87565641.2019.1688328

Swanson, H. L., Lussier, C., & Orosco, M. (2013). Effects of cognitive strategy interventions and cognitive moderators on word problem solving in children at risk for problem solving difficulties. *Learning Disabilities Research & Practice*, *28*(4), 170–183. https://doi.org/10.1111/ldrp.12019

Swanson, H. L., Moran, A., Lussier, C., & Fung, W. (2014). The effect of explicit and direct generative strategy training and working memory on word problem-solving accuracy in children at risk for math difficulties. *Learning Disability Quarterly*, *37*(2), 111–123. https://doi.org/10.1177/0731948713507264

Swanson, H. L., Orosco, M. J., & Lussier, C. M. (2014). The effects of mathematics strategy instruction for children with serious problem-solving difficulties. *Exceptional Children*, *80*(2), 149–168. https://doi.org/10.1177/001440291408000202

Tajika, H., Nakatsu, N., Neumann, E., Nozaki, H., Kato, H., Fujitani, T., & Hotta, C. (2012). Mathematical word problem solving children engaged in computer-based metacognitive support: A longitudinal study. *Japan Society for Educational Technology*, *35*, 11–19.

U.S. Department of Education. (2022). *National Assessment of Educational Progress (NAEP) 2022 Mathematics Assessment*. Institute of Education Sciences, National Center for Education Statistics.

van Garderen, D. (2007). Teaching students with LD to use diagrams to solve mathematical word problems. *Journal of Learning Disabilities*, *40*(6), 540–553. https://doi.org/10.1177/00222194070400060501

Vygotsky, L. S. (1978). *Mind in society*. Harvard University Press.

Wilson, C. L., & Sindelar, P. T. (1991). Direct instruction in math word problems: Students with learning disabilities. *Exceptional Children*, *57*(6), 512–519. https://doi.org/10.1177/001440299105700605

Xin, Y. P., & Jitendra, A. K. (1999). The effects of instruction in solving mathematical word problems for students with learning problems: A meta-analysis. *The Journal of Special Education*, *32*(4), 207–225. https://doi.org/10.1177/002246699903200402

Xin, Y. P., Kim, S. J., Lei, Q., Liu, B. Y., Wei, S., Kastberg, S. E., & Chen, Y. V. (2023). The effect of model-based problem solving on the performance of students who are struggling in mathematics. *The Journal of Special Education*, *57*(3), 181–192. https://doi.org/10.1177/00224669231157032

Xin, Y. P., Tzur, R., Hord, C., Liu, J., Park, J. Y., & Si, L. (2017). An intelligent tutor-assisted mathematics intervention program for students with learning difficulties. *Learning Disability Quarterly*, *40*(1), 4–16. https://doi.org/10.1177/0731948716648740

Yeh, C. Y., Cheng, H. N., Chen, Z. H., Liao, C. C., & Chan, T. W. (2019). Enhancing achievement and interest in mathematics learning through math-island. *Research and Practice in Technology Enhanced Learning*, *14*, 5. https://doi.org/10.1186/s41039-019-0100-9

Zhang, D., & Xin, Y. P. (2012) A follow-up meta-analysis for word-problem-solving interventions for students with mathematics difficulties. *The Journal of Educational Research*, *105*, 303–318. https://doi.org/10.1080/00220671.2011.627397

Zheng, X., Flynn, L. J., & Swanson, H. L. (2013). Experimental intervention studies on word problem solving and math disabilities: A selective analysis of the literature. *Learning Disability Quarterly*, *36*(2), 97–111. https://doi.org/10.1177/0731948712444277

Publisher's Note