

RESEARCH

Open Access

The effects of using exploratory computerized environments in grades 1 to 8 mathematics: a meta-analysis of research

Andrzej Sokolowski^{1*}, Yeping Li^{2†} and Victor Willson^{3†}

Abstract

Background: The process of problem solving is difficult for students; thus, mathematics educators have made multiple attempts to seek ways of making this process more accessible to learners. The purpose of this study was to examine the effect size statistic of utilizing exploratory computerized environments (ECEs) to support the process of word problem solving and exploration in grades 1 to 8 mathematics using meta-analysis.

Results: The findings of 24 experimental pretest and posttest studies (24 primary effect sizes) published in peer-reviewed journals between January 1, 2000, and December 31, 2013, revealed that exploratory computerized environments produced a moderate effect size (effect size (ES) = 0.60, SE = 0.03) when compared to traditional methods of instruction. A 95% confidence interval around the overall mean - $C_{\text{lower}} = 0.53$ and $C_{\text{upper}} = 0.66$ - indicated nonzero population effect and relative precision. A moderator analysis revealed differences among the effects on student achievement between traditional problem solving approaches and ECEs, favoring the latter.

Conclusions: The findings highlight the importance of providing students with opportunities to explore applications of mathematics concepts in classroom especially these supported by computers. A discussion of these findings and their potential impact on improving students' mathematical problem-solving skills, along with implications for further research, follows.

Keywords: Meta-analysis; Problem solving; Explorations; Computers; Student achievement; Grade 1 to 8 mathematics

Background

Advancement in the capabilities of varied technologies has meant that problem solving has become a domain of particular interest. While researchers have examined the use and impact of computers on presenting the content of word problems to learners (e.g., Gerofsky 2004), comparatively little research has focused on learners' use of computers to explore the relations between given problem's variables in the attempt to mathematize and solve it. Despite a wide range of interest in improving students' problem-solving skills, the rate of progress in this domain has not been satisfactory (Forster 2006; Kim & Hannafin 2011). It is still common that 'In typical elementary schools worldwide, the teaching of early arithmetic is

predominantly focused on computational proficiency' (Greer et al. 2007, p. 97).

High interactivity of contemporary software that allows for dynamically representing problem contents are not utilized fully in mathematics classrooms yet and 'The *royal road* to the educational use of computers, software and communication technology within mathematics teaching and learning is still to be discovered, if it ever exists' (Laborde & Strässer 2010, p. 125). Although more elements of explorations, such as measurement and data analysis, have received substantial attention in the newly developed common core standards (Porter et al. 2011), the process of inquiry organization in mathematics classes is not widely researched. Modern technology provides multiple opportunities for applying mathematical structures to quantify system changes (Arthur & Nance 2007; Pead et al. 2007). Moreover 'what technology has done is to return, after an absence of maybe a hundred or

* Correspondence: asokolowski@neo.tamu.edu

†Equal contributors

¹Department of Mathematics and Science, Lone Star College, Tomball, TX 77375, USA

Full list of author information is available at the end of the article

two hundred years, return mathematics back into the full spectrum of science' (Pollak 2007, p. 117).

It is hypothesized that by enriching the process of solving word problem by employing phases such as analysis, model formulation, and model verification, applications of mathematics in real-world settings will be more accessible to students. Consequently, the shift from procedural to conceptual teaching methods in mathematics, advocated by research (e.g., Hiebert 2013), might be initiated. Exploratory computerized environments (ECEs), focusing on supporting mathematical explorations and problem solving, share many commonalities with scientific discovery. Although problem solving can integrate scientific methods, this idea remains absent in the current research. As in science classes, digital technology has been proven to help students with problem-solving techniques (e.g., see Reid et al. 2003; Stern et al. 2008); searching for ways to induce congruent ideas in a mathematics classroom appears to be a promising endeavor.

The role of technologies in mathematics teaching and learning

As emphasis in teaching of mathematics is no longer placed on the knowledge of rules and calculation routines, but on building up mathematical competences such as problem-solving, modeling, and concept formation, this shift of emphasis can be significantly promoted by the use of computers (Hußmann 2007). Technology accessed both offline and online, encompasses a large range of devices, such as calculators, laptops, desktop computers, and interactive whiteboards, as well as a substantial range of software applications, such as graphing devices, data editors, spreadsheets, and dynamic geometry. Technology supporting computations and representations (e.g., geometric figures, graphs of functions, or animations) provides interactive tools and makes key relations for mathematical understanding more transparent and tangible. Several studies (Grouws & Cebulla 2000; Wander & Pierce 2009) suggest that students who develop a mathematical conceptual understanding develop the skills of knowledge transfer that enable them to perform successfully on mathematics applications. Furthermore, technologies enable complex computations and dynamic modeling that lead to more experimental forms of teaching and learning mathematics (Joyce et al. 2009; Li & Li 2009; Passey 2012).

Learning using technology posits certain challenges for software programmers. They must consider what needs to be maintained as visible and how to present the key ideas in learning sequence that will engage and challenge the learners (Bos 2009). It is unlikely that technology *per se* will affect mathematical development in any significant way; instead, what might signify viable effects is how the technology is designed to support learning and how the mathematics tools within the technology connect with the

learners. Passey (2012) formulated several domains in which technology can affect teaching and learning of mathematics: (a) devices, ranging from interactive whiteboards, to desktop computers, to iPods and iPhones; (b) learning activities, ranging from multiple-choice test to interactive videos; (c) the pedagogy embedded, ranging from knowledge acquisition, to finding specific values; (d) the learning settings and interactions, ranging from the teacher and learner alone, to learners in a group or community; and (e) the type of assessment measures used to identify the impact of a specific program on students' learning.

Explorations and problem solving in mathematics

This section summarizes the advantages of utilizing technology, focusing on its use to enhance students' competencies in explorations and problem solving. According to the National Council of Teachers of Mathematics (National Council of Teachers of Mathematics NCTM 2000), 'Technology is essential in learning mathematics' (p. 3). Applying technology to enhance students' problem-solving skills is an ultimate area of interest. A problem's setup and information component expressed in word format are often difficult for students to comprehend, analyze, and solve (Ngu et al. 2014). Such presented problems also have a low motivational factor, which consequently affects the degree to which a learner engages in the solution process. The advancement of multimedia technology has opened up new possibilities for dynamically expressing a problem's contents and extending its analysis. The process can now be externalized and amplified through digital constructions, showing more explicit properties and structures that were previously silent. Several researchers (e.g., Chen 2010; Merrill & Gilbert 2008) have found that students' word-problem-solving skills can be significantly enhanced through the integration of computer technologies.

Explorations

Exploration is defined as the act of searching with the purpose of discovery of information or resources (Kuhn 2007). Explorations give students the opportunity to appreciate applications of mathematical tools in real-life situations (Remillard & Bryans 2004). Explorations can take many forms, ranging from analyzing observed phenomena to undertaking more abstract, open-ended investigations. English and Watters (2005) found that young children are capable of exploring situations beyond those involving simple processes of counts and measures. Furthermore, researchers (e.g., English 2004; Lai & White 2012) have recommended that children receive more exposure to situations where they explore informal notions or where they quantify information, transform quantities, and deal with quantities that cannot be observed. Flum and Kaplan (2006) claimed that explorations engage the learner with

the environment through definite actions of gathering and investigating information. By inducing the use of terms that are central to scientific inquiry, like *observe*, *identify*, and *analyze* (Slough & Rupley 2010), explorations promote the transfer of knowledge, problem-solving skills, and scientific reasoning (Kuhn 2007).

Schwarz and White (2005) advocated that learning about the nature of scientific models and engaging learners in the process of creating and testing models should be a central focus of science education. It is hypothesized that by enriching mathematics curriculum via elements of such inquiry, students' problem-solving skills can be strengthened. These shifts, however, posit certain challenges. At the elementary school level, manipulatives or their interactive replicates have been extensively used to help build conceptual understanding of abstract ideas (Jitendra et al., 2007) and research (Kieran & Hillel 1990; Reimer & Moyer 2005) has proven their positive impact on students' mathematics achievement. Since manipulatives are restricted to geometrical objects, their exploratory character is limited compared to explorations, which provide a far richer context for inducing and practicing more sophisticated mathematical ideas, for instance, the concept of rate of change. Viewed through this lens, interactive exploratory learning environments dominate the previously applied drill-and-practice computer applications, and their use in mathematics has gained momentum over the past decades (Neves et al. 2011). The process of explorations usually concludes with a formulation of a mathematical model. As such, multifaceted cognitive goals are achieved by learners while they undertake such activities. Bleich et al. (2006) concluded that such activities expand students' views of mathematics by integrating mathematics with other disciplines, especially sciences, and engage students in the process of mathematization of real phenomena.

Word problems and problem solving

Situations carrying open questions that challenge learners intellectually (Blum & Niss 1991) are called *word problems* or *story problems*. The general structure of word problems is centered on three components: (a) a setup component, which provides the content (for instance, the place or story problem); (b) an information component, which provides data from which to derive a mathematical model; and (c) a question component, which is the main task directed to the solver (Gerofsky 2004). A setup component of a word problem can be externalized by a static diagram, short video, computer simulation, or physical demonstration. With the exception of static diagrams, all of these means, though not yet commonly used in mathematics classes (Kim & Hannafin 2011), assist with the visualization of problem scenarios and thus help with identifying patterns and formulating their symbolic description. Word problem solving is one area of mathematics

that is particularly difficult because it requires students to analyze content, transfer it into mathematical representations, and map it into mathematical structures. Therefore, it requires not only a retrieval of a particular problem-solving model from learners' long-term memory but also the need to create a novel solution (Zheng et al. 2011).

While word problems are often considered closed tasks that usually involve simplistic responses, problem solving and explorations gravitate toward mathematical modeling that require students to analyze a given situation, build model, and verify the model before applying it. A major contribution to the field of problem solving was made by Polya (1957), who codified four stages of the process: understanding the problem, devising a plan, carrying out the plan, and looking back. Bransford and Stein (1984) extended Polya's approach by developing a five-stage problem-solving model that encompassed identifying the problem, defining goals, exploring possible stages, anticipating outcomes, and looking back and learning. Among these phases, the phase of exploration, which leads the solver to a model formulation and validation, is of the highest importance (Arthur & Nance 2007). Once the model is validated, it can be used for forecasts, decisions, or actions determined by the problem-question component. Francisco and Maher (2005) suggested that the stage of exploration or modeling must exist in the problem-solving process for authentic mathematical problem solving to occur. A similar conclusion was previously reached by Gravemeijer and Doorman (1999), who claimed that 'the role of context problems and of symbolizing and modeling are tightly interwoven' (p. 112). The forms of the mathematical models depend on the problem content. At the elementary and middle school levels, they are often externalized by geometrical objects, ratios, and proportions (National Council of Teachers of Mathematics NCTM 2000). Over the past 30 years, the domain on teaching and learning mathematics applications has undergone modifications reflecting research advancements in the area, one of which is a change in the instructional approach to problem solving: from teaching problem solving, to teaching *via* problem solving (Lester et al. 1994). Some of the main elements of teaching *via* problem solving include (a) providing students with enough information to let them establish the background of the problem, (b) encouraging students to make generalizations about the rules or concepts, and (c) reducing teachers' role to providing guidance during the solution process (Evan & Lappan 1994). Yimer and Ellerton (2009) proposed an inclusion of a prelude phase, called *engagement*, whose role is to increase students' motivation and, consequently, their success rate. According to Kim & Hannafin (2011), these stages represent integral elements of contemporary problem-solving methods.

While technologies' engaging factors in improving student motivation have been often researched (e.g., see Lewis et al. 1998; Niss et al. 2007), their interactive features that enable the learner to hypothesize, make predictions, and verify those predictions have not yet been meta-analyzed at the elementary and middle school levels. This study sought to examine these areas and identify moderators that contribute to increasing effects on students' learning. As a result of this undertaking, we hope to formulate suggestions for a learning environment that will advance students' analytic skills and consequently improve the use of technological tools as a means of explorations in mathematics classes.

Synthesis of findings of prior meta-analytic research

The study of problem-solving methods in the domain of mathematics education has been frequently undertaken by researchers and has especially influenced mathematical practices during the past 30 years (Santos-Trigo 2007). Tall (1986) provided an insightful analysis on how computers can be used for testing mathematical concepts, claiming that 'computer programs can show not only examples of concepts, but also, through dynamic actions, they can show examples of mathematical processes' (p. 5). He questioned the formal approaches to mathematical representations used in textbooks, calling them inaccessible to students, and suggested instead using computer programs to visualize the dynamics of the processes.

Computer programs used to support problem solving were one of the moderators in a meta-analysis on methods of instructional improvement in algebra undertaken by Rakes et al. (2010). Using 82 relevant studies from 1968 through 2008, these researchers extracted five categories, of which two contained technology and computers as a medium supporting instruction and learning. Contrasting procedural and conceptual understanding of mathematics ideas, these scholars found that conceptual understanding as a separate construct, appearing initially in research in 1985, produced the highest effect size when enhanced by computer programs. The timeline of this finding corresponded with the emergence of mathematical explorations, which also exemplify mathematics conceptual understanding. In addition, Rakes et al. (2010) found that technology tools including calculators, computer programs, and java applets produced a moderate 0.30 effect size when compared to traditional methods of instruction. Another systematic review of computer technology use and its effects on K-12 students' learning in mathematics classes between 1990 and 2006 was undertaken by Li and Ma (2010). Analyzing the effects of tutorials, communication media, exploratory environments, tools, and programming language, they concluded that exploratory environments produced the highest ($ES = 1.32$) learning effect size. Li and Ma did not compute the effects of

computer technology on mathematics cognitive domains and type of learning objectives; instead they suggested the need for another review focusing on 'the nature of the use of technology' (p. 235) on student achievement. Yet problems of implementation of pieces of (educational) software, learning environments, and use of communication technology are far from being solved (Laborde & Strässer 2010), and many of the problems relate to improvement of students' problem-solving skills by the use of educational software. Artzt and Armour-Thomas (1992) reported that students' difficulties with problem solving are often attributed to their failure to initiate active monitoring and regulation of their own cognitive processes. Though several potential ways of improving students' initiation of active monitoring have already been researched (e.g., see Grouws & Cebulla 2000; Kapa 2007), in this study, we sought to uncover moderators that had been silent in the previous research. We were especially interested in learning whether extending the exploration stage of the solution process and guiding students through the phases of scientific inquiry could materialize as a construct worthy of investigation. The effect of such organized support might reduce the working memory needs and consequently free students from being overwhelmed at the start. Hart (1996) reported that students find word problems difficult because they lack motivation; thus, presenting word problems in an engaging format might increase learners' motivation factor and drive them to solve the problems. Furthermore, providing some guidance during the solution process might improve their productivity and decision-making (Stillman & Galbraith 1998). However, Blum and Niss (1991) cautioned that providing guidance in the form of ready-made software in applied problem solving may put an unintentional emphasis on routine and recipe-like procedures that neglect essential phases, such as critically analyzing and comparing models. Thus, closely examining how this concern is resolved in newly developed mathematics software was an additional focus of this meta-analysis.

Prior literature has provided many insightful conclusions about the effectiveness of exploratory computer programs on mathematics students' achievement. However, it has also led to many questions on how the content delivery methods or problem-solving settings presented by computer programs will yield the highest learning effect sizes.

Methods

A literature review can take several venues, for example: narrative, quantitative, or meta-analytic. This study took the form of the latter, using the systematic approach proposed by Glass (1976), called *meta-analysis*, which can further be described as an analysis of the analyses. A statistical meta-analysis integrates empirical studies, investigating the same outcome described as a mean

effect size statistic. Meta-analytic techniques were selected for this study because they provide tools to assess effect size considering a pool of studies as a set of outputs collected within prescribed criteria. There are two main advantages of such investigations: (a) a large number of studies that vary substantially can be integrated, and (b) the integration is not influenced by the interpretation or use of the findings by the reviewers (Gijbels et al. 2005). Meta-analysis allows also for conducting a subgroup moderator analysis that provides tools of identifying factors that affect the magnitude of mean effect. A subsequent moderator analysis is anticipated to be employed in this study to answer additional research questions.

Key term descriptions

Treatment/instrumentation

The treatment for the study was defined as an exploratory environment that was digitally delivered and displayed on a computer screen or iPod that students used to formulate and mathematize patterns or solve problems. An exploratory learning can be defined as a medium that engages the learner with the environment through the definite actions of gathering and investigating information (Flum & Kaplan 2006) and formulating a general pattern or finding a unique solution. The treatment can include specific software, such as Frizbi Mathematics 4, SimCalcMathWorlds, NeoGeo, or Dynamic Geometry Environment (DGE). For the purpose of identifying which type of treatment produces higher effect sizes, treatments were further classified as being focused on either explorations or problem solving.

Explorations The purpose of explorations is to have students experiment with models and search for underlying structures. An example of such would be having students investigate the properties of polygons through the underlying principles of congruency and similarity.

Problem solving The main component of problem solving is asking the learner to find a specific numerical solution (Gerofsky 2004). Problem solving in this meta-analysis encompassed process associated with solving word problems, story problems, or statement problems that involve developing mathematical concepts and solving mathematical equations to find a specific numerical value.

Outcome variable of the research

The outcome variable whose overall effect size was sought in this meta-analysis was student achievement, defined as scores on solving various mathematical tasks or problems embedded in various mathematical structures, such as equations, ratios, proportions, and formulas, and measured by students' performance on standardized or researcher- or teacher-developed tests expressed numerically as a ratio or percent. Student achievement scores were

further expressed as effect size computed using mean posttest scores of experimental and control groups and coupled standard deviation using Hedges (1992) formula:

$$g = \frac{\bar{x}_1 - \bar{x}_2}{s^*}, \text{ where}$$

\bar{x}_1 represents the posttest mean score of the treatment group

\bar{x}_2 represents the posttest mean score of the control group

s^* represents pooled standard deviation.

Research questions

The research questions reflected the study purpose and they were dichotomized into two groups: main and supplementary. The main research question was the following:

- What is the magnitude and direction of the effect size of using computerized exploratory environments to support the process of problem solving and explorations when compared to conventional learning methods?

As the main question led to computing the overall effect size of using ECEs, additional questions were formulated to enrich the study objective:

- Are the effect sizes of student achievement dependent on grade level?
- Are the effect sizes of student achievement different when problem solving is contrasted with exploration?
- Are the effect sizes of student achievement dependent on mathematics content domain?
- How does the type of instructional support (teacher guided or computer based) affect student achievement when computers are used?

While the answer to the main question was assessed via interpretation of the magnitude and direction of the computed mean effect size statistic, the answers to the additional research questions were based on applied subgroup moderator analysis and interpretation of the results.

Data collection criteria and procedures

Several criteria for literature inclusion in this study were established before the search was initiated. Despite the fact that computer programs as a medium supporting learning were introduced into education several decades ago (Joyce et al. 2009), a rapid increase in this field occurred around the year 2000, which was selected as the initial timeframe for the search. Thus, this synthesis intended to analyze and summarize the research published between January 1, 2000, and December 31, 2013,

on using computerized programs to support student explorations in elementary and middle school mathematics classes in either public or private schools. The minimum sample size established in this meta-analysis was ten participants. The study included only experimental research that provided pretest-posttest mean scores, standard deviation (SD), *F*-ratios, *t*-statistics, or other quantifications necessary to compute the mean effect size. As treatment groups used computerized exploratory environments, the only control groups considered were those provided with traditional instruction, meaning use of traditional teacher-centered methods, where the students are given problems to work and when they seek help when needed as described by Pilli and Aksu (2013). This reduced the confounding of effects due to hybrid treatments.

Publication bias that is a threat to any research attempting to use the published literature (Hedges 1992) was addressed by creating funnel plot for the accumulated studies and by applying Rosenthal fail-safe *N* test and computing the fail-safe number (Rosenthal 1979). The test addressing so-called *file drawer problem* estimates the number of unpublished studies required to refute significant meta-analytic means.

Studies investigating the effects of applying exploratory computerized environments and satisfying the above conditions were identified through a search of databases of ERIC, Educational Full Text (Wilson), Professional Development Collection, ProQuest Educational Journals, as well as Science Direct, and Google Scholar. The search encompassed studies conducted globally but published in English language. Due to anticipated high range of variation of sampling methods and study-level variance that produce additional source of random influence (Cooper 2010), a random-effect model is anticipated to be used to calculate the mean effect size.

Key terms were selected by the authors from the literature pertaining this study's theoretical background and prior research. In the process of extracting the relevant literature, the following queries were used: [(‘explorations’ OR ‘problem solving’ OR ‘control of variables’) AND (‘students’ achievement’ OR ‘elementary’ OR ‘middle school’ OR ‘computers’)], [(‘simulations’ AND ‘mathematics’)], and [‘exploratory environment’ AND ‘mathematics’]. The strings were arranged in a way that allowed maximizing the search engine capabilities. Thus, for example, explorations were disjoined from problem solving but both were combined with student achievement and various grade levels. Respectively, simulations and exploratory environments were joined with mathematics. This search returned 238 articles, out of which 14 satisfied the criteria discussed above.

In order to expand the pool, a further search, including PsycINFO and PsychARTICLES, was undertaken with broader conceptual definitions including synonyms. For

instance, [(‘dynamic investigations’ OR ‘techniques of problem solving’, OR ‘computerized animations’) AND (‘learning’ OR ‘student achievement’)]. These modifications, which allowed for the adjustment of the contexts and strengthening of the relevance of the literature (Cooper 2010), returned 107 studies. The additional search extracted a number of studies that, although very informative (e.g., Chen & Liu 2007; Eysink et al. 2009; Harter & Ku 2008), could not be included in this meta-analysis because computers were used in both the control and experimental groups. After further scrutiny, the pool was enhanced by 11 additional studies. Combing all search, the pool contained 25 primary studies and 25 corresponding effect sizes.

The adherence of the pool to established research criteria was supported by a double scrutiny process at the initial and the concluding stages of the selection process. Any discrepancies were resolved.

Coding features

The coding process was conducted in a two-phase mode reflecting the two-stage analysis. During the first stage, general characteristics of the studies, such as research authors, sample sizes, study dates, research design type, and pretest-posttest scores, were extracted to describe the study features. During the second phase, additional scrutiny took place to more accurately reflect on the stated research questions and seek moderators that might influence the strength of the effect sizes. The majority of the coding features, including study authors, study publication date, locale, and research design type, were extracted to support the study validity. The formulation of other coding, including grade level, instrumentation, and learning type, was enacted to apply moderator analysis that would lead to answering the supplemental research questions.

Descriptive parameters

Descriptive parameters encompassed the following: the grade level of the group under investigation, the locale where the study was conducted, the sample size representing the number of participants in experimental and control groups, the date of the study publication, and the time span of the research expressed in a common week metric.

Inferential parameters

Posttest mean scores of experimental and control groups and their corresponding standard deviations were extracted to compute study effect sizes. If these were not provided, *F*-ratios or *t*-statistics were recorded. Although most of the studies reported more than one effect size, for example, Kong (2007) and Guven (2012), who also reported on students’ change of attitude toward

computers, this study focused only on student achievement, thus reporting one effect size per study.

The research authors

A complete list of authors whose studies were selected was compiled in Table 1. As the analysis of the study progressed, each research study was labeled only by the first author and the year of research conduct.

Publication bias

All studies included in this meta-analysis were peer reviewed and published as journal articles; thus, no additional category in the summaries was created to distinguish the publication mode of the studies. We also examined the authorship and author group membership as a consideration of a possible publication bias in certain studies being over-represented, yet no publication bias was found. Publication bias was quantified and justified through using Rosenthal fail-safe N test and by creating and examining a funnel plot.

Group assignment

This categorization was supported by the way the research participants were assigned to treatment and control groups, as defined by Shadish et al. (2002). During the coding process, two main categories emerged: (a) randomized, where the participants were randomly selected and assigned to the treatment or control group, and (b) quasi-experimental, where the participants were assigned by the researchers.

Type of research design

Only experimental studies that provided pretest-posttest means or other statistic parameters representing the means were utilized in this study.

Type of instructional support

Two subcategories were identified to classify and evaluate the effects of the type of instructional support: (a) teacher-guided support, where the teacher served as a source of providing support during student explorations or problem solving, or (b) computer-based support,

Table 1 General characteristics of the studies' features

Authors	Date	Locale	RD	SS	Grade level	RTL (in wks)	Treatment approach	IS
Pilli & Aksu	2013	Cyprus	R	55	4th	12	EX	TG
Kong	2007	Hong Kong	QE	72	4th	5	EX	TG
Hwang & Hu	2013	Taiwan	R	58	5th	8	PS	CB
Lai & White	2012	USA	QE	12	6th & 7th	1	EX	CB
Chang, Sung, & Lin	2006	Taiwan	QE	132	5th	6	PS	TG
Erbas & Yenmez	2011	Turkey	QE	134	6th	2	EX	TG
Roschelle et al.	2010	USA	R	1621	7th	40	EX	CB
Roschelle et al.	2010	USA	R	825	8th	80	EX	CB
Kapa	2007	Israel	R	107	8th	8	PS	CB
Papadopoulous & Dagdilelis	2008	Greece	QE	98	5th & 6th	4	PS	TG
Eid	2005	Kuwait	QE	62	5th	1	PS	CB
Huang, Liu, & Chang	2012	Taiwan	QE	28	2nd & 3rd	1	PS	CB
Lan, Sung, Tan, Lin, & Chang	2010	Taiwan	R	28	4th	4	PS	CB
Van Loon-Hillen, van Gog, & Brand-Gruwel	2012	Netherlands	QE	45	4th	3	PS	CB
Guvenc	2012	Turkey	QE	68	8th	40	EX	CB
Chen & Liu	2007	Taiwan	QE	165	4th	4	PS	TG
Ku & Sullivan	2002	Taiwan	QE	136	4th	1	PS	CB
Suh & Moyer-Packenham	2007	USA	QE	36	3rd	1	PS	CB
Panaoura	2012	Cyprus	QE	255	5th	8	PS	CB
Kanive et al.	2013	USA	R	90	4th & 5th	1	EX	CB
Shin et al.	2013	USA	R	41	2nd	5	EX	CB
Hwang et al.	2010	Taiwan	QE	56	6th	1	EX	CB
Kesan at al.	2013	Turkey	R	42	7th	2	EX	TG
Cakir & Simsek	2010	Turkey	R	90	7th	2	PS	TG

R, randomized; QE, quasi-experimental; RD, research design; EX, explorations; SS, sample size; RTL, research time length; wks, weeks; CB, computer based; TG, teacher guided; PS, problem solving; IS, Instructional Support.

where the primary source of support was provided by the software and was available on the computer screen. In both of these settings, the medium of learning was digitally delivered by the computer.

Length of treatment

Three main categories were established for this moderator: short - 2 weeks or less; intermediate - between 2 and 5 weeks inclusively; and long - more than 5 weeks.

We were initially interested in examining the magnitude of the exploratory learning environments on improving students' problem-solving skills, specifically focusing on analyzing the effects of scientific inquiry; however, we encountered limited research findings for extracting such features. Thus, the effect of scientific empirical methods on building theoretical mathematical models could not be investigated to the scale it was intended. The objective of the research was then augmented to focus on comparing traditional methods of teaching problem solving and explorations to ones using digital technology as a medium for such.

Results and discussion

Homogeneity verification and summary of data characteristics

The data analysis in this study was initially performed using SPSS 21 with verification of homogeneity of the study pool as suggested by Hedges (1992). A standardized mean difference effect size was calculated using posttest means on experimental and control groups. The

individual effect sizes were then weighted, indicated by ES in this study, and an overall weighted mean effect size of the study pool was calculated. In studies including multiple independent subgroups (e.g., separate data is presented for girls and boys in the experimental and control condition), first summary statistics for both conditions was recreated, and then this data was used to calculate the effect size. If different learning methods were assessed within the same study, the effects were not combined but instead the type of learning method that best matched the study's goals and research question was selected.

The homogeneity statistics ($Q_T = 117.78$, with $d_f = 23$, $p < 0.01$) showed that the set of effect sizes varied statistically significantly. This finding further supports the adoption of a random-effect model for the data analysis. If the Q_T had not been statistically significant, a fixed-effect model would have been adopted for the analysis. In order to detect potential bias due to underrepresentation of studies with small subject samples, a funnel plot was generated (see Figure 1). The funnel plot visualizes the position of the individual effect sizes, the mean effect size, as well as the confidence intervals for each study around the computed mean of the pool with 95th percentile confidence interval around the mean. Due to a high range of studies populations, the magnitudes of the populations were converted to a logarithmic scale.

The visual inspection of the funnel plot (see Figure 1) shows that the results from smaller sample size studies are more widely spread around the mean than the studies with larger sample size studies which according to

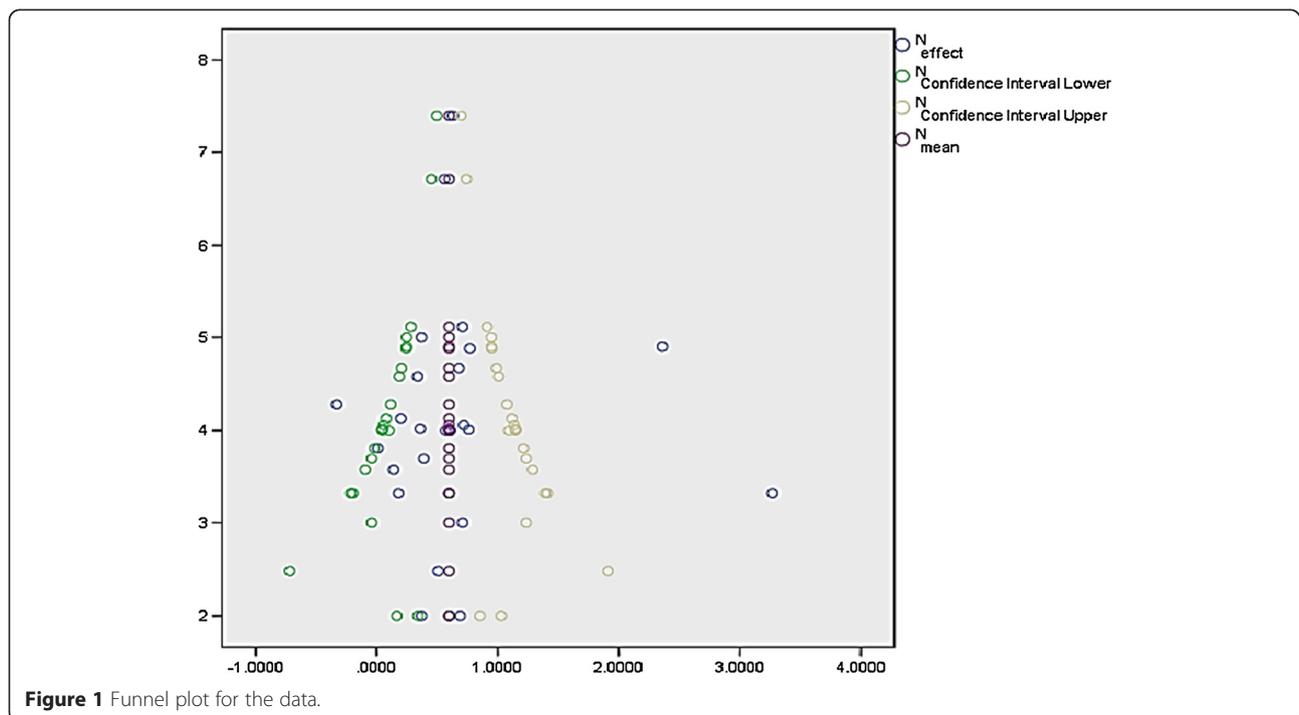


Figure 1 Funnel plot for the data.

Rothstein et al. (2006) minimizes the existence of publication bias in this meta-analysis. The funnel plot also showed some of the means located outside of the area of the funnel graph (see Figure 1), indicating a lack of homogeneity of distributions within the pool, which was also depicted by the significant p value ($p < 0.01$). As the main purpose of a meta-analysis is to compute overall effect size (Willson 1983), this deficiency did not undermine the validity of the calculated mean effect; rather, it explicated the characteristics of the studies, revealing that some of them, or their linear combinations, came from different distributions. For instance, Figure 1 illustrates three labeled studies - one by Kong (2007), labeled as 1; another by Erbas and Yenmez (2011), labeled as 2; and a third by Huang et al. (2012), labeled as 3 - whose means fell outside of the funnel graph. While Kong (2007) investigated the effects of digitally presented explorations on fourth grader understanding of fraction operations, Erbas and Yenmez (2011) investigated the effects of digitally presented explorations on sixth grader geometry concept understanding, and Huang et al. (2012) investigated the effects of digital explorations on second grader problem solving skills. These studies do not have specific common features though. Due to their complex study designs and valuable research findings, all of them were included in this meta-analysis, and thus all contributed through their weighted effect sizes to the overall effect size.

Table 1 summarizes the extracted general characteristics of the studies. The studies were further aggregated into classes to reflect the objectives of the research questions.

The calculated confidence intervals (CIs) for each study effect size (ES) consist of a range of values that act as good estimates of the ES of an unknown population. The selected 95% level indicates the probability that the CI captures this true population ES given a distribution of sample. It is to note that an ES of 17 of the studies (68%) placed within 95% CI, indicating their high precision. The duration of experimental treatment usually lasted for one unit lesson (45 min) with an average application frequency of twice a week. The majority of the studies (14, or 56%) were conducted quasi-experimentally, while the remaining ten (44%) were randomized. The study duration was expressed in a common (weeks) metric scale, although some of the studies reported the duration in months or by semesters. The highest sample size of 1,621 students was reported for a study conducted by Roschelle et al. (2010), and the lowest sample size of 12 participants was reported by Lai and White (2012). While analyzing the studies from a grade-level point of view, students whose primary level was grade 4 (28%) dominated the pool. Since this study focused on gathering research on exploratory environments provided by computer programs or the Internet, the examined studies were aggregated by their focus on supporting either problem solving or explorations in

mathematics. For example, the study conducted by Lai and White (2012) was classified as an exploration because students explored the space and constructed shapes without constraints, and then provided their definitions for the various quadrilaterals. Thus, the explorations led them to formulate general patterns and descriptions. Similarly, a study by Kong (2007) was also classified as an exploration because it used a general partitioning model to have students explore the subject of common fractions and fraction operations. In contrast, problem-solving studies, such as the one conducted by Kapa (2007), followed the traditional four stages of problem solving: (1) understanding the problem, (2) making a plan, (3) executing the plan and thus finding a unique solution, and (4) reviewing the solution. The study-highlighted characteristics were further aggregated into subgroups (see Table 2) and corresponding effect sizes were computed.

Descriptive analysis

The analysis of the data was organized deductively. It began with a synthesis of the general features of the studies, furnished by a descriptive analysis, and then moved to an examination of the differences of the effect sizes mediated by the type of instrumentation, cognitive domain, study duration, grade level, and content domain.

The research pool encompassed 4,256 elementary and middle school students. The average sample size was 112 participants. Applied descriptive analysis provided information about the frequencies of the studies per year (see Figure 2) and the locale distribution where the studies were conducted (see Figure 3).

The majority of the studies (17, or 71%) were conducted within the past 5 years, which indicates a growing interest in using ECEs to support the learning of mathematics. In terms of research locale, Taiwan dominated the pool with seven studies (29%), followed by the United States with five studies (21%). The distributions show that applying and investigating the effects of ECEs in mathematics classrooms has accumulated a global interest.

Inferential analysis

Quantitative inferential analysis was performed on the primary studies to find individual weighted effect size and the mean weighted effect size of the study pool. The mean effect size for the 24 primary studies (24 effect sizes) was reported to have a magnitude of 0.60 ($SE = 0.03$) and a positive direction, which according to Lipsey and Wilson (2001) can be classified as of a medium size. A 95% confidence interval around the overall mean - $C_{lower} = 0.53$ and $C_{upper} = 0.66$ - supported its statistical significance and its relative precision as defined by Hunter and Schmidt (1990). When applied to school practice, it indicated that the score of an average student in the experimental

Table 2 Effect sizes of using ECEs in grades 1 to 8 mathematics

Study (first author)	ES	SE	95% CI		Reliability of measure	Program used, research findings, research specifications
			Lower	Upper		
Pilli (2013)	0.76	0.24	0.05	1.09	Researcher developed, Cronbach's $\alpha = 0.9$	Used Frizbi Mathematics 4. Explored arithmetic operations
Kong (2007)	-0.33	0.27	0.12	1.15	Teacher developed	Used Graphical Partitioning Model (GPM). Fraction operations were explored. GPM has a potential for promoting collaborative learning
Hwang (2013)	0.72	0.59	0.07	1.93	Researcher developed	Used virtual manipulative and 3D objects. Investigated the effect of peer learning
Lai (2012)	0.51	0.18	0.71	0.97	California Mathematics Standard Test	Used NeoGeo. Interactive environment helped make the applications meaningful. Investigated a peer effect
Chang (2006)	0.77	0.18	0.26	0.96	Researcher developed	Used schemata-developed problem solving. Provided teacher guidance to support phases of problem solving
Erbas (2011)	2.36	0.05	0.26	0.71	Researcher developed	Used DGE. Dynamic environment contextualized scenarios well
Roschelle (2010)	0.63	0.07	0.51	0.75	Researcher developed	Used SimCalcMathWorlds. Explored the concepts of change
Kapa (2007)	0.68	0.20	0.20	1.00	Ministry of Education guided	Used three-step problem-solving and open-ended scenarios
Papadopoulos (2008)	0.34	0.21	0.22	1.02	Researcher developed	Used computers to help explore hypotheses and verify the solutions
Eid (2005)	0.20	0.16	0.19	0.92	Standardized	Contrasted students' performance using computerized scenarios and traditional representations
Huang (2012)	3.27	0.26	0.29	1.13	Researcher designed	Used onscreen presented solutions to walk students through the course of thinking
Lan (2010)	0.18	0.40	0.09	0.42	CEA assessment	Used Group Scribbles (GS) platform that enhances collaboration. Developed stages of problem solving
Van Loon-Hillen (2012)	-0.01	0.39	0.20	1.40	Researcher developed	Worked examples to help with following procedures
Guyen (2012)	0.61	0.32	0.19	1.22	Researcher developed	Used dynamic geometry software (DGS). Developed four stages of difficulty: recognition, analysis, deductive, and rigorous
Chen (2010)	0.71	0.34	0.01	1.30	Teacher developed	Incorporated personal contexts that helped students relate mathematics concepts with their experience
Ku (2002)	0.23	0.18	0.26	0.96	Teacher developed	Used personalized context to help students with mathematics concept understanding
Suh (2007)	0.14	0.34	0.09	1.30	Researcher developed	Incorporated principle of balance scale to model linear equations
Panaoura (2012)	0.37	0.13	0.34	0.85	Researcher developed	Developed program that divided problem into stages that focused the students' attention of cognitive processes
Kanive (2013)	0.36	0.27	0.05	1.14	Researcher developed	Computer program provides intermediate feedback
Shin (2013)	0.39	0.31	-0.04	1.23	Researcher developed	Computer program provided a game-like environment
Hwang (2010)	0.57	0.27	0.03	1.14	Teacher developed	Used computer programs to provide students with tasks descriptions
Kesan (2013)	0.71	0.32	-0.04	1.25	Researcher developed	Used Sketchpad geometry software
Cakir (2010)	0.69	0.22	0.17	1.03	Used PISA test bank	Used computer program to personalized and scaffold tasks

ES, effect size; SE, standard error; PISA, Program for International Student Assessment; DGE, Dynamic Geometry Environment.

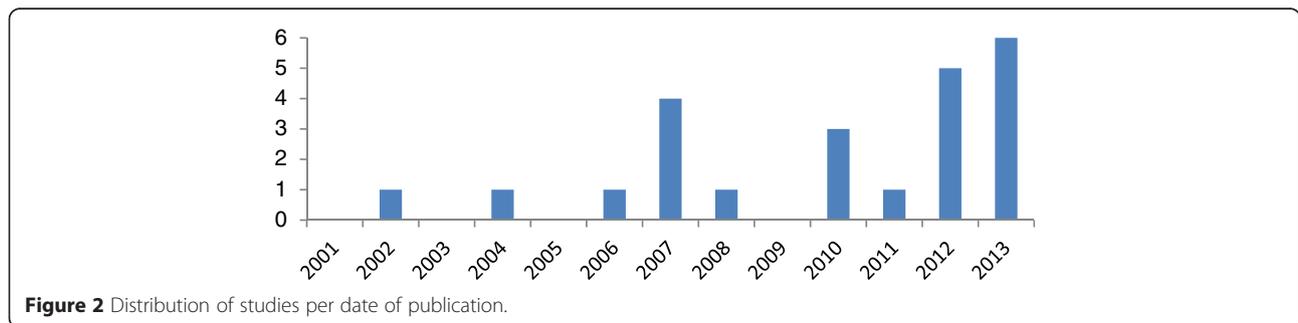


Figure 2 Distribution of studies per date of publication.

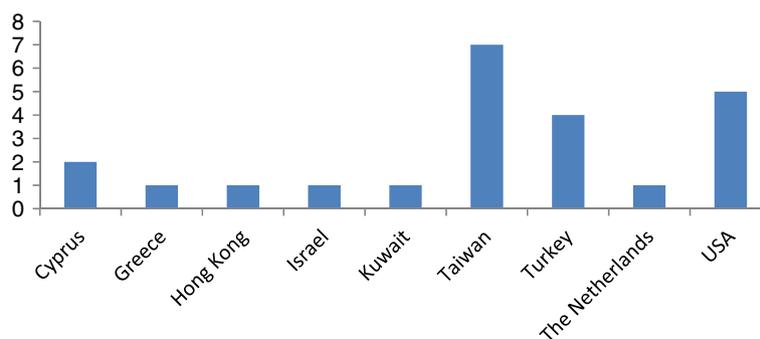


Figure 3 Distribution of studies per locale.

groups, who learned using ECEs, was 0.60 of standard deviation above the score of an average student in the control groups, who was taught using traditional methods of instruction. In order to quantify publication bias, Rosenthal's fail-safe procedure was used. The test has showed that an additional 480 unpublished null-effect studies would be required to bring the p level beyond the 0.05 threshold of significance. Further calculations show that only 130 unpublished research papers are needed for this study to nullify the mean of 0.6 below 0.05 level. Combining the results of both tests, the inspection of the funnel plot (see Figure 1) and procedure of fail-safe, we claim that the publication bias has minimal effects of the mean effect size calculations in this study. Further examination of the computed effect size and incorporation of the U_3 effect size matrix (Cooper 2010) led to the conclusion that the average pupil who learned mathematical structures using exploratory environments scored higher on unit tests than 70% of students who learned the same concepts using traditional textbook materials. It can thus be deduced that using exploratory environments as a medium of support in the teaching of mathematics has a significant impact on students' mathematics concept understanding when compared to conventional methods of teaching. Table 2 provides a summary of the individual effect sizes of the meta-analyzed studies along with their confidence intervals. The table also contains qualitative research findings, the computer programs used as the instruments, and the reliability of measures used to compute the individual mean scores, expressed by indicating whether the test was researcher developed or standardized. Where it was available, Cronbach's alpha (α) was also listed, along with additional information provided by the primary researchers that distinguish the given study within the pool.

The majority of the studies (16, or 84%) used researcher- or teacher-developed evaluation instruments, and only one (Pilli & Aksu 2013) reported a Cronbach's α coefficient of reliability measure. In addition, the majority

of the studies (17, or 89%) reported positive effect sizes when an exploratory environment was used as a medium of learning. Only two studies - one conducted by Van Loon-Hillen et al. (2012) and one conducted by Kong (2007) - reported negative effect sizes favoring traditional instruction, illustrating that exploratory environments cannot replace good teaching and that some concepts, like operations on fractions (Kong 2007), require the instructor to deliver the concept and its stages and to suggest ways of overcoming obstacles that students may face. Exploratory environments seemed to produce high effect sizes in cases where students applied already-learned mathematics concepts in new situations (e.g., Chang et al., 2006; Guven, 2012; Roschelle et al., 2010) but not when students simultaneously explored new concepts and applied them. The highest effect size on explorations was reported by Erbas and Yenmez (2011; ES = 2.36), who examined the effect of open-ended geometry investigations, and the highest effect size on problem solving was noted by Huang and colleagues (2012; ES = 3.27), who investigated the effect of embedded support during the process of problem solving. Although an influx of onscreen instructional support might work well in many classroom settings, we believe that the elements of mathematical explorations induced in the study by Erbas and Yenmez (2011) more accurately supported the objectives of this study.

Possible moderators and analysis of their effects

Just as a mean effect size provides certain evidence for potential duplications of the study findings, subgroup analysis allows for uncovering moderators that optimize the effect. Since student mathematical achievement was the main construct under investigation, during the process of moderator formulation, attention was directed toward formulating and extracting the study features that could affect achievement when ECEs were utilized. We anticipated that through identifying such features, an optimum learning environment would emerge. By analyzing the settings of the primary research, a set of five moderators was

identified: grade level, instrumentation, treatment duration, content domain, and type of learning setting. This categorization resulted in 12 subgroups whose effects were individually computed. The moderator analyses were performed for random-effects model because this meta-analysis sought also to reveal the relative effectiveness of ECE due to formulated moderators, and hence it did not assume that the effect size will be the same for all studies. The computations of effect sizes followed Cooper (2010), who suggested giving more weight to effect sizes with larger sample populations ($w = \text{inverse of the variance in the effect calculations}$). Along with calculating subgroup effects, corresponding confidence intervals and standard errors were also computed. Table 3 displays the effect sizes according to the formulated moderators and the subgroups. In order to provide a common metric for the subgroup effect magnitude comparisons, the effect sizes were weighted by the sample sizes. Although we realized that the effects of using ECEs might be strongly influenced by the degree of interactivity of the educational programs used and the applied scaffolding necessary to have students assimilate tasks presented in contexts, such extractions were not feasible due to the limited information provided in this regard by the primary researchers.

All of the magnitudes of the calculated effect sizes place within their confidence intervals, which proves the significance of the effect sizes and their relative precision (Hunter & Schmidt, 1990). Furthermore, considering, for example, teacher-guided support during explorations ($ES = 0.75$), one can conclude that the practitioners using such an approach can be 95% confident that the effect size of students' achievement will be 60% to 93% higher than when compared to traditional level of instruction. The categorization into subgroups and the descriptive analysis provided a more insightful picture about the effects of ECEs on the achievement of students in grades 1 to 8 mathematics classes and helped answer the research questions of this study, as discussed next.

Are the effect sizes of student achievement dependent on grade level?

A block of *grade level* was created to answer this question. Following NCTM (2000), three subgroup levels were formulated: lower elementary, which included grades 1 to 3; upper elementary, which included grades 4 and 5; and middle school, which encompassed grades 6 to 8. The computed effect size showed differences across grade levels, with middle school producing the highest effect size ($ES = 0.65$), which according to Lipsey and Wilson (2001) can be classified as moderate followed by lower elementary school ($ES = 0.61$) and upper elementary ($ES = 0.41$). It is inferred that this result can be attributed to the fact that students at the middle school level often

Table 3 Summary of subgroups' weighted effect sizes

Moderator and subgroups	N	ES	SE	95% CI	
				Lower	Upper
Grade level					
Lower elementary: 1 through 3	3	0.61	0.03	0.54	0.67
Upper elementary: 4 to 5	12	0.41	0.07	0.27	0.54
Middle school: 6 to 8	9	0.65	0.04	0.57	0.73
Instrumentation					
Problem solving	12	0.54	0.07	0.41	0.67
Explorations	12	0.62	0.03	0.54	0.69
Treatment duration					
Short	8	0.47	0.14	0.11	0.74
Intermediate	9	0.63	0.09	0.45	0.81
Long	7	0.62	0.04	0.55	0.70
Content domain					
Geometry	9	0.67	0.22	-0.07	0.79
Arithmetic and algebra	15	0.61	0.03	0.43	0.56
Type of instructional support					
Teacher guided	8	0.75	0.08	0.59	0.92
Computer based	16	0.56	0.04	0.49	0.63

N, sample size; ES, effect size; SE, standard error.

use manipulatives to support their mathematics concept understanding (e.g., see Jitendra et al., 2007); thus, these students' transition to ECEs occurs more spontaneously, resulting in the highest score gain. The effect sizes in the other grades also showed a moderate magnitude.

Are the effect sizes of student achievement different when problem solving is compared to exploration?

The moderator category of *instrumentation* was used to conclude whether ECEs affect student achievement differently through supporting problem solving or exploration. As explorations have often led students to pattern formulations (e.g., see Panaoura, 2012; Suh & Moyer-Packenham, 2007), problem solving was usually constructed within defined stages, leading students toward finding numerical answers or unique solutions to the stated problems (e.g., see Chen & Liu, 2007; Hwang & Hu, 2013). When contrasted with problem solving, learning supported by explorations produced a higher effect size of $ES = 0.62$ (as opposed to $ES = 0.54$ for problem solving). This finding generated several conjectures. As the process of explorations resonates better with students' natural curiosity (Stokoe, 2012) and their prior experiences, working on explorations might ignite a higher student's motivation level, thus their higher achievement. Despite the fact that efforts to help students understand the solution process are multidimensional, ranging from creating schemas (Kapa, 2007) to inducing personalization (Chen & Liu,

2007; Ku & Sullivan, 2002), attempts at helping students learn the process of problem solving by embedding explorations in some of the transitioning stages are rare. With exceptions of researches conducted by Roschelle et al. (2010) and Panaoura (2012), an inquiry process is not emphasized in the accumulated pool of studies on problem solving despite a strong support by other researchers (e.g., see English, 2004). The current research on problem solving gravitates toward creating and examining the effects of cognitive support or showing worked-out solutions that students can follow (e.g., see Van Loon-Hillen et al., 2012). As illustrated by the computed effect sizes, all of these attempts seem to produce desirable positive results; however, by concentrating on simplifying or following the mechanics of the problem-solving process, the meaning of the context embedded in the problems is diminished. The question that arises here is How to convert word problems to explorations? Bonotto (2007) suggested 'Change the type of activity aimed at creating interplay between the real world and mathematics towards more realistic and less stereotyped problem situations' (p. 86) and 'change classroom culture by establishing new classroom socio-mathematical norms' (p. 86). (Greer et al. 2007) proposed to 'Valorize forms of answer other than single, exact numerical answers' (p. 92).

Are the effect sizes dependent on the mathematics content domain?

Two mathematics content domains dominated in this study pool: geometry and algebra. Geometry, traditionally supported by visualization, showed a higher effect size ($ES = 0.67$) compared to algebra ($ES = 0.61$). As geometric objects can also be easily externalized by their real embodiments, more effort should be placed on contextualization and visualization of other, more abstract mathematical structures such as functions. Teaching algebraic structures via exploratory environments is being practiced and researched, yet embodying algebraic structures by context-driven scenarios seems to be a challenge, which is reflected by locating only eight (33%) such studies.

How does the type of instructional support (teacher guided or computer based) affect student achievement when computers are used?

There were two main categories of instructional support provided to the students in the study pool: computer-based support displayed on the computer screen or teacher-centered support provided by the instructor. Computer-based instructional support dominated the study pool (16, or 67%) compared to teacher-based instruction (8, or 33%). When compared by learning effects, teacher-centered support produced a higher effect ($ES =$

0.75) than computer-based support ($ES = 0.56$). This result signifies the importance of the teacher's role in developing students' understanding of mathematics structures and helping them apply the structures to solve problems, and it corresponds with Li and Li's (2009) finding who claimed that 'teacher transfers the knowledge development and justification responsibilities to students' (p. 275). A particular instance that needs further investigation is transitioning from verifying to explaining (Hähkiöniemi & Leppäaho, 2012). Merchant et al. (2014) claimed that 'It is essential that teachers are made knowledgeable about the features and situations that make feedback effective' (p. 37).

Programmed tips are important and readily available to students, yet the expertise, encouragement, and support from a live person appear to have a higher impact on students' learning. Further research contrasting learning effect sizes by using as a moderator, for example, the frequency of seeking help available on the computer screen versus frequency of seeking help from the teacher along with quality of answers sought, would likely shed more light on the cause of the differences.

In addition to analyzing the effects of moderators that reflected the research questions, the effect of treatment length was also computed. The analysis showed that the treatment of length between 2 and 5 weeks, called *intermediate* herein, produced the highest effect size of ($ES = 0.63$). A similar conclusion was reached in a meta-analysis by Xin et al. (2005), who also proved that longer treatment results in higher student achievement. The student needs to be acquainted with the mechanics of the new learning medium; thus, it is important that the first contact and experience with an ECE be absorbed into a learner's working memory. It is hypothesized that longer and more frequent exposure to the new environment allows a higher focus on task-driven objectives related to the content analysis, which consequently results in better context understanding and higher learning effects. However, as Guven (2012) and Roschelle et al. (2010) found, there is an achievement saturation level, which perhaps suggests that in order to further increase learning effects, ECEs need to mediate with other factors, not necessarily related to content knowledge, for instance different forms of analysis, synthesis, or evaluation as suggested by Anderson and Krathwohl (2001).

When linking the subgroups with the highest learning effects, it appears that month-long geometry explorations in grades 1 to 3 mathematics classes, guided by the teacher, would produce the highest learning effects.

Conclusions

While this study found a moderate positive effect size ($ES = 0.59$) associated with ECE, this finding does not diminish the importance of good teaching. Several studies (Christmann et al. 1997; Clark, 1994; Povey & Ransom,

2000) found that using computers purely as a method of instruction does not improve students' mathematics understanding. Hence, although computers have been used in mathematics classrooms for several decades now, the question regarding to what extent they can impact the teaching and learning of mathematics seems to be open for further investigations. This meta-analysis of up-to-date literature allowed for formulations of some inferences based on implementations of technology; however, many new questions emerged, such as the following: How do exploratory environments help students with the transfer of mathematics concepts to new situations? How can we assure that the methods of quantitative scientific modeling that students apply in their science classes are coherent with the ones used in mathematics, and vice versa? Mathematics provides tools for phenomena quantifications; thus, unifications of the techniques of modeling seem to benefit the transition of knowledge between mathematics and science and consequently affect the learners' perception of mathematics as a subject with a high applicability range. Will such unification prompt students to increase their engagement in mathematics? More detailed studies in these domains are worthy of consideration, and the availability of ECEs will be very helpful in organizing such studies.

The impact of ECEs on students' problem-solving techniques

Problem-solving techniques are developed on the basis of understanding the context through identifying the principles of the system's behavior. However, it is a highly intertwined process that might include verbal and syntactic processing, special representation storage and retrieval in short- and long-term memory, algorithmic learning, and its most complex element - conceptual understanding (Goldin, 1992). Computerized programs offering a basis for investigation offer a great potential for improving conceptual understanding of problems; however, this study shows that this area is not yet fully explored, and taking full advantage of such learning environments to examine their impact on student achievement is a possible extension of this undertaking. More specifically, enriching the problem analysis through explorations to focus learners' attention more on the underpinning principles emerges as a possible objective of such studies. Higher student achievement on explorations ($ES = 0.62$), compared to problem solving ($ES = 0.54$), encourages designs of more comprehensive research about inducing an exploratory approach to problem solving also to solving standard textbook problems as opposed to current schemata-driven methods. Would giving students more ownership in exploring a given system's behavior, hypothesizing a solution, testing, and proving or disproving their hypothesis be a possible moderator affecting learning? Will these types of activities help solidify a notion and belief in the power of

mathematics as sense-making subject? There also seems to be more work needed to evaluate how learners link mathematics concepts with principles embedded in given context and how they initiate applications of the procedures that they select. 'Only an analysis of the instrument, i.e., the interaction of the artifact and the utilization schemes of its users (teachers and students), the analysis of its *instrumental genesis* will help in the implementation of computers, software and communication technology in the mathematics classroom' (Laborde and Str  ber 2010, p. 131).

Limitations and suggestions for future research

This meta-analytic research has certain limitations, primarily because this study could not be conducted in an experimental fashion where ECEs constituted instrumentation provided by computer programs and a direct contrast between two different modes of learning - digital and traditional - were exploited. Furthermore, the limited count of studies available to be meta-analyzed affected the study generalizability. Although sensitivity to smaller sample sizes was restored by the process of weighing, the impact of the mean effect would validate the replication of the findings more significantly by being computed over a larger study pool.

The other factor affecting validity of the computed effect sizes is the high span of interactivity of the software used in the primary studies and their exploratory nature, ranging from linear equation exploration supported by an interactive balance scale (Suh & Moyer-Packenham, 2007) to investigation of rate of change supported by the SimCalcMathWorlds program (Roschelle et al. 2010). A metric for incorporating this moderator to effect size calculations could have been furnished by evaluating the designs of the interventions through the lens of the *multi-media principles* defined by Clark and Mayer (2011). This task, however, was not possible to accomplish due to the lack of the software detailed descriptions.

The validity of the research would have been higher if the calculated homogeneity statistics were not statistically significant. In this meta-analysis, $Q_T = 117.78$, with $d_f = 23$, $p < 0.01$), which implied a random-effect model for the data analysis instead of a more precise fixed-effect model.

Another factor limiting the study findings involves the widely varied student assessment methods, ranging from traditional multiple-choice questions mostly locally developed to new assessment techniques such as standardized-based assessments. Although one of the studies (Pilli & Aksu, 2013) reported a Cronbach's alpha reliability coefficient, most did not, thereby decreasing the reliability of the measuring instrument. The degree of diversity further extends due to obvious differences in mathematics curricula, objectives, and expectation levels in the nine countries whose research studies were represented herein. Even

though control groups in the extracted pool of studies were taught traditionally, the term *traditional teaching* might have been interpreted differently depending on the country. For example, Lan et al. (2010) defined traditional instruction as being without the support of technological devices, whereas Papadopoulos and Dagdilelis (2008) defined traditional teaching as a paper-and-pencil environment. Both descriptions imply that technology was not being used, but the treatment applied in the control groups might have varied in terms of degree of representations and method used, or teacher qualifications, which potentially could have mediated with the control groups' posttest scores. We concluded, however, that these fluctuations did not affect the overall effect size in a manner that would question the validity of the computed overall effect size.

Though at first we intended to examine the effects of embedded scientific inquiry methods in exploratory environments on students' problem-solving skills, we encountered a limited number of research studies addressing this domain. Thus, we modified the study focus. We realized that exploratory environments used in both types of interventions - explorations and problem solving - contained, to a certain degree, some elements of scientific inquiry and affected students' problem-solving skills, not just students' problem-solving performance as measured by testing. Further studies focusing primarily on the effects of inducing scientific inquiry processes in mathematical modeling and problem solving would serve to extend this paper. Technology has encouraged researchers to consider not only how to best adapt tools to the learning of mathematics but also how to adapt the content of mathematics in light of new, tool-rich possibilities to enable learners to perform tasks that would not previously have been possible (Hoyle & Noss, 2009). The task of explorations seems to provide basis for inducing such adaptations.

This meta-analysis, to a certain extent, exposed the focus of the existing primary studies on the effects of exploratory environments on problem solving in mathematics education. We advocate for searching and formulating more constructs to quantify students' problem-solving techniques with ECEs as a medium of context.

Competing interests

The authors declare that they have no competing interests.

Authors' contribution

AS was conducted research as part of his dissertation. YL and VW co-chairman of the research dissertation.

Author details

¹Department of Mathematics and Science, Lone Star College, Tomball, TX 77375, USA. ²College of Education and Human Development, Department of Teaching Learning and Culture, Texas A&M University, College Station, TX 77843-4225, USA. ³College of Education and Human Development, Department of Educational Psychology, Texas A&M University, College Station, TX 77843-4225, USA.

Received: 4 September 2014 Accepted: 12 April 2015

Published online: 20 May 2015

References

- Anderson, L, & Krathwohl, DA. (2001). *Taxonomy for learning, teaching and assessing: a revision of Bloom's Taxonomy of Educational Objectives*. New York, NY: Longman.
- Arthur, JD, & Nance, RE. (2007). Investigating the use of software requirements and engineering techniques in simulation modeling. *Journal of Simulation*, 1(3), 159–174.
- Artzt, AF, & Armour-Thomas, E. (1992). Development of a cognitive-metacognitive framework for protocol analysis of mathematical problem solving in small groups. *Cognition and Instruction*, 9(2), 137–175.
- Bleich, L, Ledford, S, Hawley, C, Polly, D, & Orrill, C. (2006). An analysis of the use of graphical representation in participants' solutions. *The Mathematics Educator*, 16(1), 22–34.
- Blum, W, & Niss, M. (1991). Applied mathematical problem solving, modeling, applications, and links to other subjects: state trends and issues in mathematics instruction. *Educational Studies in Mathematics*, 22(1), 37–68.
- Bonotto, C. (2007). How to replace word problems with activities of realistic mathematical modelling. In W Blum, P Galbraith, HW Henn, & M Niss (Eds.), *Modelling and applications in mathematics education: the 14th ICMI Study* (pp. 45–56). New York, NY: Springer.
- Bos, B. (2009). Virtual mathematics objects with pedagogical, mathematical, and cognitive fidelity. *Computers in Human Behavior*, 25(2), 521–528.
- Bransford, JD, & Stein, BS. (1984). *The ideal problem solver. A guide for improving thinking, learning, and creativity*. New York, NY: Freeman.
- Cakir, O, & Simsek, N. (2010). A comparative analysis of the effects of computer and paper-based personalization on student achievement. *Computers & Education*, 55(4), 1524–1531.
- Chang, KE, Sung, YT, & Lin, SF. (2006). Computer-assisted learning for mathematical problem solving. *Computers & Education*, 46(2), 140–151.
- Chen, CH. (2010). Promoting college students' knowledge acquisition and ill-structured problem solving: Web-based integration and procedure prompts. *Computers & Education*, 55(1), 292–303.
- Chen, CJ, & Liu, PL. (2007). Personalized computer-assisted mathematics problem-solving program and its impact on Taiwanese students. *Journal of Computers in Mathematics and Science Teaching*, 26(2), 105–121.
- Christmann, E, Badgett, J, & Lucking, R. (1997). Microcomputer-based computer-assisted instruction within differing subject areas: a statistical deduction. *Journal of Educational Computing Research*, 16(3), 281–296.
- Clark, RE. (1994). Media will never influence learning. *Educational Technology Research and Development*, 42(2), 21–29.
- Clark, RC, & Mayer, RE. (2011). *e-Learning and the science of instruction: proven guidelines for consumers and designers of multimedia learning*. San Francisco, CA: Pfeiffer.
- Cooper, H. (2010). *Research synthesis and meta-analysis* (4th ed.). Thousand Oaks, CA: Sage.
- Eid, GK. (2005). An investigation into the effects and factors influencing computer-based online mathematics problem-solving in primary schools. *Journal of Educational Technology Systems*, 33(3), 223–240.
- English, L. (2004). Mathematical modeling in the primary school. In *Mathematics education for the third millennium: towards 2010 (Proceedings of the 27th Annual Conference of the Mathematics Education Research Group of Australasia)* (Vol. 1, pp. 207–214). Sydney, Australia: MERGA.
- English, LD, & Watters, JJ. (2005). Mathematical modeling in the early school years. *Mathematics Education Research Journal*, 16(3), 58–79.
- Erbas, AK, & Yenmez, AA. (2011). The effect of inquiry-based explorations in a dynamic geometry environment on sixth grade students' achievements in polygons. *Computers & Education*, 57(4), 2462–2475.
- Evan, R, & Lappan, G. (1994). Constructing meaningful understanding of mathematics content. In D Aichele & A Coxford (Eds.), *Professional development for teachers of mathematics* (pp. 128–143). Reston, VA: NCTM.
- Eysink, TH, de Jong, T, Berthold, K, Kolloffel, B, Opfermann, M, & Wouters, P. (2009). Learner performance in multimedia learning arrangements: an analysis across instructional approaches. *American Educational Research Journal*, 46(4), 1107–1149.
- Flum, H, & Kaplan, A. (2006). Exploratory orientation as an educational goal. *Educational Psychologist*, 4(2), 99–110.
- Forster, PA. (2006). Assessing technology-based approaches for teaching and learning mathematics. *International Journal of Mathematical Education in Science and Technology*, 37(2), 145–164.

- Francisco, JM, & Maher, CA. (2005). Conditions for promoting reasoning in problem solving: insights from a longitudinal study. *The Journal of Mathematical Behavior*, 24(3), 361–372.
- Gerofsky, S. (2004). *A man left Albuquerque heading east: word problem as genre in mathematics education* (Vol. 5). New York, NY: Peter Lang.
- Gijbels, D, Dochy, F, Van den Bossche, P, & Segers, M. (2005). Effects of problem-based learning: a meta-analysis from the angle of assessment. *Review of Educational Research*, 75(1), 27–61.
- Glass, GV. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher*, 5, 3–8.
- Goldin, GA. (1992). Meta-analysis of problem-solving studies: a critical response. *Journal for Research in Mathematics Education*, 23(3), 274–283.
- Gravemeijer, K, & Doorman, M. (1999). Context problems in realistic mathematics education: a calculus course as an example. *Educational Studies in Mathematics*, 39(1), 111–129.
- Greer, B, Verschaffel, L, & Mukhopadhyay, S. (2007). Modelling for life: Mathematics and children's experience. In W Blum, P Galbraith, HW Henn, & M Niss (Eds.), *Modelling and applications in mathematics education: the 14th ICMI Study* (pp. 89–98). New York, NY: Springer.
- Grouws, DA, & Cebulla, KJ. (2000). *Improving student achievement in mathematics*. Geneva, Switzerland: International Academy of Education.
- Güven, B. (2012). Using dynamic geometry software to improve eighth grade students' understanding of transformation geometry. *Australasian Journal of Educational Technology*, 28(2), 364–382.
- Häkkiöniemi, M, & Leppäaho, H. (2012). Prospective mathematics teachers' ways of guiding high school students in GeoGebra-supported inquiry tasks. *International Journal for Technology in Mathematics Education*, 19(2), 45–58.
- Hart, JM. (1996). The effect of personalized word problems. *Teaching Children Mathematics*, 2(8), 504–505.
- Harter, CA, & Ku, HY. (2008). The effects of spatial contiguity within computer-based instruction of group personalized two-step mathematics word problems. *Computers in Human Behavior*, 24(4), 1668–1685.
- Hedges, LV. (1992). Meta-analysis. *Journal of Educational and Behavioral Statistics*, 17(4), 279–296.
- Hiebert, J (Ed.). (2013). *Conceptual and procedural knowledge: the case of mathematics*. New York, NY: Routledge.
- Hoyle, C, & Noss, R. (2009). The technological mediation of mathematics and its learning. *Human Development*, 52(2), 129–147.
- Huang, TH, Liu, YC, & Chang, HC. (2012). Learning achievement in solving word-based mathematical questions through a computer-assisted learning system. *Educational Technology & Society*, 15(1), 248–259.
- Hunter, JE, & Schmidt, FL. (1990). *Methods of meta-analysis: correcting error and bias in research findings*. Newbury Park, CA: Sage.
- Hußmann, S. (2007). Building concepts and conceptions in technology-based open learning environments. In W Blum, P Galbraith, HW Henn, & M Niss (Eds.), *Modelling and applications in mathematics education: the 14th ICMI Study* (pp. 341–348). New York, NY: Springer.
- Hwang, WY, & Hu, SS. (2013). Analysis of peer learning behaviors using multiple representations in virtual reality and their impacts on geometry problem solving. *Computers & Education*, 62, 308–319.
- Hwang, GJ, Wu, PH, Zhuang, YY, & Huang, YM. (2013). Effects of the inquiry-based mobile learning model on the cognitive load and learning achievement of students. *Interactive Learning Environments*, 21(4), 338–354.
- Jitendra, AK, Griffin, CC, 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.
- Joyce, B, Weil, M, & Calhoun, E. (2009). *Models of teaching* (8th ed.). Boston, MA: Pearson.
- Kanive, R, Nelson, PM, Burns, MK, & Ysseldyke, J. (2014). Comparison of the effects of computer-based practice and conceptual understanding interventions on mathematics fact retention and generalization. *The Journal of Educational Research*, 107(2), 83–89.
- Kapa, E. (2007). Transfer from structured to open-ended problem solving in a computerized metacognitive environment. *Learning and Instruction*, 17(6), 688–707.
- Kesan, C, & Caliskan, S. (2013). The effect of learning geometry topics of 7th grade in primary education with dynamic Geometer's Sketchpad geometry software to success and retention. *Turkish Online Journal of Educational Technology-TOJET*, 12(1), 131–138.
- Kieran, C, & Hillel, J. (1990). "It's tough when you have to make the triangles angle": insights from a computer-based geometry environment. *The Journal of Mathematical Behavior*, 9(2), 99–127.
- Kim, MC, & Hannafin, MJ. (2011). Scaffolding problem solving in technology-enhanced learning environments (TELEs): bridging research and theory with practice. *Computers & Education*, 56(2), 403–417.
- Kong, SC. (2007). The development of a cognitive tool for teaching and learning fractions in the mathematics classroom: a design-based study. *Computers & Education*, 51(2), 886–899.
- Ku, HY, & Sullivan, HJ. (2002). Student performance and attitudes using personalized mathematics instruction. *Educational Technology Research and Development*, 50(1), 21–34.
- Kuhn, D. (2007). Is direct instruction the answer to the right question? *Educational Psychologist*, 42(2), 109–114.
- Laborde, C, & Sträßer, R. (2010). Place and use of new technology in the teaching of mathematics: ICMI activities in the past 25 years. *ZDM*, 42(1), 121–133.
- Lai, K, & White, T. (2012). Exploring quadrilaterals in a small group computing environment. *Computers & Education*, 59(3), 963–973.
- Lan, YJ, Sung, YT, Tan, NC, Lin, CP, & Chang, KE. (2010). Mobile-device-supported problem-based computational estimation instruction for elementary school students. *Educational Technology & Society*, 13(3), 55–69.
- Lester, FK, Jr, Masingila, JO, Mau, ST, Lambdin, DV, dos Santos, VM, & Raymond, AM. (1994). Learning how to teach via problem solving. In D Aichele & A Coxford (Eds.), *Professional development for teachers of mathematics* (pp. 152–166). Reston, VA: NCTM.
- Lewis, R, Stoney, S, & Wild, M. (1998). Motivation and interface design: maximizing learning opportunities. *Journal of Computer Assisted Learning*, 14(1), 40–50.
- Li, Y, & Li, J. (2009). Mathematics classroom instruction excellence through the platform of teaching contests. *ZDM*, 41(3), 263–277.
- Li, Q, & Ma, X. (2010). A meta-analysis of the effects of computer technology on school students' mathematics learning. *Educational Psychology Review*, 22(3), 215–243.
- Lipsey, MW, & Wilson, DB. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.
- Merchant, Z, Goetz, ET, Cifuentes, L, Keeney-Kennicutt, W, & Davis, TJ. (2014). Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: a meta-analysis. *Computers & Education*, 70, 29–40.
- Merrill, MD, & Gilbert, CG. (2008). Effective peer interaction in problem-centered instructional strategy. *Distance Education*, 29(2), 199–207.
- National Council of Teachers of Mathematics (NCTM). (2000). Principles and standards for school mathematics. Reston, VA: NCTM.
- Neves, RG, Silva, JC, & Teodoro, VD. (2011). Improving learning in science and mathematics with exploratory and interactive computational modeling. *Trends in Teaching and Learning of Mathematical Modeling*, 1(4), 331–339.
- Ngu, BH, Yeung, AS, & Tobias, S. (2014). Cognitive load in percentage change problems: unitary, pictorial, and equation approaches to instruction. *Instructional Science*, 42(5), 685–713.
- Niss, M, Blum, W, & Galbraith, P. (2007). Introduction. In W Blum, P Galbraith, HW Henn, & M Niss (Eds.), *Modelling and applications in mathematics education: the 14th ICMI Study* (pp. 3–33). New York, NY: Springer.
- Panaoura, A. (2012). Improving problem solving ability in mathematics by using a mathematical model: a computerized approach. *Computers in Human Behavior*, 28(6), 2291–2297.
- Papadopoulos, I, & Dagdilelis, V. (2008). Students' use of technological tools for verification purposes in geometry problem solving. *The Journal of Mathematical Behavior*, 27(4), 311–325.
- Passey, D. (2012). Educational technologies and mathematics: signature pedagogies and learner impacts. *Computers in the Schools*, 29(1–2), 6–39.
- Pead, D, Ralph, B, & Muller, E. (2007). Uses of technologies in learning mathematics through modelling. In W Blum, P Galbraith, HW Henn, & M Niss (Eds.), *Modelling and applications in mathematics education: the 14th ICMI Study* (pp. 309–318). New York, NY: Springer.
- Pilli, O, & Aksu, M. (2013). The effects of computer-assisted instruction on the achievement, attitudes and retention of fourth grade mathematics students in North Cyprus. *Computers & Education*, 62, 62–71.
- Pollak, H. (2007). Mathematical modelling—a conversation with Henry Pollak. In W Blum, P Galbraith, HW Henn, & M Niss (Eds.), *Modelling and applications in mathematics education: the 14th ICMI Study* (pp. 109–120). New York, NY: Springer.
- Polya, G. (1957). *How to solve it*. Princeton, NJ: Lawrence Erlbaum.

- Porter, A, McMaken, J, Hwang, J, & Yang, R. (2011). Assessing the common core standards: opportunities for improving measures of instruction. *Educational Researcher*, 40(4), 186–188.
- Povey, H, & Ransom, M. (2000). Some undergraduate students' perceptions of using technology for mathematics: tales of resistance. *International Journal of Computers for Mathematical Learning*, 5(1), 47–63.
- Rakes, CR, Valentine, JC, McGatha, MB, & Ronau, RN. (2010). Methods of instructional improvement in algebra: a systematic review and meta-analysis. *Review of Educational Research*, 80(3), 372–400.
- Reid, DJ, Zhang, J, & Chen, Q. (2003). Supporting scientific discovery learning in a simulation environment. *Journal of Computer Assisted Learning*, 19, 9–20.
- Reimer, K, & Moyer, PS. (2005). Third-graders learn about fractions using virtual manipulatives: a classroom study. *Journal of Computers in Mathematics and Science Teaching*, 24(1), 5–25.
- Remillard, JT, & Bryans, MB. (2004). Teachers' orientations toward mathematics curriculum materials: implications for teacher learning. *Journal for Research in Mathematics Education*, 35(5), 352–388.
- Roschelle, J, Shechtman, N, Tatar, D, Hegedus, S, Hopkins, B, Empson, S, & Gallagher, LP. (2010). Integration of technology, curriculum, and professional development for advancing middle school mathematics: three large-scale studies. *American Educational Research Journal*, 47(4), 833–878.
- Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological bulletin*, 86(3), 638.
- Rothstein, HR, Sutton, AJ, & Borenstein, M. (2006). *Publication bias in meta-analysis: prevention, assessment and adjustments*. Chichester, UK: Wiley & Sons.
- Santos-Trigo, M. (2007). Mathematical problem solving: an evolving research and practice domain. *Mathematics Education*, 39(5/6), 523–536.
- Schwarz, CV, & White, BY. (2005). Metamodeling knowledge: developing students' understanding of scientific modeling. *Cognition and Instruction*, 23(2), 165–205.
- Shadish, WR, Cook, TD, & Campbell, DT. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Belmont, CA: Wadsworth Cengage Learning.
- Shin, N, Sutherland, LM, Norris, CA, & Soloway, E. (2012). Effects of game technology on elementary student learning in mathematics. *British journal of educational technology*, 43(4), 540–560.
- Slough, SW, & Rupley, WH. (2010). Re-creating a recipe for science instructional programs: adding learning progressions, scaffolding, and a dash of reading variety. *School Science and Mathematics*, 110(7), 352–362.
- Stern, L, Barnea, N, & Shauli, S. (2008). The effect of a computerized simulation on middle school students' understanding of the kinetic molecular theory. *Journal of Science Education and Technology*, 17(4), 305–315.
- Stillman, GA, & Galbraith, PL. (1998). Applying mathematics with real-world connections: metacognitive characteristics of secondary students. *Educational Studies in Mathematics*, 36, 157–195.
- Stokoe, R. (2012). Curiosity, a condition for learning. *The International Schools Journal*, 32(1), 63.
- Suh, J, & Moyer-Packenham, P. (2007). Developing students' representational fluency using virtual and physical algebra balances. *Journal of Computers in Mathematics and Science Teaching*, 26(2), 155–173.
- Tall, D. (1986). Using the computer as an environment for building and testing mathematical concepts: a tribute to Richard Skemp. In *Papers in honor of Richard Skemp* (pp. 21–36). Retrieved from <http://homepages.warwick.ac.uk/staff/David.Tall/pdfs/dot1986h-computer-skemp.pdf>
- Van Loon-Hillen, N, van Gog, T, & Brand-Gruwel, S. (2012). Effects of worked examples in a primary school mathematics curriculum. *Interactive Learning Environments*, 20(1), 89–99.
- Wander, R, & Pierce, R. (2009). Marina's Fish Shop: a mathematically- and technologically-rich lesson. *Australian Mathematics Teacher*, 65(2), 6.
- Willson, VL. (1983). A meta-analysis of the relationship between science achievement and science attitude: kindergarten through college. *Journal of Research in Science Teaching*, 20(9), 839–850.
- Xin, YP, Jitendra, AK, & Deatline-Buchman, A. (2005). Effects of mathematical word problem-solving instruction on middle school students with learning problems. *The Journal of Special Education*, 39(3), 181–192.
- Yimer, A, & Ellerton, NF. (2009). A five-phase model for mathematical problem solving: identifying synergies in pre-service-teachers' metacognitive and cognitive actions. *ZDM*, 42, 245–261.
- Zheng, X, Swanson, HL, & Marcoulides, GA. (2011). Working memory components as predictors of children's mathematical word problem solving. *Journal of Experimental Child Psychology*, 110(4), 481–498.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com
