



ELSEVIER

Contents lists available at ScienceDirect

Computers & Education

journal homepage: <http://www.elsevier.com/locate/compedu>

The potential of digital tools to enhance mathematics and science learning in secondary schools: A context-specific meta-analysis

Delia Hillmayr^{a,*}, Lisa Ziernwald^a, Frank Reinhold^a, Sarah I. Hofer^b,
Kristina M. Reiss^a

^a Technical University of Munich, Arcisstr. 21, 80333, Munich, Germany

^b Ludwig Maximilians University Munich, Leopoldstr. 13, 80802, Munich, Germany

ARTICLE INFO

Keywords:

Computer-based science learning
Mathematics
Media in education
Meta-analysis
Secondary education

ABSTRACT

Based on systematic research of studies published since the year 2000, this comprehensive meta-analysis investigated how the use of technology can enhance learning in secondary school mathematics and science (grade levels 5–13). All studies ($k = 92$) compared learning outcomes of students using digital tools to those of a control group taught without the use of digital tools. Overall, digital tool use had a positive effect on student learning outcomes ($g = 0.65, p < .001$). The provision of teacher trainings on digital tool use significantly moderated the overall effect. Use of intelligent tutoring systems or simulations such as dynamic mathematical tools was significantly more beneficial than hypermedia systems. On a descriptive level, the effect size was larger when digital tools were used in addition to other instruction methods and not as a substitute. The results open up new directions for future research and can inform evidence-based decision-making on the use of digital tools in education.

1. Introduction

There has been a growing body of research on the use of digital tools in school settings, recently (e.g., Cheung & Slavin, 2013; Ma, Adesope, Nesbit, & Liu, 2014; Steenbergen-Hu & Cooper, 2013). Numerous studies examined the effect of using digital tools on students' achievement. However, the media debate (Clark, 1994; Kozma, 1994) strengthened the argument that it might not be the mere medium that has an effect on learning outcomes. In consequence, contextual factors—such as teachers' views of digital media for teaching and learning or specific instructional design features of digital tools—came into focus. In the International Computer and Information Literacy Study (ICILS), for instance, teachers of 8th-grade students across 12 countries were asked for their attitudes toward digital tools (Fraillon, Ainley, Schulz, Friedman, & Duckworth, 2019, p. 183). The ICILS study shows that 87% of teachers across the participating countries think that ICT helps students to work at a level appropriate to their learning needs, and 78% state that ICT enables students to collaborate more effectively. Additionally, 91% of teachers agreed with the statement that information and communications technology (ICT) helps students develop greater interest in learning (Fraillon et al., 2019, p. 184). However, at least 23% of the participating teachers agreed that ICT also impedes concept formation by students and 37% state that ICT distracts students from learning (Fraillon et al., 2019, p. 185).

This ambiguity is also apparent when we take a closer look at studies investigating if and how digital tools impact teaching and

* Corresponding author. TUM School of Education, Germany.

E-mail address: delia.hillmayr@tum.de (D. Hillmayr).

<https://doi.org/10.1016/j.compedu.2020.103897>

Received 13 September 2018; Received in revised form 8 April 2020; Accepted 10 April 2020

Available online 15 April 2020

0360-1315/© 2020 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license

(<http://creativecommons.org/licenses/by/4.0/>).

learning. Despite the wide range of research studies and meta-analyses published in the last years (e.g., Al-Balushi, Al-Musawi, Ambusaidi, & Al-Hajri, 2017; Bayraktar, 2001/2002; Özyurt, Özyurt, Güven, & Baki, 2014; Perry & Steck, 2015; Van der Kleij, Feskens, & Eggen, 2015), the impact of using digital tools is not yet fully clear, as numerous studies differ in their findings. However, research shows that the use of digital tools can especially enhance learning in the context of technology-related subject matters within mathematics and science courses (Cheung & Slavin, 2013; Lesh, Post, & Behr, 1987; Ozdemir, Sahin, Arcagok, & Demir, 2018; Steenbergen-Hu & Cooper, 2014; Sung, Yang, & Lee, 2017; Van der Kleij et al., 2015). Today mathematical literacy, as well as science literacy, can be considered fundamental for social participation and “necessary for finding solutions to complex (real-world) problems” (OECD, 2016a, p. 6). However, international large scale studies as—for instance—PISA show that a substantial amount of students worldwide struggle with learning mathematics (e.g., OECD, 2016b, p. 192; OECD, 2019a, p. 106) and natural sciences (e.g., OECD, 2016b, p. 71; OECD, 2019a, p. 114). Here, the use of digital media yields high potential for the teaching and learning of mathematics (e.g., Gunbas, 2015), and science (e.g., Buckley et al., 2004; Chang, Chen, Lin, & Sung, 2008; Frailich, Kesner, & Hofstein, 2009, see Section 2.2.2). Therefore, this meta-analysis examines if and how the use of digital tools can especially enhance secondary school mathematics, biology, chemistry and physics. To this end, we analyzed existing published primary studies investigating the impact of using digital tools compared to instruction methods without the use of digital tools.

2. Theoretical rationale for using digital tools in teaching and learning

In Section 2, we first describe theoretical rationales for learning with interactive digital tools (Section 2.1), and second, how different types of interactive digital tools—and their corresponding characteristics—may enhance student learning (Section 2.2). In addition, we give an overview of prior research on digital tool use within mathematics and science learning as well as the influence of contextual factors on computer-supported student learning (Section 2.3).

2.1. Learning with interactive digital tools

With the three assumptions underlying the *cognitive theory of multimedia learning*, Mayer (2014) described why learning with digital tools can be beneficial: According to the *dual-channel assumption*, learners can organize information into two different cognitive structures, namely the visual and the auditory channel. The second assumption is the *limited capacity* of information processing in one channel. Therefore, it is favorable if learning environments stimulate the activation of both channels, the visual and auditory channel, in order to prevent a cognitive overload. This is possible, for example, by presenting sound images or spoken texts in combination with written texts or visual images. The third assumption is that learners need to engage *actively* with learning content in order to comprehend new information (Mayer, 2014). This is possible by the use of interactive learning environments, where the learner can actively and directly influence their own learning processes. In other words, “the defining feature of interactivity is responsiveness to the learner’s action during learning” (Moreno & Mayer, 2007, p. 310).

Such interactivity can be further categorized into dialoguing, controlling, and manipulating: Dialoguing means that the learner receives additional information on demand or feedback on his or her entered solutions. Interactivity by controlling occurs when the learner determines his or her individual learning pace or the preferred order of presentation. Finally, the learner can interact with learning environments by manipulating the presented information. This means that he or she “can control aspects of the presentation, such as setting parameters before a simulation runs, zooming in or out, or moving objects around the screen” (Moreno & Mayer, 2007, p. 311). Thus, in contrast to other instruction methods without these interactive features—where the learner passively receives information—an interactive learning environment enables learners to act as sense-makers constructing their own knowledge. Because “deep learning depends on cognitive activity” (Moreno & Mayer, 2007, p. 312) interactive tools are supposed to support student learning by offering specific characteristics such as the previously described dialoguing, controlling or manipulating.

2.2. Different types of interactive digital tools

In the following section, we describe different types of interactive digital tools with their corresponding characteristics (Section 2.2.1) and outline in greater detail how interactive features can enhance student learning of mathematics and science (Section 2.2.2).

Different types of interactive digital tools vary in regard to the instructional design features they provide and can hence be expected to differ in their impact on student learning. Therefore, research on the effectiveness of using digital tools in teaching and learning should focus more sharply on different *types* of tools (e.g., Higgins, Huscroft-D’Angelo, & Crawford, 2019). We follow the characterization by Nattland & Kerres, 2009 and divide digital tools into five categories—which are commonly used in educational settings—bearing different instructional design features: *drill and practice programs*, *tutoring systems*, *intelligent tutoring systems*, *simulations*, and *hypermedia systems*. Common to all of them is the availability of interactive features, however, interactivity occurs in different ways:

Whereas drill and practice programs serve to strengthen content knowledge acquired before by giving the learner the opportunity to practice at his or her own pace, and to repeat certain exercises as often as needed, tutoring systems can provide *new* knowledge content in small units and *additionally* offer opportunities for the learners to practice. The common characteristic of these two types is that learners get immediate feedback (Hattie & Timperley, 2007) on the correctness of the submitted solution. Intelligent tutoring systems additionally have adaptive features: they can present new content in consideration of the learners’ prior knowledge. Moreover, they allow individual adaptation of task difficulty or pace of presenting new content to learner needs. Moreover, they provide differentiated feedback or hints in order to support optimal learning processes (Nattland & Kerres, 2009). Especially tools with

adaptive features, like intelligent tutoring systems, can be expected to have a positive impact on student performance (Ma et al., 2014). Hypermedia systems are not designed to teach learning units in a structured way but serve as non-linear hypertext systems that are mostly used as encyclopedias within educational settings, providing information connected through hyperlinks. Simulations represent complex real-world situations that can be changed by manipulating different parameters. In general, simulations can be used to apply or expand knowledge by explorative learning. Prominent examples of such simulation programs are virtual laboratories. In addition, computer algebra systems—such as GeoGebra (see for instance Hohenwarter, Jarvis, & Lavicza, 2009)—also enable to learn in a comparable explorative way and can be considered as one kind of simulation (Hegedus, Dalton, & Tapper, 2015; Lichti & Roth, 2018). However, we use a separate category for these tools within our meta-analysis, because “a genre of software development called ‘Dynamic Mathematics’ has created a suite of tools to construct and interact with mathematical objects and configurations” (Hegedus et al., 2015, p. 205). For the purpose of this study, we divided the simulation category into *virtual reality* and *dynamic mathematical tools*. The latter stands for simulation programs that allow the manipulation of mathematical expressions.

2.2.1. Instructional design features within interactive learning environments

In the following section, we describe in greater detail how three instructional design features (i.e., feedback, pacing, and guided activity, see Moreno & Mayer, 2007) can enhance learning with interactive tools:

Feedback can be implemented into digital tools (e.g., Van der Kleij et al., 2015). There is empirical evidence that feedback can have a positive effect on students’ learning, yet, explanatory feedback is more beneficial than corrective feedback alone (Hattie & Timperley, 2007). Corrective feedback provides the correct answer only, whereas explanatory feedback provides information about *why* students’ answers are correct or incorrect and therefore can help overcome existing misconceptions (Moreno & Mayer, 2007). Regarding digital tools, Van der Kleij et al. (2015) investigated the impact of different methods of providing feedback in computer-based learning environments, such as providing the correct answer, providing the correctness of the answer or providing an explanation why the answer is correct or incorrect. They reported positive effects across all types of feedback; however, elaborated feedback, which e.g. provided explanations, showed largest effect sizes. Feedback providing the correct answer yielded larger effect sizes than feedback that only showed the correctness of the answer. In line with these findings, studies suggest that using intelligent tutoring systems—which give explorative feedback—can have a positive impact on learning outcomes (Belland, Walker, Kim, & Lefler, 2017; Ma et al., 2014; Van der Kleij et al., 2015), whereas negative effects for drill and practice programs—which only give corrective feedback—were found (Bayraktar, 2001/2002).

There is some evidence that pacing as well as guided activity can be beneficial features for student learning. Pacing as an example of controlling interactivity can be favorable for students, because they can take control over their own learning speed: Moreno (2006) found in her study that learning with the opportunity for pacing resulted in lower ratings of difficulty compared to learning without pacing. In a recent meta-analysis, Belland et al. (2017) analyzed the impact of scaffolding, which is one type of guided activity. They reported a positive overall effect of computer-based scaffolding on learning outcomes. More general, “guided activity enables students to interact with a pedagogical agent who guides their cognitive processing during learning” (Moreno & Mayer, 2007, p. 315).

Whereas all mentioned types of digital tools do offer control of speed by the learner, drill and practice programs as well as hypermedia systems do not offer guided activity. Intelligent tutoring systems combine *all* mentioned features (feedback, pacing and guided activity). Thus, they can be expected to yield high potential for improving learning.

Empirically, several meta-analyses focused on intelligent tutoring systems or on the impact of specific features of this type of tool, such as adaptivity, scaffolding or feedback. Ma et al. (2014) as well as Steenbergen-Hu and Cooper (2014) found positive overall effects of using intelligent tutoring systems ($g = 0.41$ and $g = 0.37$, respectively).

2.2.2. How the use of digital tools can enhance mathematics and science learning

Several studies show that the use of digital tools can especially enhance learning and teaching within technology-related subjects such as mathematics (e.g., Gunbas, 2015), physics (e.g., Chang et al., 2008), biology (e.g., Buckley et al., 2004) or chemistry (e.g., Fraulich et al., 2009). We illustrate the potential of digital tools for the teaching and learning in these subjects with a focus on mathematics, exemplarily, bearing in mind that there is a certain degree of distinctiveness in each of these domains regarding learning as well as teaching processes.

The use of digital tools can support skills and strategies that are highly relevant in the scientific and mathematical content area, such as real-world problem solving (Greefrath, Hertleif, & Siller, 2018; Huppert, Lomask, & Lazarowitz, 2002) or visualizing complex relationships (e.g., Koklu & Topcu, 2012). It can support learning through interactive and scaffolded activities (e.g., Reinhold et al., 2020). In addition, manipulating representations in computer simulations can support model-based learning—as students may understand mathematics and science concepts more elaborately because they observe direct consequences of the changes they make (Buckley et al., 2004). Furthermore, it can help students overcome “cognitive constraints originating from various misconceptions” (Jimoyiannis & Komis, 2001, p. 184). In mathematics, for instance, dynamic tools such as GeoGebra (Greefrath et al., 2018; Hohenwarter et al., 2009) enable students to learn abstract subjects, such as geometry, algebra and calculus in an interactive and explorative manner (Bhagat & Chang, 2015; Lichti & Roth, 2018; Shadaan & Leong, 2013). Such dynamic mathematical tools—as well as computer algebra systems which are still used remarkably low in schools (OECD, 2015, p. 56, p. 56)—can support mathematical learning, such as “understanding algebraic reasoning, finding patterns, and reflecting on the solution process” (Kramarski & Hirsch, 2003, p. 250). For instance, Bhagat and Chang (2015) found that the use of GeoGebra enhanced reasoning and visualization skills of students.

Regarding more general features, adaptive digital tools allow students to receive content according to their individual learning style (Section 2.2.1), which can especially be fruitful when students learn new and abstract mathematical concepts (Reinhold, Hoch,

Werner, Richter-Gebert, & Reiss, 2020; Özyurt et al., 2014). In addition, digital tools can also provide opportunities for students to practice content knowledge acquired before, which is important—for example—for fostering mathematical principles at a more basic level (Soliman & Hilal, 2016; Tienken & Wilson, 2007). By providing individual feedback to the learner immediately, specific tools aim to avoid developing typical misconceptions (Reinhold et al., 2020), which are often a problem in learning mathematics (e.g., Lortie-Forgues, Tian, & Siegler, 2015; Obersteiner, Van Hoof, Verschaffel & Van Dooren, 2016) and science (Jimoyiannis & Komis, 2001).

With a focus on not mere cognitive, but affective learning outcomes, there is evidence that the use of digital tools in teaching and learning mathematics can increase student motivation (e.g., Özyurt et al., 2014; Turk & Akyuz, 2016). One commonly used argument for this positive effect is derived from self-determination theory (Deci & Ryan, 2010, pp. 1–2): the opportunity to make own choices during the learning process (i.e., autonomy) and experiencing tasks as challenging but not overly complicated (i.e., competence) can be achieved via implementing educational features—such as feedback, pacing, and guided activity (Section 2.2.1)—into digital tools (Wouters, van Nimwegen, van Oostendorp, & van der Spek, 2013). Here, Özyurt et al. (2014) showed that students were more satisfied when they learned mathematics with the use of an intelligent tutoring system, and they were of the opinion that it facilitated understanding mathematics. In the same line, other researchers found a similar result when they asked students for their attitudes toward learning with the dynamic geometry system GeoGebra: “drawing in paper pencil environment could require drawing the shapes again and again. However, in computer it is easier and enjoyable” (Turk & Akyuz, 2016, p. 100).

2.3. Prior meta-analyses on learning with digital tools in mathematics and science subjects

In the following, we first illustrate the potential of meta-analysis in the field of educational research (Section 2.3.1). Second, we describe general findings of prior meta-analyses investigating the effects of digital tool use in secondary school mathematics and science subjects (Section 2.3.2). Third, we specifically describe the results of prior meta-analyses regarding contextual factors of computer-supported learning (Section 2.3.3). To interpret the effect sizes reported in prior studies, we apply a rule of thumb that is used frequently in other meta-analyses within the field (e.g., Bayraktar, 2001/2002; Belland et al., 2017; Sung et al., 2017): values for Cohen's d and Hedges's g between 0.20 and 0.50 can be considered small effects, values between 0.50 and 0.80 medium, and values over 0.80 large effects (Cohen, 1988). Although many different rules of thumb for effect-size interpretation exist, we restrict to this one in order to warrant better readability and comparability. Regarding the derivation of practical implications, categorizing differently sized effects should be done with caution (Lipsey et al., 2012). Thus, we always report the exact values of effect sizes, too.

2.3.1. The potential of meta-analysis

Due to a large number of studies and great interest in the impact of using digital tools on teaching and learning, as well as in the efficacy of the mentioned contextual factors, different researchers have conducted meta-analyses on the effects of computer-based mathematics and science learning. Research syntheses offer the opportunity to describe the status quo of a certain research field, which additionally allows for detecting research gaps (Borenstein, Hedges, Higgins, & Rothstein, 2009). Furthermore, meta-analyses can improve the possibility of generalizing the results, since different studies focus on different samples or were conducted in diverse settings at different times (Borenstein et al., 2009). In addition to contributing to developing and verifying theories, e.g., by testing for potential moderating effects, meta-analyses can also have an impact on political and practical decision processes by disseminating relevant research findings (Cumming, 2012).

2.3.2. Overall effects of using digital tools within mathematics and science learning

A comprehensive search for prior meta-analyses in the field of mathematics and science learning shows that they either investigate the general effects of digital tool use versus instruction methods without digital tool use or they focus on the impact of specific characteristics of computerized learning. Prior meta-analyses that compared the outcomes of technology-supported learning to those of other instruction methods found overall effects of different sizes.

Bayraktar (2001/2002) investigated the effect of computer-assisted instruction on student achievement in science subjects such as biology, chemistry or physics, and found a small positive overall effect of $d = 0.27$. An earlier meta-analysis by Kulik and Kulik (1991) showed the advantage of computer-based science and mathematics instruction compared to other methods as having a moderate but significant effect of $d = 0.30$. Sung, Chang, and Liu (2016) focused on the effect of integrating mobile devices in teaching and learning in comparison to instruction methods without digital tools and also found a moderate but slightly larger effect of $g = 0.52$.

Two other recent meta-analyses investigated the potential of computer use to support collaborative learning. Sung et al. (2017) found an effect of $g = 0.47$ for computer-supported collaborative learning compared to instruction methods without computer use. A significant positive effect of $g = 0.45$ for computer use in collaborative learning environments on knowledge gain was also reported by Chen, Wang, Kirschner, and Tsai (2018).

2.3.3. Contextual factors of learning with digital tools in science and mathematics

Such positive overall effects of digital tool use in mathematics and science classrooms may be one reason for the effort to improve computer equipment at schools today. However, evidence shows that “accessibility does not always imply an improved learning environment” (Higgins et al., 2019, p. 285, see also the media debate; Clark, 1994; Kozma, 1994). Thus, several contextual factors of learning with digital tools should be considered as well because they can influence the effectiveness of digital tool use on student achievement (e.g., Bayraktar, 2001/2002; Cheung & Slavin, 2013; Steenbergen-Hu & Cooper, 2013; Sung et al., 2017). Here, prior meta-analyses considered the student-to-computer-ratio, compared different grade levels and school subjects or replacing versus supplemental use of digital tools.

The student-to-computer-ratio and its beneficial impact on student achievement—which can be understood as an indicator for collaborative learning within computerized learning environments—was reported by [Chen et al. \(2018\)](#) as well as [Sung et al. \(2017\)](#). [Frailich et al. \(2009\)](#) stated that the use of digital tools in groups resulted in more productive collaborative work among students. However, [Bayraktar \(2001/2002\)](#) found that the effect was larger when students used digital tools on their own compared to pairwise use or use in larger groups.

Regarding different grade levels, prior meta-analyses yield diverse results: in one prior meta-analysis different effect sizes between grade levels were reported ([Sung et al., 2017](#)), whereas other meta-analyses found no differences in the effect of digital tools on student learning between different grade levels ([Bayraktar, 2001/2002](#); [Steenbergen-Hu & Cooper, 2013](#)).

Comparing the effects of computer use in different school subjects, the effects for mathematics and sciences was found to be largest at $g = 0.89$, respectively $g = 0.77$ ([Sung et al., 2017](#)). However, [Cheung and Slavin \(2013\)](#) only detected a small effect of the use of digital tools on student achievement in mathematics, $d = 0.16$, and [Steenbergen-Hu and Cooper \(2013\)](#) an overall effect of $g = 0.09$, which can be considered negligible.

Regarding replacing versus supplemental use, there is evidence that a supplemental use of digital tools in the regular classroom produced greater learning outcomes than the complete replacement of other instruction methods ([Bayraktar, 2001/2002](#); [Cheung & Slavin, 2013](#)).

3. The present study

Previous meta-analyses investigating the effects of computer-supported learning either focused on one type of digital tool, such as intelligent tutoring systems ([Ma et al., 2014](#); [Steenbergen-Hu & Cooper, 2014](#)), were limited to only one school subject ([Cheung & Slavin, 2013](#)), or focused on studies from one specific country ([Bayraktar, 2001/2002](#)). Moreover, there is a rapidly developing market of educational software and appropriate hardware. Thus, the consistently increasing number of studies investigating the impact of learning with digital tools make it difficult to maintain a current overview of this topic. Here, it is important for researchers and practitioners to understand “what the comparative effects of different approaches are” ([Chen et al., 2018](#), p. 804). In addition, the diverse results for contextual factors of learning with digital tools in science and mathematics (Section 2.3.3) suggest that there is need for further clarification regarding the impact of student-to-computer-ratio, grade levels, school subjects, or replacing versus supplemental use of digital tools. Moreover, other contextual factors—such as the duration of using digital tools ([Jeno, Vandvik, Eliassen, & Grytnes, 2019](#)), or the level of guidance by teacher or peers (see [Frailon et al., 2014](#), p. 22) during learning with digital tools—are of equal importance but seem underrepresented in recent meta-analyses.

Hence, there is a need for an updated comprehensive meta-analysis on the effects of using digital tools particularly in mathematics, chemistry, physics and biology on secondary school students' academic performance that is not limited to certain countries or particular types of tools and that considers a broad variety of contextual factors.

3.1. Potential moderators considered

The current study does not only analyze the overall effect of the use of digital tools in mathematics and science classrooms, but also takes into account several specific conditions, which are expected to be more or less favorable for student learning. These contextual factors are treated as potential moderators affecting the described overall effect.

First, in order to analyze the impact of tool-specific features, *type of digital tool* is treated as a potential moderator. As portrayed in Section 2.2, the scheme used in this study is built upon the classification scheme of [Nattland & Kerres, 2009](#), discriminating the types of digital tools most commonly used in schools. This allows an overview and direct comparison of the effects regarding frequently used digital tools that were divided into (a) drill and practice programs, (b) tutoring systems, (c) intelligent tutoring systems, (d) dynamic mathematical tools (e) virtual reality, and (f) hypermedia systems. Dynamic mathematical tools include computer algebra systems or dynamic geometry software to manipulate geometric constructions. The category virtual reality covers, for example, virtual chemistry laboratories representing complex real world situations, which can also be changed by manipulating parameters ([Nattland & Kerres, 2009](#)). The six types of digital tools are not always exclusive because they partially overlap due to similar features ([Nattland & Kerres, 2009](#)). Thus, they were classified according to their dominant function. Game-based learning was not considered in this study, since educational games can be very different from each other in terms of their concept or features and there is a lack of well-designed boundaries within the educational research field ([Gros, 2007](#)). Therefore, an exact categorization into our scheme would be difficult, because the features always overlap with at least one other type of digital tool. Yet, recent meta-analyses cover the topic of gamification of learning (e.g., [Sailer & Homner, 2019](#)).

As mentioned earlier, contextual factors of learning environments are considered as potential moderator variables within the current study.

The first contextual factor is the *student-to-computer-ratio*. We investigate whether the impact of learning with digital tools is more beneficial if students use tools on their own, pairwise or in groups.

The second contextual moderator variable is *school subject*. To gain more insight into the divergent results between school subjects (mathematics, physics, chemistry, and biology) within prior studies, we also consider this aspect in the analyses.

Another potential moderator is the *school level* in which digital tools are used. To clarify the empirical inconsistency within prior studies, grade level is considered within the current analyses, too. Since we focus on the advanced primary and the secondary level, studies were ranked into even categories of grade levels 5 to 7, 8 to 10, and 11 to 13.

A further interest lies in the importance of *support for students* when using digital tools in class. As technology constantly develops

and computer-based instruction in schools is still rather unfamiliar for many teachers (Fraillon, Ainley, Schulz, Friedman, & Gebhardt, 2014, p. 207) as well as for students (Fraillon et al., 2014, p. 22), it follows that specific support for using media is pivotal. Thus, we explore the impact of support for students provided by teachers and/or peers versus no support as another potential moderator variable.

In order to investigate whether a supplemental use of digital tools in the regular classroom has greater impact on learning outcomes than if they completely replace methods without digital tools, the *type of computer use* is considered within moderator analyses in this study.

The teacher's competence in using digital tools in class is a central determinant for successful student learning. However, teacher self-assessments show that a considerable proportion of teachers do not feel sufficiently educated in this area (Fraillon et al., 2014, p. 207). For this reason, providing *teacher training* as a potential factor for positive student learning outcomes is also investigated here. Providing specific teacher trainings are recognized as a study quality issue and means a larger expense for the researchers. It is assumed that if teacher training is available, that the authors would also report it. Thus, if a study does not report on the presence of a specific training, it is coded as *no teacher training*.

Furthermore, there are some methodological features that were already identified by several prior meta-analyses as potential moderators (e.g., Chen et al., 2018). These study characteristics are *duration of the interventions*, *sample size*, *study design* and the *instructor effect*.

The *duration of interventions* moderated the effect in prior studies, as shorter interventions resulted in higher learning outcomes (Bayraktar, 2001/2001; Sung et al., 2017). In the present study, duration is categorized into interventions with durations of one day to six days, one week to four weeks, four weeks to six months, and more than 6 months, because prior studies found significant differences between these categories (Chen et al., 2018; Sung et al., 2017).

Empirically, studies with smaller *sample sizes* produce larger effects (Chen et al., 2018; Cheung & Slavin, 2013; Slavin and Smith, 2009). Therefore, in order to explore differences between sample sizes within the current study, we categorized the studies into sample sizes of 100 or fewer, sample sizes between 101 and 500, or more than 500. For further analyses, all studies were additionally categorized into only two categories of more or fewer than 100.

In prior meta-analyses considering *study design*, such as in Belland et al. (2017) and Chen et al. (2018), the type of randomization was reported as a significant moderator. They found larger effects for quasi-experimental studies than for randomized studies, which is thus considered within our study as well.

The last potential moderator analyzed in our study is the so-called *instructor effect*, focusing on whether the intervention in the experimental and control groups were conducted by the same or by different persons. Prior studies found larger effect sizes when the treatment and the control group were instructed by different people rather than by the same teacher or researcher (Bayraktar, 2001/2002; Kulik & Kulik, 1991).

Some studies additionally reported findings on student outcomes such as emotional and motivational orientations, which were considered in an additional explorative analysis within the current study. The term *attitude* was used to summarize these student outcomes since it generally describes "different psychological concepts with diverse theoretical backgrounds, such as emotional and motivational orientations, self-related cognitions, strategies, and value beliefs" (Schiepe-Tiska, Roczen, Müller, Prenzel, & Osborne, 2017, p. 306). Motivational student outcomes such as interest in science and mathematics or instrumental forms of motivation can influence student career choices later in life (Dowker, Sarkar, & Looi, 2016; Schiepe-Tiska et al., 2017). Because of the low student interest in sciences in many OECD countries found in PISA 2015 (OECD, 2016b, p. 125), and the current need for qualified employees in the field of science and engineering (OECD, 2016b, p. 110), it seems important to investigate the potential of digital tool use to have a positive effect on students' attitude in these fields.

3.2. Research questions

The purpose of this meta-analysis is to clarify the impact of learning mathematics and sciences with digital tools on the performance (and attitudes toward the taught subject) of secondary school students. It addresses the following research questions:

- Do secondary school students learning with digital tools in mathematics and science classes have different learning outcomes (and attitudes) compared to students learning without the use of digital tools?
- Which conditions of learning with digital tools in mathematics and science classes are favorable with regard to student learning outcomes?

4. Method

In the following section we provide a detailed overview of the inclusion criteria (Section 4.1) and describe the method of the literature search (Section 4.2), the process of coding studies (Section 4.3) including measures for interrater reliability (Section 4.4), and data analysis (Section 4.5). For a detailed introduction to meta-analysis see—for example—Borenstein et al. (2009).

4.1. Inclusion criteria

All studies had to meet the following pre-defined inclusion criteria in order to ensure maximum quality standards:

- Primary data were reported in the study.
- Digital tools (computer, tablet, smartboard, mobile phone, notebook, or CAS computer) were used during mathematics, physics, chemistry, biology or science class in general (and were *not* additionally used at home).
- The sample consisted of secondary-school students (grade levels 5 to 13).
- The sample did *not* consist only of students with special educational needs (e.g., only gifted or only disabled students).
- The dependent variable was student performance and *optionally* student attitudes in addition.
- The study had a pre-post-control-group design.
- The study did not investigate the effects of computer games.
- Effect sizes or data necessary for effect size calculation were reported in the study.
- The control group consisted of students taught with instruction methods not using digital tools.
- The study was published between 2000 and October 5, 2018.
- The study was published in a peer-reviewed journal and a full text was available in English or German.

4.2. Literature search

The literature search was conducted in three relevant major databases, specifically, *Web of Science*, *Scopus* and *ERIC*. The databases were chosen to cover different disciplines, with *Web of Science* providing studies in the field of social sciences, arts, and humanities, *Scopus* providing studies in the field of natural sciences and technology, and *ERIC* providing studies in the field of education research.

We searched for studies published between 2000 and October 5, 2018. This timeframe was chosen because the databases used in this study show the first significant increase in the number of publications on the effectiveness of learning with digital tools between 2000 (e.g., 5 results in *Scopus*) and 2005 (e.g., 43 results in *Scopus*). In the *Scopus* database, for example, the literature search for studies published between 1975 and 2000 yielded one to a maximum three results per year.

With the following syntax, $k = 6572$ studies were found in all three databases: “study” OR “*empiric**” OR “research” AND “digital media” OR “tablet” OR “computer” OR “whiteboard” OR “smartboard” OR “ipad” OR “pc” OR “cas” OR “ict” OR “netbook” OR “software” AND “stem” OR “math*” OR “mint” OR “*physic**” OR “chemistry” OR “biology” OR “science” AND “secondary school” OR “high school” OR “secondary education” OR “middle school” NOT “computer science” NOT “informatics” NOT “engineering.” The 6572 studies were then limited to articles published in peer-reviewed journals, since peer review is “a generally accepted criterion to ensure scientific quality” (Van der Kleij et al., 2015, p. 502). The terms were searched within titles, abstracts or keywords. Three filters were used to refine the results, namely document types (limited to articles), publication years (between 2000 and October 5, 2018), and research areas, such as psychology, education, mathematics, chemistry, biology and physics.

After initially screening titles and abstracts as well as removing duplicates, 474 articles remained that met the inclusion criteria. During the detailed coding process using the full texts, additional studies were excluded, because they did not fit the inclusion criteria that are stated above. Therefore, the final dataset consisted of 92 studies from 91 articles (in one of the articles, two independent studies were reported) with a total of $N = 14,910$ students. The entire study selection process is described in a flowchart in Fig. 1, following the guidelines of The PRISMA Group (Moher, Liberati, Tetzlaff, & Altman, 2009).

Most studies were conducted in Turkey ($k = 22$), followed by studies conducted in the US ($k = 21$), Taiwan ($k = 10$), Israel ($k = 5$), Slovenia ($k = 5$), Malaysia ($k = 4$), Indonesia ($k = 3$), The Netherlands ($k = 3$), Nigeria ($k = 3$), Germany ($k = 2$), Australia ($k = 1$), Canada ($k = 1$), Ghana ($k = 1$), India ($k = 1$), Iran ($k = 1$), Korea ($k = 1$), Kenya ($k = 1$), Kuwait ($k = 1$), South Africa ($k = 1$), Oman ($k = 1$), Lebanon ($k = 1$) and Singapore ($k = 1$). Two articles did not report where the study was conducted.

4.3. Coding strategy

The coding form included the following items: type of digital tool, grade level (5–7, 8–10 or 11–13), subject (mathematics, biology, chemistry, physics or science in general), student-to-computer-ratio, sample size, duration of the study, whether digital tools were used supplemental to other existing methods in class or if they replaced them, student support by teacher and/or peers, provision of teacher training, randomization (quasi-experimental or experimental design), and whether the interventions in the experimental and control group were conducted by different persons or by the same person. All 92 studies were coded independently by at least two raters.

4.4. Interrater reliability

The detailed coding form developed for this synthesis was piloted by coding six studies that were excluded from the final data set. As a value for interrater reliability, Cohen’s Kappa (Cohen, 1968) was calculated for each variable separately, because it was expected that some of the variables would yield lower values than others because of heterogeneous underlying theoretical concepts used in the analyzed studies. Indeed, values for Cohen’s Kappa ranged from $\kappa = 0.24$ (type of digital tool), which is considered a fair agreement, to $\kappa = 0.87$ (subject), which represents nearly perfect agreement (Landis & Koch, 1977). The mean of all values for Cohen’s Kappa was $\kappa = 0.43$ ($SD = 0.23$). The agreement for variables without any scope for interpretation, such as data for effect size calculation, was almost perfect. Each particular disagreement was discussed in regular team meetings in which three of the authors took part until a consensus was reached. For this purpose, the relevant full texts were considered.

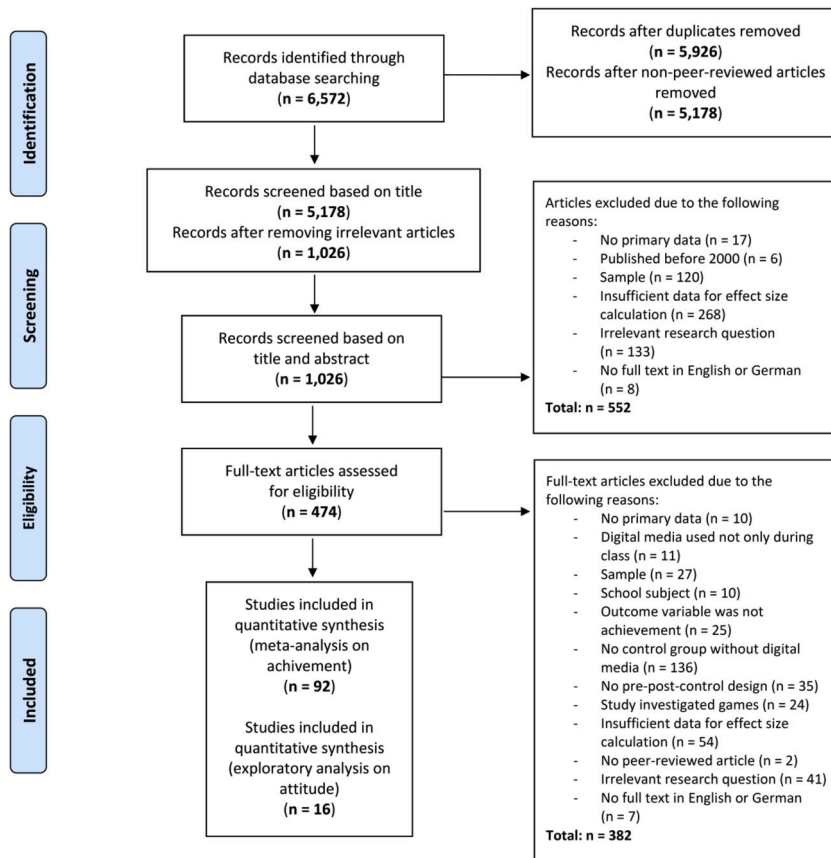


Fig. 1. Flowchart of the study selection process following the guidelines of The PRISMA Group (Moher et al., 2009).

4.5. Effect size calculation and data analysis

After the data set was completed, the data was imported to the *Comprehensive Meta-analysis* (CMA) program for further analyses (Borenstein et al., 2009). In order to calculate the overall effects, all available effect sizes were transformed to the effect size Hedges's g (Hedges, 1981). If studies did not report any effect size, relevant data was used for calculations. In addition to the calculations of the weighted mean effect size of Hedges's g with its standard errors and 95% confidence intervals around each mean, CMA was also used to test for homogeneity by calculating Q , p , and I^2 . The Q statistic was inspected to examine whether all analyzed studies share a common population effect size. Q is approximately chi-square-distributed with $k - 1$ degrees of freedom. If Q exceeds the critical value of the distribution, this indicates that the effect sizes significantly vary between the studies (Shadish & Haddock, 2009). In addition, I^2 was used as a descriptive statistic that represents the amount of real variance within the observed variance of the effect sizes. An I^2 of over 75% can be considered high variance. In the latter case, subgroup analyses are appropriate for finding reasons for the variance of the effect sizes (Borenstein et al., 2009).

Hedges's g is comparable to Cohen's d and therefore represents the standardized mean difference, which is calculated by dividing the difference between experimental and control group means by the pooled standard deviation (Hedges, 1981). Although the two effect sizes are closely similar, the difference between them is that "Cohens's d uses N for the denominator of the estimated variance to obtain the standard deviation, whereas Hedges's g uses $N - 1$, that is, the pooled within-sample unbiased estimate of the population variance to obtain the standard deviation" (Rosnow & Rosenthal, 2003, p. 223). Thus, because Cohen's d tends to overestimate the standardized mean differences in smaller samples (Borenstein et al., 2009), and we only have access to the standard deviation of the sample and not to the population standard deviation, it is more evident to use Hedges's g within the current analyses. The interpretation of Hedges's g is equivalent to Cohen's d (see Section 2.3).

Some studies were special in that they compared, for example, two experimental groups with only one control group, and therefore the data was not independent. In case of multiple dependent comparisons within a study, effect sizes were combined by calculating their mean. Moreover, the correlations between the subgroups were computed based on the exact group sizes. These correlations were taken into account when calculating the variance of the composite effect size (see Borenstein et al., 2009).

The second case of dependent data was the report of multiple outcomes per study and therefore of one and the same group. In this case, the correlations of the outcomes were set to one ($r = 1$). This approach tends to overestimate the standard error and to underestimate precision, which is therefore a more conservative method for combining dependent outcomes than to assume completely

independent outcomes (see Borenstein et al., 2009).

For all analyses, a *random-effects model* was used. Differences in the effect sizes among studies under the use of a random-effects model are not only due to sampling error but also due to systematic variance between the studies (Borenstein et al., 2009). Because the 92 studies were conducted by independent researchers with different ways of implementing interventions and of using digital tools, one cannot assume that heterogeneity among studies might not be influenced by these different circumstances. Moreover, different variables and study designs were used in the analyzed studies. Therefore, a random-effects model seems more appropriate than a fixed-effect model (Borenstein et al., 2009; Cooper, 2010; Lipsey & Wilson, 2001). Furthermore, the results of this meta-analysis should be generalizable to students of various grade levels and in various school subjects, hence, this is why a random-effects model is the more obvious choice.

Moderator analyses were conducted to investigate whether different study features (e.g., type of digital tool or randomization) have a differently sized impact on the overall effect. Therefore, each study was classified into one of the corresponding categories of every potential moderator variable. If there was no information available, the category *not reported* was used. The difference between the respective subgroups was tested with Q_B , which stands for the heterogeneity among the groups and is equivalent to the F -value within ANOVA. Thus, a significant Q_B means that there is a statistically significant heterogeneity among the subgroups and that the moderator variable can partially explain the heterogeneity among the effect sizes (Hedges & Pigott, 2004). For all analyses, the significance level was set at $\alpha = 0.05$.

To analyze whether a publication bias could have influenced the results of the study, fail-safe N (Rosenthal, 1979) was computed. Furthermore, we applied a rank correlation test, which analyzed the correlation between the standard error and the effect size. This inverse correlation is expected because studies with smaller sample sizes are more often included in meta-analysis if they show large treatment effects (Begg & Mazumdar, 1994). The funnel-plot-based method of trim-and-fill was used as a third approach to assess the threat of potential publication bias. With this method, unmatched values in the distribution can be trimmed to obtain a more symmetric funnel plot and missing values are imputed. Then, the overall effect can be recalculated by taking the additional values into consideration (Duval & Tweedie, 2000).

5. Results

5.1. Descriptive statistics

In total, the 92 studies yielded 117 effect sizes regarding the student learning outcomes of $N = 14,910$ students. All studies with their characteristics and effect sizes are presented in Table 1. We combined multiple outcome measures per study as described above and therefore used one effect size per study as independent units for the analyses. These 92 effect sizes range from $g = -0.33$ to $g = 2.46$, representing the minimum and maximum effect sizes, respectively. Eighty of the effect sizes (87%) are positive, whereas twelve studies (13%) yielded negative effect sizes, indicating that digital tools adversely affected student learning outcomes. Analyses regarding the effects of digital tool use on student attitudes revealed a total of 16 effect sizes ranging from $g = -2.24$ to $g = 1.59$. The minimum effect size of $g = -2.24$ is the only negative one, whereas 15 of the 16 effect sizes indicate a positive impact of digital tool use on student attitudes toward the subject. Fig. 2 shows the forest plot of the 92 effect sizes regarding student learning outcomes with the standardized mean difference for each study and their corresponding 95% confidence intervals.

5.2. Outlier analysis

Based on the one-study-removed-approach (Borenstein et al., 2009), the outlier analysis in CMA shows that all 92 studies have effect sizes related to student learning outcomes that fall within the 95% confidence interval of the average effect size. Thus, no study was identified as an outlier. The analysis of the 16 effect sizes related to student attitudes also revealed no obvious outlier.

5.3. Overall effect on student learning outcomes

The first research question focused on the benefits of using digital tools in science or mathematics classes on student learning outcomes in comparison to classes learning without digital tools. The overall effect shows that the use of digital tools had a medium positive and statistically significant effect on student learning, $g = 0.65$, 95% CI [0.54, 0.75], $p < .001$. Hence, secondary school students who learned with the use of digital tools in science or mathematics classes had significantly greater learning outcomes than students that were taught without the use of digital tools.

The test for heterogeneity showed that the effect sizes varied significantly between the studies with $Q(91) = 757.92$, $p < .001$, and with $I^2 = 87.99$, indicating large heterogeneity.

5.4. Publication bias

To test for potential publication bias, fail-safe N was computed (Rosenthal, 1979). The value of fail-safe N shows how many non-significant studies are missing in the analyses to nullify the positive effect found: 8891 non-significant studies are therefore needed to invalidate the observed overall effect. The limit of $5k + 10$ studies suggested by Rosenthal (1979), which is 470 in the current study, was therefore far exceeded. Additionally, we used the rank correlation test as another method for analyzing the threat of potential publication bias by analyzing the rank correlation between the study size and the effect size (Begg & Mazumdar, 1994). The positive

Table 1
Study Characteristics with Corresponding Effect Size (Hedges's g) and Standard Error.

Study	Grade Level	Subject	Type of Digital Tool	Use of Tool by	Support by	Type of Computer Use	Teacher Training	Instruction in Treatment & Control Group by	Study Design	Effect Size	Standard Error
Abanikamda (2016)	Not reported	Chemistry	Hypermedia Learning	Not reported	Teacher	Replacing	No	Different teachers/researchers	Quasi-experimental	1.43**	0.29
Adegoke (2010)	8-10	Physics	Virtual Reality	Not reported	Not reported	Supplemental	Yes	Different teachers/researchers	Quasi-experimental	0.82**	0.11
Aktas and Aydın (2016)	7	Physics	Not reported	Not reported	Teacher	Supplemental	Yes	Different teachers/researchers	Quasi-experimental	0.56*	0.23
Al-Batushi et al. (2017)	12	Chemistry	Virtual Reality	One student	No support	Supplemental	No	Different teachers/researchers	Quasi-experimental	0.29	0.26
Aliasgari, Riahinia, and Mojdehavar (2010)	11-13	Mathematics	Tutoring System	Not reported	Teacher	Not reported	No	Different teachers/researchers	Quasi-experimental	0.95**	0.30
Ardac and Sezen (2002)	8-10	Chemistry	Virtual Reality	One student	Teacher	Replacing	No	Same teacher/researcher	Quasi-experimental	-0.01	0.27
Ayub, Tarmizi, Bakar, and Luan Wong (2012)	8-10	Mathematics	Drill & Practice	Not reported	No support	Not reported	No	Not reported	Quasi-experimental	1.27**	0.27
Baki and Güvelci (2008)	8-10	Mathematics	Hypermedia Learning	Not reported	Teacher	Supplemental	No	Same teacher/researcher	Quasi-experimental	0.45*	0.22
Baran and Maskan (2013)	11-13	Physics	Virtual Reality	More than 3 students	Teacher and peers	supplemental	Yes	Not reported	Quasi-experimental	1.48**	0.27
Basaran and Gönen (2012)	8-10	Physics	Virtual Reality	Not reported	Not reported	Supplemental	No	Different teachers/researchers	Quasi-experimental	0.73**	0.18
Basham and Kotrlík (2008)	8-10	Mathematics	Virtual Reality	Two students	Not reported	Supplemental	Yes	Different teachers/researchers	Quasi-experimental	0.05	0.11
Beal, Arroyo, Cohen, and Woolf (2010)	5-7	Mathematics	ITS	Not reported	Teacher and peers	Supplemental	No	Different teachers/researchers	Quasi-experimental	0.20	0.39
Bhagat and Chang (2015)	8-10	Mathematics	Dynamic mathematical tool	Not reported	Not reported	Replacing	Yes	Same teacher/researcher	Quasi-experimental	1.12**	0.30
Buckley et al. (2004)	8-10	Biology	Virtual Reality	Not reported	Not reported	Replacing	No	Same teacher/researcher	Quasi-experimental	0.49*	0.24
Çepni, Taş, and Köse (2006)	11-13	Biology	Tutoring System	More than 3 students	Peers	Replacing	No	Same teacher/researcher	Quasi-experimental	0.64*	0.28
Chandra and Briskey (2012)	8-10	Mathematics	Hypermedia Learning	One student	Teacher and peers	Supplemental	Yes	Different teachers/researchers	Quasi-experimental	0.58	0.29
Chang et al. (2008)	8-10	Physics	Virtual Reality	One student	Not reported	Replacing	No	Not reported	Quasi-experimental	0.80**	0.19
Chang, Sung, and Chen (2001)	5-7	Biology	Tutoring System	One student	Not reported	Supplemental	No	Not reported	Quasi-experimental	0.32	0.30
Cheng, Cheng, and Chen (2012)	Not reported	Biology	Tutoring System	Not reported	Teacher	Replacing	No	Not reported	Quasi-experimental	0.45*	0.21
Dhindsa and Shahrizal-Emran (2011)	11-13	Chemistry	Virtual Reality	Not reported	Teacher and peers	Replacing	No	Different teachers/researchers	Quasi-experimental	-0.10	0.23
Drijvers, Doorman, Kirschner, Hoogveld, and Boon (2014)	8-10	Mathematics	Tutoring System	One student	Not reported	Replacing	No	Same teacher/researcher	Quasi-experimental	-0.14	0.07
Ektici and Pekmezci (2015)	5-7	Physics	Virtual Reality	More than 3 students	Not reported	Supplemental	No	Not reported	Quasi-experimental	0.63*	0.30
Faour and Ayoubi (2018)	10	Physics	Virtual Reality	Two students	Teacher and peers	Supplemental	Yes	Same teacher/researcher	Experimental	0.62*	0.29

(continued on next page)

Table 1 (continued)

Study	Grade Level	Subject	Type of Digital Tool	Use of Tool by	Support by	Type of Computer Use	Teacher Training	Instruction in Treatment & Control Group by	Study Design	Effect Size	Standard Error
Frailich et al. (2009)	8–10	Chemistry	Hypermedia Learning	Two students	Teacher and peers	Supplemental	Yes	Different teachers/researchers	Quasi-experimental	0.76**	0.14
Funkhouser (2002)	11–13	Mathematics	Not reported	Not reported	Not reported	Supplemental	No	Not reported	Not reported	0.41	0.29
Gelbart, Brill, and Yarden (2009)	11–13	Biology	Virtual Reality	Not reported	Teacher	Supplemental	Yes	Different teachers/researchers	Quasi-experimental	0.82**	0.22
Graff, Mayer, and Lebens (2008)	5–7	Mathematics	ITS	One student	Teacher	Supplemental	No	Different teachers/researchers	Quasi-experimental	1.18**	0.17
Gregorius, Santos, Dano, and Gutierrez (2010)	5–7	Chemistry	Tutoring System	One student	Teacher	Supplemental	No	Not reported	Not reported	0.10**	0.18
Gunbas (2015)	5–7	Mathematics	Tutoring System	One student	Teacher	Supplemental	No	Not reported	Experimental	0.53*	0.22
Gürbüz and Birgin (2012)	5–7	Mathematics	Tutoring System	More than 3 students	Teacher and peers	Supplemental	No	Same teacher/researcher	Experimental	0.78*	0.34
Harskamp and Suhre (2007)	Not reported	Mathematics	ITS	One student	Not reported	Supplemental	Yes	Different teachers/researchers	Quasi-experimental	1.56**	0.16
Hegedus et al. (2015)	Not reported	Mathematics	Dynamic mathematical tool	Not reported	Teacher and peers	Replacing	Yes	Different teachers/researchers	Quasi-experimental	0.35**	0.07
Herga and Grmek (2014)	5–7	Chemistry	Virtual Reality	Not reported	Not reported	Supplemental	No	Different teachers/researchers	Quasi-experimental	1.70**	0.22
Huppert et al. (2002)	8–10	Biology	Virtual Reality	One student	Teacher	Supplemental	No	Different teachers/researchers	Quasi-experimental	0.34*	0.15
Johnson-Glenberg, Birchfield, Tolentino, and Koziupa (2014)	11–13	Chemistry	Virtual Reality	Two students	Teacher and peers	Supplemental	No	Same teacher/researcher	Quasi-experimental	0.91**	0.31
Johnson-Glenberg et al. (2014)	11–13	Biology	Virtual Reality	More than 3 students	Peers	Supplemental	No	Same teacher/researcher	Quasi-experimental	0.49	0.29
Jones (2002)	8–10	Biology	Hypermedia Learning	More than 3 students	Teacher and peers	Replacing	No	Different teachers/researchers	Quasi-experimental	−0.10	0.23
Kaheru and Kriek (2016)	11–13	Physics	Virtual Reality	More than 3 students	Teacher	Supplemental	Yes	Different teachers/researchers	Quasi-experimental	0.67**	0.21
Kara and Yeşilyurt (2008)	8–10	Biology	Tutoring System	One student	Teacher	Replacing	Yes	Same teacher/researcher	Quasi-experimental	0.88**	0.30
Koedinger, McLaughlin, and Heffernan (2010)	5–7	Mathematics	Tutoring System	One student	No support	Supplemental	No	Different teachers/researchers	Quasi-experimental	−0.23**	0.07
Koklu and Topcu (2012)	8–10	Mathematics	Virtual Reality	Not reported	Teacher	Supplemental	Yes	Same teacher/researcher	Quasi-experimental	0.64*	0.30
Kolcak and Mogol (2014)	8–10	Physics	Virtual Reality	Not reported	Not reported	Not reported	No	Same teacher/researcher	Quasi-experimental	1.07**	0.30
Kramarski and Hirsch (2003)	8–10	Mathematics	Dynamic mathematical tool	Not reported	Not reported	Supplemental	No	Not reported	Quasi-experimental	0.49	0.30
Kwon and Cifuentes (2007)	8–10	Not reported	Not reported	Not reported	Teacher	Replacing	Yes	Same teacher/researcher	Quasi-experimental	0.88**	0.28
Kwon and Cifuentes (2009)	5–7	Not reported	Not reported	Not reported	Teacher	Replacing	No	Same teacher/researcher	Quasi-experimental	1.30**	0.19
Lalley, Piotrowski, and Battaglia (2010)	11–13	Biology	Virtual Reality	Two students	Peers	Replacing	No	Different teachers/researchers	Quasi-experimental	0.93**	0.21

(continued on next page)

Table 1 (continued)

Study	Grade Level	Subject	Type of Digital Tool	Use of Tool by	Support by	Type of Computer Use	Teacher Training	Instruction in Treatment & Control Group by	Study Design	Effect Size	Standard Error
Lamb and Annetta (2013)	11–13	Chemistry	Virtual Reality	One student	Teacher	Supplemental	No	Different teachers/researchers	Quasi-experimental	0.21	0.11
Leong (2013)	11–13	Mathematics	Dynamic mathematical tool	One student	Teacher and peers	Supplemental	No	Not reported	Quasi-experimental	1.70**	0.35
Lin, Cheng, and Chang (2002)	8–10	Biology	Hypermedia Learning	One student	Not reported	Replacing	No	Same teacher/researcher	Quasi-experimental	0.16	0.23
Lin, Chang, Hou, and Wu (2016)	8–10	Physics	Not reported	One student	Teacher and peers	Replacing	No	Same teacher/researcher	Quasi-experimental	0.49	0.40
Looi and Lim (2009)	5–7	Mathematics	ITS	Two students	Teacher and peers	Supplemental	No	Different teachers/researchers	Quasi-experimental	1.05**	0.26
Lundeberg, Bergland, and Klyczek (2003)	Not reported	Biology	Tutoring System	Not reported	Peers	Supplemental	No	Same teacher/researcher	Quasi-experimental	1.39**	0.41
Marbach-Ad, Rotbain, and Stavay (2008)	11–13	Biology	Virtual Reality	One student	Teacher and peers	Supplemental	No	Different teachers/researchers	Quasi-experimental	0.86**	0.16
McLaughlin and Munsell (2012)	8–10	Biology	Tutoring System	Not reported	Not reported	Replacing	Yes	Different teachers/researchers	Quasi-experimental	1.73**	0.16
Mihindo, Wachanga, and Anditi (2017)	Not reported	Chemistry	Virtual Reality	Not reported	Not reported	Not reported	Yes	Different teachers/researchers	Quasi-experimental	1.08**	0.23
Moyer-Packenham and Suh (2012)	5–7	Mathematics	ITS	One student	Teacher	Replacing	No	Same teacher/researcher	Quasi-experimental	−0.21	0.40
Mustafa and Trudel (2013)	11–13	Physics	Virtual Reality	More than 3 students	Teacher and peers	Not reported	No	Same teacher/researcher	Quasi-experimental	−0.15	0.28
Nugraini, Choo, Hin, and Hoon (2013)	8–10	Biology	Hypermedia Learning	One student	Not reported	Replacing	No	Not reported	Quasi-experimental	0.44**	0.13
Onder and Aydin (2016)	8–10	Biology	Not reported	Not reported	Not reported	Not reported	Yes	Same teacher/researcher	Quasi-experimental	2.45**	0.37
Osman and Lee, 2013	11–13	Chemistry	Tutoring System	One student	Teacher	Replacing	No	Same teacher/researcher	Quasi-experimental	0.40*	0.18
Owusu, Monney, Appiah, and Wilmot (2010)	11–13	Biology	Tutoring System	One student	No support	Replacing	No	Different teachers/researchers	Quasi-experimental	−0.08	0.23
Özmen (2008)	11–13	Chemistry	Virtual Reality	Two students	Teacher	Supplemental	Yes	Same teacher/researcher	Quasi-experimental	0.58*	0.26
Özmen, Demircioğlu, and Demircioğlu (2009)	11–13	Chemistry	Virtual Reality	Two students	Teacher	Supplemental	Yes	Different teachers/researchers	Quasi-experimental	2.23**	0.36
Özyurt et al. (2014)	11–13	Mathematics	ITS	One student	No support	Replacing	No	Same teacher/researcher	Quasi-experimental	0.58**	0.20
Perry and Steck (2015)	Not reported	Mathematics	Virtual Reality	Not reported	Teacher	Supplemental	No	Different teachers/researchers	Quasi-experimental	−0.33	0.19
Poehnl and Bogner (2013)	8–10	Biology	Hypermedia Learning	Two students	Peers	Supplemental	No	Different teachers/researchers	Quasi-experimental	−0.03	0.11
Poł, Harskamp, and Suhre (2005)	Not reported	Physics	Tutoring System	One student	Not reported	Supplemental	No	Not reported	Quasi-experimental	0.63	0.36
Puhek, Perse, and Sorgo (2012)	8–10	Biology	Virtual Reality	Not reported	Peers	Replacing	No	Different teachers/researchers	Quasi-experimental	−0.12	0.14
Rahman, Ismail, and Nasir (2014)	11–13	Physics	Virtual Reality	Not reported	Not reported	Not reported	No	Different teachers/researchers	Quasi-experimental	2.46**	0.31
Roschelle et al. (2010)	8-10	Mathematics	Virtual Reality	Not reported	Teacher	Supplemental	Yes	Different teachers/researchers	Quasi-experimental	0.59**	0.04

(continued on next page)

Table 1 (continued)

Study	Grade Level	Subject	Type of Digital Tool	Use of Tool by	Support by	Type of Computer Use	Teacher Training	Instruction in Treatment & Control Group by	Study Design	Effect Size	Standard Error
Sengel and Özden (2010)	8–10	Physics	Virtual Reality	Two students	Peers	Supplemental	Yes	Same teacher/researcher	Quasi-experimental	0.88**	0.26
Shadaan and Leong (2013)	8–10	Mathematics	Dynamic mathematical tool	Not reported	Teacher	Supplemental	No	Same teacher/researcher	Quasi-experimental	1.08**	0.29
She and Lee (2008)	5–7	Physics	ITS	Not reported	Peers	Replacing	No	Not reported	Quasi-experimental	1.28**	0.28
Shieh (2012)	11–13	Physics	Tutoring System	Not reported	Peers	Supplemental	No	Different teachers/researchers	Quasi-experimental	0.43*	0.21
Shim, Park, Kim, and Kim (2003)	8–10	Biology	Virtual Reality	One student	Not reported	Supplemental	No	Not reported	Quasi-experimental	0.78**	0.24
Shirvani (2010)	Not reported	Mathematics	Tutoring System	Not reported	Teacher	Supplemental	Yes	Same teacher/researcher	Quasi-experimental	0.18	0.18
Soliman and Hilal (2016)	5-7	Mathematics	Drill & Practice	One student	Teacher	Supplemental	Yes	Same teacher/researcher	Quasi-experimental	0.74**	0.26
Starbek, StarčićErjavec, and Peklaj (2010)	11–13	Biology	Virtual Reality	Not reported	Not reported	Supplemental	No	Different teachers/researchers	Quasi-experimental	0.64**	0.14
Stieff (2011)	11–13	Chemistry	Virtual Reality	Two students	Teacher and peers	Supplemental	Yes	Different teachers/researchers	Quasi-experimental	0.25*	0.10
Taasobshirazi, Zuiker, Anderson, and Hickey (2006)	11–13	Physics	Tutoring System	Two students	Teacher and peers	Replacing	Yes	Different teachers/researchers	Not reported	0.53*	0.23
Tatli and Ayas (2013)	8–10	Chemistry	Tutoring System	Not reported	No support	Replacing	No	Same teacher/researcher	Quasi-experimental	0.65	0.34
Tienken and Wilson (2007)	5–7	Mathematics	Drill & Practice	One student	Teacher	Not reported	No	Different teachers/researchers	Quasi-experimental	0.34*	0.12
Tunaboylu and Demir (2017)	5-7	Mathematics	Not reported	Not reported	Teacher	Supplemental	Yes	Different teacher/researchers	Quasi-experimental	0.67*	0.27
Turk and Akyuz (2016)	8-10	Mathematics	Dynamic mathematical tool	Not reported	Not reported	Not reported	No	Same teacher/researcher	Experimental	1.84**	0.39
Ülen, Çagran, Slavinec, and Gerlič (2014)	11–13	Physics	Tutoring System	Two students	Teacher and peers	Supplemental	No	Not reported	Quasi-experimental	0.88**	0.23
Vrtačnik, Sajovec, Dolničar, Pučko-Razdevsek, Glazar, and Brouwer (2000)	11–13	Chemistry	Tutoring System	Two students	Peers	Supplemental	No	Not reported	Not reported	1.30**	0.34
Wang (2011)	5–7	Mathematics	Drill & Practice	One student	Not reported	Supplemental	No	Same teacher/researcher	Quasi-experimental	0.09	0.22
Yang and Heh (2007)	8-10	Physics	Tutoring System	More than 3 students	Peers	Supplemental	No	Same teacher/researcher	Quasi-experimental	0.57**	0.17
Yenmez, Özpınar, and Gökçe (2017)	8-10	Mathematics	Hypermedia Learning	More than 3 students	Teacher and peers	Supplemental	No	Same teacher/researcher	Quasi-experimental	0.20	0.22
Yildirim, Ozden, and Aksu (2001)	8–10	Biology	Hypermedia Learning	One student	Not reported	Replacing	No	Not reported	Experimental	0.24	0.38
Yuan, Lee, and Wang (2010)	8–10	Mathematics	Virtual Reality	More than 3 students	Teacher and peers	Replacing	No	Not reported	Quasi-experimental	–0.04	0.25
Zulnaidi and Zakaria (2012)	Not reported	Mathematics	Virtual Reality	Not reported	Not reported	Not reported	No	Same teacher/researcher	Quasi-experimental	0.41*	0.18

Note. ITS = Intelligent tutoring system.

* $p < .05$. ** $p < .01$.

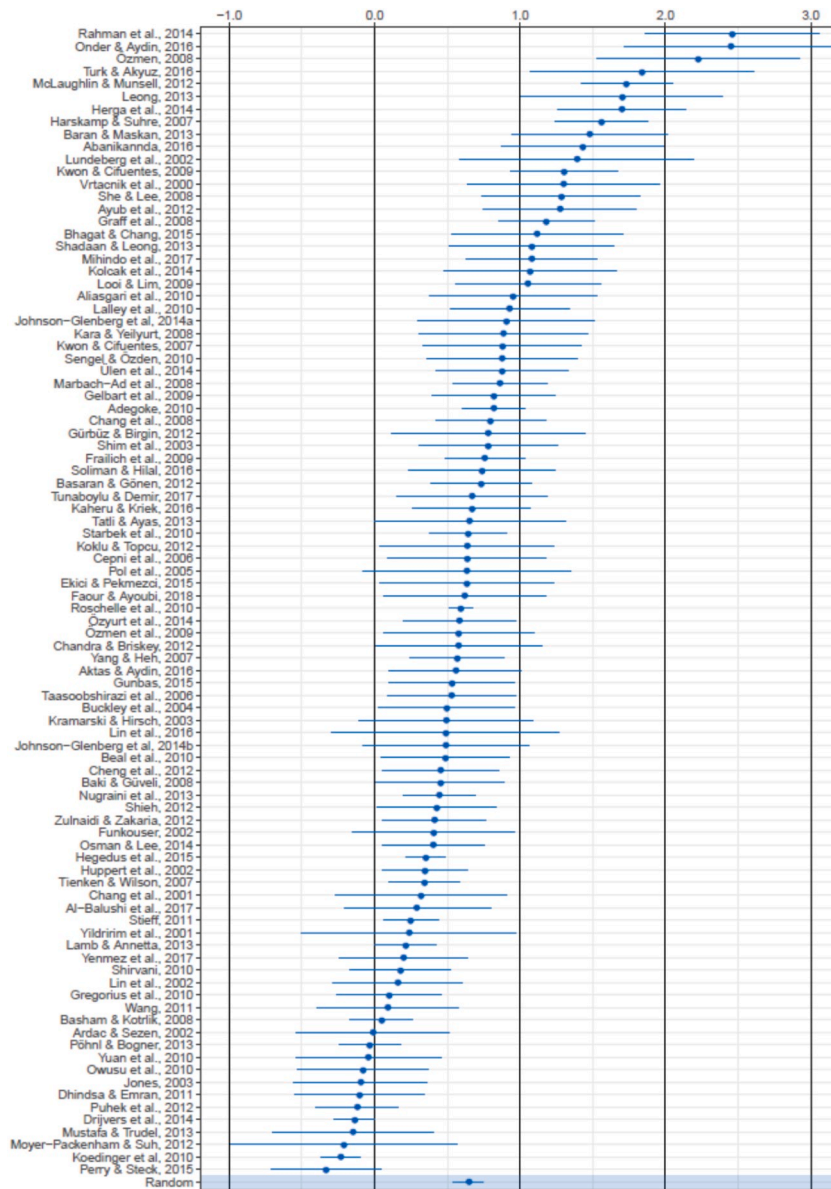


Fig. 2. Forest plot of effect sizes (Hedges's g) for the 92 studies and overall effect regarding student learning outcomes and their corresponding 95% confidence intervals.

result of the test with Kendall's $\tau = 0.18, p < .01$, indicated a significant correlation between the treatment effect and the standard error, and thus suggested that the data basis of this meta-analysis might be skewed by publication bias. Furthermore, we applied the trim-and-fill method (Duval & Tweedie, 2000) as a third approach to assess the possible effect of publication bias. First, a funnel plot was used to visualize the distribution of the effect sizes around the mean, which is presented in Fig. 3. The analysis indicated that twenty-eight studies on the left side of the distribution were missing to make it symmetrical. Next, the overall effect size was computed again by considering the 28 additional values, and an adjusted effect size of $g = 0.35, p < .05$ was found. The funnel plot with the imputed studies is presented in Fig. 4. Although the adjusted effect size is smaller than the observed effect, there is no evidence that the positive overall effect size found in this study was affected significantly by publication bias. In other words, the observed effect size in this meta-analysis is robust, because the adjusted effect is still significantly positive.

5.5. Conditions of learning with digital tools

The second research question focused on the conditions of learning with digital tools in mathematics or science classes, and which one is more or less favorable relative to student learning outcomes. Because the test for heterogeneity was significant, moderator

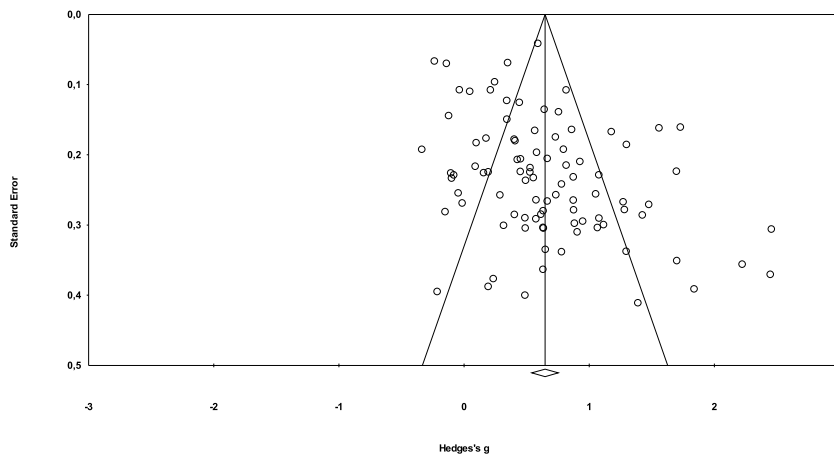


Fig. 3. Funnel plot of standard error by Hedges's g . Distribution of all 92 included studies.

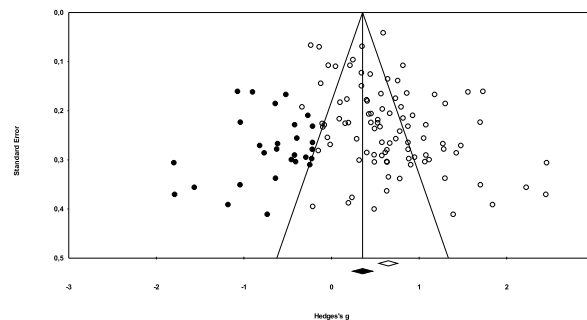


Fig. 4. Funnel plot of standard error by Hedges's g after trim-and-fill. Distribution of 92 observed and 28 imputed studies (black circles).

analyses were conducted to find out if potential moderating variables were responsible for the significant variance among studies and which conditions are more favorable for student learning outcomes than others.

Most of the studies analyzed the effects of using digital tools supplementary to non-digital material ($k = 53$), in mathematics classes ($k = 33$), and in grade levels 8 to 10 ($k = 37$). Simulations (dynamic mathematical tools and virtual reality) were most often used as type of tool ($k = 42$). In most of the analyzed studies, students used digital tools on their own ($k = 29$) and were supported by their teachers when learning with digital tools ($k = 28$). All coded variables (including the number of studies k in each category) and the results of the moderator analyses are presented in Table 2. The table shows the results of using a random-effects model, including the number of students (N), the number of studies (k), the weighted mean effect size (g), standard errors (SE), 95% confidence intervals, and the tests for heterogeneity (Q_B). Studies that did not report on the respective variable were excluded from the moderator analyses.

We tested whether learning outcomes of students were greater if their teachers had received specific training before they used digital tools in their class. The variable teacher training produced statistically significant between-levels variance, $Q_B(1) = 5.53$, $p < .05$. Hence, the positive effect on student learning outcomes is significantly larger if teachers received specific training for using digital tools before the intervention was conducted ($g = 0.84$, $p < .05$) compared to interventions without specific teacher training ($g = 0.56$, $p < .05$).

Moderator analyses were conducted with all coded variables and revealed that teacher training was the only variable that significantly moderated the overall effect in the current meta-analysis when using a random-effects model. Although the between-levels variance was not statistically significant for the remaining variables, the following results show tendencies of more or less beneficial learning conditions and methodological study features.

Regarding the use of different types of digital tools, we expected a positive effect for tools with adaptive or feedback features. The analyses show that dynamic mathematical tools ($g = 1.02$, $p < .05$) and intelligent tutoring systems ($g = 0.89$, $p < .05$) produced the largest and statistically significant effect sizes. Smaller but still medium effects resulted for the use of virtual reality ($g = 0.63$, $p < .05$) and the use of tutoring systems ($g = 0.55$, $p < .05$). The use of drill and practice programs also produced a medium effect size ($g = 0.58$, $p < .05$). The smallest effect size was found for studies that investigated the use of hypermedia learning systems ($g = 0.40$, $p < .05$). Pairwise comparative analyses showed significantly greater learning outcomes for the use of dynamic mathematical tools than for hypermedia systems ($p < .05$) and also significantly greater outcomes for intelligent tutoring systems than for hypermedia systems ($p < .05$).

Table 2
Moderator Analyses and Weighted Mean Effect Sizes for all Coded Variables.

Moderator Variables	N	k	Random-effects model				Q _B	P
			Effect size		95% CI			
			g	SE	Lower	Upper		
Subject								
Biology	2580	22	0.59*	0.11	0.38	0.81	3.22	.36
Chemistry	1986	16	0.69*	0.13	0.44	0.94		
Mathematics	8086	33	0.55*	0.09	0.37	0.72		
Physics	2023	19	0.80*	0.12	0.57	1.03		
Not reported	235	2						
Grade level							0.68	.71
5–7	2775	18	0.62*	0.13	0.37	0.87		
8–10	7464	37	0.61*	0.09	0.44	0.78		
11–13	2963	27	0.71*	0.10	0.51	0.92		
Not reported	1708	10						
Use of tool by							2.68	.26
One student	5069	29	0.46*	0.09	0.28	0.64		
Two students	2138	14	0.72*	0.14	0.45	0.99		
More than 3 students	767	11	0.46*	0.16	0.15	0.77		
Not reported	6936	38						
Type of digital tool							7.62	.18
Drill & Practice	492	4	0.58*	0.25	0.10	1.07		
Hypermedia Learning	1281	10	0.40*	0.16	0.09	0.71		
Intelligent Tutoring System	676	7	0.89*	0.20	0.50	1.27		
Tutoring System	3812	22	0.55*	0.11	0.33	0.76		
Virtual Reality	7087	36	0.63*	0.08	0.47	0.79		
Dynamic mathematical tools	1071	6	1.02*	0.22	0.59	1.44		
Not reported	491	7						
Type of computer use							1.16	.28
Supplemental to other methods	9918	53	0.64*	0.07	0.50	0.77		
Replacing other methods	4142	29	0.51*	0.09	0.33	0.70		
Not reported	850	10						
Student support by							1.43	.70
teacher	5258	28	0.61*	0.08	0.44	0.77		
peers	1321	11	0.63*	0.14	0.36	0.89		
Teacher and peers	2618	20	0.54*	0.10	0.34	0.74		
No support	1536	5	0.37	0.20	-0.01	0.76		
Not reported	4177	28						
Teacher training							5.53	.02
Yes	6643	27	0.84*	0.10	0.65	1.03		
No or not reported	8267	65	0.56*	0.06	0.44	0.69		
Duration of study							1.75	.63
1–6 days	709	5	0.86*	0.22	0.43	1.28		
1–4 weeks	3077	36	0.67*	0.09	0.50	0.84		
4 weeks–6 months	3366	21	0.63*	0.11	0.42	0.84		
More than 6 months	3811	5	0.46*	0.22	0.02	0.89		
Not reported	3947	25						
Sample size I							3.04	.08
N ≤ 100	3472	59	0.72*	0.07	0.59	0.86		
N > 100	11,438	33	0.54*	0.08	0.37	0.70		
Sample size II							4.84	.09
N ≤ 100	3472	59	0.72*	0.07	0.59	0.86		
N > 100 & ≤ 500	5579	28	0.59*	0.09	0.41	0.77		
N > 500	5859	5	0.28	0.21	-0.13	0.68		
Randomization							0.23	.63
Experimental design	235	5	0.77*	0.25	0.28	1.27		
Quasi-experimental design	14,366	83	0.65*	0.06	0.53	0.76		
Not reported	309	4						
Instruction in treatment and control group by							0.57	.90
Same teacher	2788	28	0.60*	0.10	0.40	0.80		
Different teachers	9786	37	0.64*	0.09	0.47	0.80		
Same researcher	288	6	0.79*	0.24	0.33	1.25		
Different researchers	592	2	0.70	0.36	-0.01	1.41		
Not reported	1456	19						

Note. CI = confidence interval. Studies that did not report on the moderator variable were excluded from the analyses.

*p < .05.

< .05). Hence, students had the greatest learning gains when they used simulation programs, such as dynamic geometry software or adaptive tools, and the effects are considered large when compared to learning gains of students taught without the use of digital tools.

The investigated use of digital tools was either supplementary or other existing methods have been completely replaced. The supplementary use ($g = 0.64, p < .05$) comes along with a larger effect size than if other existing instruction methods were substituted completely ($g = 0.51, p < .05$), however, there is no significant difference.

Digital tools were used either by students individually, by students in groups, or pairwise. The effect size is largest if digital tools were used pairwise ($g = 0.72, p < .05$), but no significant differences were found. All effect sizes are presented in Table 2.

In most of the studies, students were supported by the teacher and/or their peers while they learned with digital tools. Learning with support by peers ($g = 0.63, p < .05$), by teachers ($g = 0.61, p < .05$) or by teachers and peers ($g = 0.54, p < .05$) yielded larger effect sizes than learning without any support by peers or by teachers ($g = 0.37$), but the differences are not significant.

The largest effect sizes resulted from interventions that took place in physics lessons ($g = 0.80, p < .05$), followed by chemistry ($g = 0.69, p < .05$), biology ($g = 0.59, p < .05$), and mathematics lessons ($g = 0.55, p < .05$). All effect sizes are statistically significant, however, the between-levels variance was not significant.

Secondary school students in grades 5 to 13 were considered in the current meta-analysis. The use of digital tools shows a statistically significant positive effect across all grade levels 5 to 7 ($g = 0.61, p < .05$), 8 to 10 ($g = 0.55, p < .05$), and 11 to 13 ($g = 0.73, p < .05$). Although the effect is larger for the latter than for the lower levels, there is no significant difference between them.

The following variables relate to methodological features of the studies. The first methodological feature analyzed in the current study is the duration of the intervention. The analyses show that the shortest interventions, that is, one day (there was no study with a duration of two to six days), produced the largest effect size ($g = 0.86, p < .05$), whereas the longest interventions with use of digital tools for over six months produced the smallest effect size ($g = 0.47, p < .05$). Effect sizes are significant across all other durations, as shown in Table 2, however, there is no significant difference among the categories.

Type of assignment of students to either the experimental or the control group is the second methodological feature analyzed in the current meta-analysis. Most of the studies used a quasi-experimental design ($k = 83$), whereas only five studies used an experimental design with randomized assignment to the different groups. The effect size for experimental design studies ($g = 0.77, p < .05$) was larger than for quasi-experimental design studies ($g = 0.65, p < .05$), but was not significantly different.

Effect sizes were slightly larger if interventions in experimental group and control group were conducted by the same researcher ($g = 0.79, g < 0.05$) than if interventions were conducted by different external researchers ($g = 0.70$), and the latter effect is not significant. Conversely, the effect of learning with digital tools was slightly larger if interventions were conducted by different teachers ($g = 0.64, p < .05$) than by the same teacher ($g = 0.60, p < .05$). There was no significant difference between the characteristics of this potential moderator.

The last methodological feature analyzed in this study is the sample size. Studies with sample sizes of more than 500 ($g = 0.28$) showed non-significant and smaller effect sizes than studies with sample sizes of 100 or fewer ($g = 0.72, p < .05$). The between-levels variance was not significant, however, there is a significant difference between large sample sizes of more than 500 and small sample sizes of 100 or fewer ($p < .05$).

5.6. Overall effect on student attitudes

Additional explorative analysis of the analyzed studies focused on the benefits of using digital tools in science or mathematics classes on student attitudes toward the subject taught, in comparison to classes learning without digital tools. Sixteen of the 92 studies provided effect sizes regarding the attitude of students ($N = 1639$). It is important to note that these 16 studies investigated attitudes *in addition* to student learning outcomes. The overall effect shows that the use of digital tools had a small positive and statistically significant effect on student attitudes toward the subject taught, $g = 0.45, p < .05$. Hence, secondary school students who were taught using digital tools in science or mathematics classes had significantly more positive attitudes toward the *subject taught* than students who learned without the use of digital tools. The small number of studies that investigated student attitudes ($k = 16$) did not allow for further moderator analyses.

6. Discussion

6.1. Summary of results

This meta-analysis examined how the use of digital tools influenced student learning outcomes in mathematics and science classrooms, as well as student attitudes related to the subject. The overall effects show that the use of digital tools had a medium, significantly positive effect on student learning outcomes and a small, significantly positive effect on student attitudes. Moderator analyses were conducted for student learning outcomes. The variable teacher training moderated the overall effect significantly: Interventions that provided teacher training in the digital tool used in class produced significantly larger effects than studies that did not provide specific training. Looking at the different types of digital tools, the meta-analysis shows that dynamic mathematical tools and intelligent tutoring systems tended to yield larger effect sizes than drill and practice and were significantly more effective than hypermedia learning systems. Although the sample size did not significantly moderate the overall effect within the current study, comparative pairwise analyses showed the significantly larger effect for smaller samples than for large samples.

All further analyses yield no significant differences, however, they show that the effects were significantly positive across all analyzed subjects and grade levels, although the benefit seemed greatest for physics classes and for grade levels 11 to 13. Students

tended to profit most if digital tools were used pairwise, and in addition to other instruction methods. Moreover, the effect was larger for students who received support from their teachers or peers when learning with digital tools, as compared to those who did not receive support.

The small number of studies in most subgroups and the use of a random-effects model resulted in reduced statistical power, which may partly explain non-significant effects and between-levels variance in the moderator analyses (Jackson & Turner, 2017). Apart from that, reporting effect sizes is important to interpret research findings with regard to practical relevance (Sink & Stroh, 2006) whereas “statistical significance does not necessarily imply that the result is important in practice” (Ellis & Steyn, 2003, p. 51).

6.2. Comparison of results with prior studies

6.2.1. Overall effect

In general, the overall effect of learning with digital tools found in the current meta-analysis is in line with prior studies that have also found advantages in favor of learning with digital tools (Ma et al., 2014; Steenbergen-Hu & Cooper, 2014). The somewhat larger overall effect size in comparison to the prior meta-analyses like Bayraktar (2001/2002) with $d = 0.27$, Ma et al. (2014) with $g = 0.41$, Steenbergen-Hu and Cooper (2014) with $g = 0.37$, Steenbergen-Hu and Cooper (2013) with $g = 0.09$, and Cheung and Slavin (2013) with $d = 0.16$ could result from further development of digital tools and learning programs in recent years.

6.2.2. Duration of digital tool use

One reason for the relatively small impact on student learning outcomes for longer interventions could be that researchers are usually more involved in shorter interventions than in long-term use of digital tools in schools. Therefore, “the degree of implementation might have impacted the effectiveness” (Steenbergen-Hu & Cooper, 2013, p. 984). Another explanation could be the occurrence of a novelty effect (Clark & Sugrue, 1990). Among others, the fact that students were taught with the use of digital tools, and therefore with new technology and different methods than usual, could be responsible for short-term increases in motivation or interest, and as a consequence, indirectly lead to better performance. The results are in line with the findings of Bayraktar (2001/2002) and—except for interventions lasting only one week or less—with those of Sung et al. (2017). Conversely, the low overall effect found by Cheung and Slavin (2013) could be explained by the fact that they only analyzed studies with intervention durations greater than 12 weeks. The assumption of a novelty effect is not in line with the result of Ma et al. (2014), who found that interventions with durations of less than four weeks denote less-positive student learning outcomes. However, the opposite result of Ma et al. (2014) could be linked with the fact that their analysis was limited to intelligent tutoring systems, which are intended to systematically build up new knowledge, and therefore may not result in positive learning outcomes in a relatively short period of less than four weeks.

6.2.3. Grade level

On a descriptive level, the current study shows that the learning outcomes for students using digital tools are slightly greater for students in grade levels 11 to 13 than for students in grade levels 5 to 10. Hence, although the difference is not significant, this result is in line with the assumption of Steenbergen-Hu and Cooper (2014): They argued in their discussion that the use of intelligent tutoring systems might be more beneficial for older students who have greater self-regulation skills and computer literacy as well as greater prior knowledge and learning motivation. However, the results of Ma et al. (2014) did not comply with this assumption. Further research is needed to find answers to this open question.

6.2.4. Type of digital tool

With regard to different types of digital tools, the study shows the largest effect sizes for simulations, such as dynamic mathematical tools and for intelligent tutoring systems, which conforms to the results of Bayraktar (2001/2002), since he also found that simulations and tutorials were more beneficial than drill-and-practice programs. The small effect size regarding hypermedia systems found in the current study might be due to the less structured way of learning with such systems and is therefore in line with the result that learning without guidance is also less beneficial. Interactive features such as feedback, activation of relevant knowledge, and adaption of learning content to prior student knowledge, which are part of intelligent tutoring systems, could be responsible for the stronger effect of intelligent tutoring systems. This is in line with the findings of studies by Belland et al. (2017) or Van der Kleij et al. (2015) and also with the *guided discovery principle* within the *cognitive theory of multimedia learning*, which emphasizes the beneficial effect of feedback and hints provided by digital learning environments (Mayer, 2014). Simulation programs such as dynamic geometry software also show large effect sizes, and simulations like virtual laboratories show medium positive effect sizes. This result is in accordance with both the *modality principle* and the *multimedia principle* within the *cognitive theory of multimedia learning*, which imply that the presentation of information in terms of sounds and words as well as words and pictures lead to better learning results than solely using words (Mayer, 2014). Furthermore, the greater extent of learner control and learning through discovery and exploration might also be responsible factors for beneficial learning with simulation programs (Karich, Burns, & Maki, 2014). One reason for the smaller effect size regarding drill-and-practice programs could be that such programs do not impart new knowledge, but seek to strengthen already learned content. This could, for instance, result in redundant information, especially for learners with above-average prior knowledge, and lead to less efficient learning according to the *redundancy principle* of the *cognitive theory of multimedia learning* (Mayer, 2014). Drill-and-practice programs do not adapt to prior knowledge, which could be one reason for the smaller effect sizes and could explain why Bayraktar (2001/2002) even found negative effects for these programs.

6.2.5. Student-to-computer-ratio

Unlike the results of Bayraktar (2001/2002), the current study shows on a descriptive level that pairwise use of digital tools by students yields larger effect sizes than if they use media on their own. The greater learning gains when students worked together in pairs might have been a consequence of more interactive and communicative learning. Frailich et al. (2009) reported in their study that students were more cooperative and discussed difficulties more often when they used digital tools in small groups. However, the results of the current meta-analysis did not yield greater learning outcomes when digital tools were used in groups compared to individual use, unlike the results of Chen et al. (2018). Yet, the latter did not differentiate between two or more students within their analyses. One reason for our result could also be that computer-based collaborative learning can only be effective if there are specifically designed tasks for more than one student (Chen et al., 2018). Further research is needed to examine which student-to-computer-ratio works best, and especially whether this ratio is correlated with certain types of digital tools.

6.3. Teacher training

Studies that provided specific teacher training before the interventions were conducted produced significantly larger effect sizes than studies that did not provide such training. Although this result seems obvious, it emphasizes the importance of particular trainings for teachers to successfully use digital tools. Given the fact that a large number of teachers do not feel competent in the area of digital tool use (Fraillon et al., 2014, p. 207), this finding might be a consequence of lacking educational and pedagogical content during teacher training at university and should be considered in further research and teacher education (Mishra & Koehler, 2006).

6.4. Limitations of the present study

6.4.1. Missing information in primary studies

A large number of studies was not considered in the current meta-analysis because necessary statistical data was not reported. Out of 474 potentially relevant journal articles found in the three databases, only 92 studies could finally be used for the analyses. Studies were excluded not only because they lacked statistical data, but also because of other missing information that is important for meta-analyses, and because these studies lacked sufficient quality criteria. As already stated by Ma et al. (2014), this result emphasizes the need for greater transparency in reporting standards, which is especially important in an interdisciplinary research field such as the area of technology-supported learning. Moreover, the low values of Cohen's Kappa for some coded variables show how important it is not only to use consistent reporting standards but also to use coherent definitions of equivalent theoretical concepts and to describe study features as detailed as possible.

6.4.2. Threat of publication bias

This study focused on peer-reviewed publications, because they ensure high quality standards. Apart from that, there is "no way of objectively retrieving (unpublished) work" (Van der Kleij et al., 2015, p. 502). Therefore, the threat of publication bias was taken into account by calculating and reporting fail-safe N , the rank correlation test, and the trim-and-fill method, which partly indicated that the exclusion of unpublished studies could have resulted in a biased overall effect of learning with digital tools. One limitation of the rank correlation method is that a significant correlation can indicate the existence of publication bias but does not say anything about its direct impact on the study results. Therefore, we also applied the trim-and-fill method, which showed that the adjusted effect size was still significantly positive. Moreover, there is empirical evidence that the threat of publication bias is frequently overestimated (Hunter & Schmidt, 2004). The fact that "most studies examine multiple hypotheses, and, hence, there are multiple significance tests ... reduces (and may eliminate) the possibility of publication bias based on statistical significance, because the probability that all such tests would be significant is quite low" (Hunter & Schmidt, 2004, p. 497). Indeed, there are 31 studies within the current meta-analysis that examined more than one hypothesis, and which yielded non-significant or negative results for at least one research question.

6.4.3. Impact of digital tool use on student attitudes

The low number of studies that investigated student attitudes did not allow for moderator analyses to find out more about beneficial learning conditions with regard to attitudes. For the current meta-analysis, only studies that investigated student attitudes in addition to student performance were considered. The promising explorative findings on student attitudes reported in this meta-analysis, however, call for future research on favorable aspects of using digital tools with regard to student attitudes.

6.4.4. Heterogeneity among analyzed primary studies

The heterogeneity among studies could not be completely explained by the results of the moderator analyses. Although the analyses revealed several interesting tendencies regarding more or less favorable conditions of using digital tools, there was only one variable that significantly influenced the overall effect due to differences between content-related categories, which was teacher training. The use of a random-effects model can increase statistical power regarding the overall effect, but can also lead to power loss within moderator analyses, especially if the heterogeneity among studies is large (Jackson & Turner, 2017). Moreover, additional factors that could not be considered within the current study could have been responsible for the differently sized effectiveness between studies. Such potential variables are different types of learners, different kinds of learning content (procedural, conceptual, or declarative knowledge), or the quality of the assessment of student performance. As already stated by Van der Kleij et al. (2015), studies often do not report on the reliability of test scores. The features mentioned could not be examined within the current study because of insufficient information provided in the primary studies. Therefore, the results of the study should be interpreted in consideration of

these limitations. However, this finding is important for future research because it calls for greater consideration of the potential moderator variables that are mentioned.

6.4.5. Characteristics of analyzed primary studies

Taking a closer look at the countries in which the analyzed studies were conducted, we have to notice that some countries are overrepresented, such as Turkey (22 studies) or Taiwan (10 studies), whereas European countries as well as Canada are quite rare, and Latin-American countries are not represented at all. This may be linked to the fact that in some countries this specific type of comparative educational research studies is funded more frequently (see OECD, 2019b) and is more highly valued by officials in charge of education compared to other countries. This could have resulted in a biased data basis and should be taken into account when interpreting the findings of the present study.

We focused on quantitative primary studies with experimental research design. Numerous studies used the same experimental design, because they compared learning outcomes of one group using digital tools to those of a control group taught without digital tools, and mostly within regular school lessons. However, to investigate whether the didactic processing of the material used per se leads to the better learning outcomes of students using digital tools, it is important to involve a second experimental group using exactly the same material as the first experimental group but without computer support (see Reinhold et al., 2020). We recommend considering this aspect in further research studies.

Furthermore, it seems noteworthy that most of the studies considered within this meta-analysis were published in interdisciplinary educational and psychological journals. With the subject-specific focus of our study, we expected to integrate more studies published in subject-specific journals. The inclusion criteria of our meta-analysis—considering a very specific type of research design—might be responsible for that: Subject-specific journals often have a stronger focus on didactic aspects and comparative research designs may be not that common. It should be further examined whether other types of research designs—published in subject-specific journals—bring more insight regarding aspects that could not be considered within this meta-analysis (e.g., different types of learners, see Section 6.4.4).

Thus, the results of this meta-analysis are limited to a certain extract of the whole research area investigating the impact of learning and teaching with digital tools. In order to complete the whole picture, further research syntheses that could also consider qualitative research designs would be highly rewarding.

6.5. Practical implications

According to Hattie (2012), the effectiveness of educational interventions in school settings should be interpreted in relation to the hinge point of $d = 0.40$, which he found to be the average effect size in his second-order meta-analysis. Slavin, Cheung, Holmes, Madden, and Chamberlain (2013) recommended considering interventions with effect sizes of at least $d = 0.20$ as meaningful results for practical implications. Although the overall effect of learning with digital tools in the current meta-analysis is greater than the suggested hinge points of $d = 0.40$ (Hattie, 2012, p. 3) and $d = 0.20$ (Slavin et al. (2013)), the results should not lead to premature decisions in school practice or completely replace other existing teaching methods. In fact, digital tools show the largest positive effects on student learning outcomes if they are used *in addition* to non-digital material. Despite the potential of using digital tools in mathematics and science classes, teachers should always assess additional benefits in regard to the context they want to use it in, and learning content should still take center stage (e.g., Clark, 1994; Kozma, 1994). Nevertheless, the results of the current study can be used as an orientation guide as well as for informational purposes for teachers or developers of learning environments (see Hillmayr, Reinhold, Zierwald, & Reiss, 2017). Politicians as well as other decision-makers should be aware of the fact that different contextual factors within a computer learning environment can have differently sized impact on student learning outcomes. Practitioners and software developers should consider time aspects of using or developing digital learning tools, the level of guidance during computer-supported learning and the respective subject content. Considering the current efforts of an evidence-based practice within the educational field, teachers can use this research review in order to reflect their own work routine and to refine the way and extent of computer usage within teaching and learning.

Because “digital technologies are ever-changing, not always predictable, and can take on many forms” (Hamilton, Rosenberg, & Akcaoglu, 2016, p. 2), continuous teacher training is important. In the same line, the current meta-analysis shows how important specific teacher training is in supporting student learning with digital tools. However, the international average of teacher participation in digital tool use training is 68% for internal school training and 39% for external school training (Fraillon et al., 2014, p. 187). For that reason, acquiring media competency during teacher training at university should be addressed more comprehensively. Furthermore, school principals are urged to offer appropriate training for teachers, and the latter are encouraged to take part proactively.

All in all, regarding the current state of published research, there is a need for future studies, and to examine in greater detail how digital tools can specifically enhance different aspects of learning mathematics and science subjects. Because the use of intelligent tutoring systems or simulation programs was more effective than drill and practice, the added value of using digital tools in these domains might be related to learning complex ideas or abstract concepts rather than solely repeating and strengthening already learned subject content.

7. Conclusion

Despite the heterogeneous research situation on the use of digital tools in educational settings, the current study brings insight into

the potential of technology-supported teaching and learning. After a comprehensive literature research, 92 primary studies that investigated the effects of using digital tools in secondary school science and mathematics classes were analyzed. The resulting overall effect of digital tool use on student learning outcomes and attitudes toward the taught subject was significantly positive. This shows the potential of learning with digital tools, especially because students often struggle to understand mathematical (OECD, 2016b, p. 192) or scientific subject matter in schools (OECD, 2016b, p. 71). Moreover, the meta-analysis indicated the importance of teacher training before using digital tools in class, and, on a descriptive level, that complementary use of digital tools is more beneficial than replacing other instruction methods completely. Although these results are an essential contribution to the current state of research on the effectiveness of using digital tools in mathematics and science subjects, further studies are needed to gain additional insights about more or less beneficial learning conditions. In particular, future studies should directly compare certain features of interactive digital tools, such as different forms of feedback, and therefore greater consideration of the learning material used by control groups is recommended. Because analyses within the current study indicated that the data might have been biased because of missing unpublished studies, there is further need of a meta-analysis that also searches for gray literature, such as dissertations. Apart from that, future studies should be published more often, *regardless* whether they *do* or *do not report* statistically significant results, since this could counteract the potential problem of publication bias.

Overall, the fact that digital learning environments will be used increasingly often in educational settings is undoubted. Consequently, researchers, teacher educators, and politicians need to be constantly informed so that they can provide best conditions for maximizing the potential of learning with digital tools in school.

Funding

This research was supported by grants from the Stiftung Mercator, Essen, Germany.

Declaration of competing interest

None.

CRediT authorship contribution statement

Delia Hillmayr: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing - original draft, Writing - review & editing, Visualization, Project administration. **Lisa Ziernwald:** Conceptualization, Methodology, Investigation, Data curation, Writing - review & editing. **Frank Reinhold:** Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing - review & editing, Visualization. **Sarah I. Hofer:** Writing - review & editing. **Kristina M. Reiss:** Supervision, Funding acquisition, Writing - review & editing.

References*

- *Abanikannda, M. O. (2016). Enhancing effective chemistry learning through hypermedia instructional mode of delivery. *European Journal of Educational Research*, 5(1), 27–34.
- *Adegoke, B. A. (2010). Integrating animations, narratives and textual information for improving Physics learning. *Electronic Journal of Research in Educational Psychology*, 8(2), 725–748.
- *Aktas, S., & Aydin, A. (2016). The effect of the smart board usage in science and technology lessons. *European Journal of Educational Research*, 64, 125–138.
- *Al-Balushi, S. M., Al-Musawi, A. S., Ambusaidi, A. K., & Al-Hajri, F. H. (2017). The effectiveness of interacting with scientific animations in chemistry using mobile devices on grade 12 students' spatial ability and scientific reasoning skills. *Journal of Science Education and Technology*, 26(1), 70–81. <https://doi.org/10.1007/s10956-016-9652-2>.
- *Aliasgari, M., Riahinia, N., & Mojdehavar, F. (2010). Computer-assisted instruction and student attitudes towards learning mathematics. *Education, Business and Society: Contemporary Middle Eastern Issues*, 3(1), 6–14. <https://doi.org/10.1108/17537981011022779>.
- *Ardac, D., & Sezen, A. H. (2002). Effectiveness of computer-based chemistry instruction in enhancing the learning of content and variable control under guided versus unguided conditions. *Journal of Science Education and Technology*, 11(1), 39–48.
- *Ayub, A. F. M., Tarmizi, R. A., Bakar, K. A., & Luan Wong, S. (2012). WxMaxima computer software as an aid to the study of calculus by students with different learning approaches. *Procedia - Social and Behavioral Sciences*, 64, 467–473. <https://doi.org/10.1016/j.sbspro.2012.11.055>.
- *Baki, A., & Güveli, E. (2008). Evaluation of a web based mathematics teaching material on the subject of functions. *Computers & Education*, 51(2), 854–863. <https://doi.org/10.1016/j.compedu.2007.09.003>.
- *Baran, M., & Maskan, A. (2013). Examining the influence of technology and project-supported thinking journey on achievement. *The Turkish Online Journal of Educational Technology*, 12(2), 122–130.
- *Basaran, B., & Gönen, S. (2012). Investigation of the effects of scorm adapted whiteboard movie techniques on students' success in physics. *Asia-Pacific Forum on Science Learning and Teaching*, 13(2), 1–19.
- *Basham, K. L., & Kotrlík, J. W. (2008). The effects of 3-dimensional CADD modeling on the development of the spatial ability of technology education students. *Journal of Technology Education*, 20(1), 32–47. <https://doi.org/10.21061/jte.v20i1.a.3>.
- Bayraktar, S. (2001/2002). A meta-analysis of the effectiveness of computer-assisted instruction in science education. *Journal of Research on Technology in Education*, 34(2), 173–188. <https://doi.org/10.1080/15391523.2001.10782344>.
- *Beal, C., Arroyo, I., Cohen, P., & Woolf, B. (2010). Evaluation of AnimalWatch: An intelligent tutoring system for arithmetic and fractions. *The Journal of Interactive Online Learning*, 9(1), 64–77.
- Begg, C. B., & Mazumdar, M. (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics*, 50(4), 1088–1101. <https://doi.org/10.2307/2533446>.

* References marked with an asterisk indicate studies included in the meta-analysis.

- Belland, B. R., Walker, A. E., Kim, N. J., & Lefler, M. (2017). Synthesizing results from empirical research on computer-based scaffolding in STEM education: A meta-analysis. *Review of Educational Research*, 87(2), 309–344.
- *Bhagat, K. K., & Chang, C.-Y. (2015). Incorporating GeoGebra into geometry learning-A lesson from India. *Eurasia Journal of Mathematics, Science and Technology Education*, 11(1), 77–86. <https://doi.org/10.12973/eurasia.2015.1307a>. Eurasia J. Math. Sci. Tech.
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester U.K.: John Wiley & Sons.
- *Buckley, B., Gobert, J., Kindfield, A. C. H., Horwitz, P., Tinker, R. F., Gerlits, B., ... Willett, J. (2004). Model-based teaching and learning with BioLogica™: What do they learn? How do they learn? How do we know? *Journal of Science Education and Technology*, 13(1), 23–41.
- *Çepni, S., Taş, E., & Köse, S. (2006). The effects of computer-assisted material on students' cognitive levels, misconceptions and attitudes towards science. *Computers & Education*, 46(2), 192–205. <https://doi.org/10.1016/j.compedu.2004.07.008>.
- *Chandra, V., & Briske, J. (2012). ICT driven pedagogies and its impact on learning outcomes in high school mathematics. *International Journal of Pedagogies and Learning*, 39(6), 1471–1498. <https://doi.org/10.5172/ijpl.2012.1471>.
- *Chang, K.-E., Chen, Y.-L., Lin, H.-Y., & Sung, Y.-T. (2008). Effects of learning support in simulation-based physics learning. *Computers & Education*, 51(4), 1486–1498.
- Chang, K. E., Sung, Y. T., & Chen, S. F. (2001). Learning through computer-based concept mapping with scaffolding aid. *Journal of Computer Assisted Learning*, 17(1), 21–33. <https://doi.org/10.1046/j.1365-2729.2001.00156.x>.
- *Cheng, Y.-H., Cheng, J.-T., & Chen, D.-J. (2012). The effect of multimedia computer assisted instruction and learning style on learning achievement. *WSEAS Transactions on Information Science and Applications*, 9(1), 24–35.
- Chen, J., Wang, M., Kirschner, P. A., & Tsai, C. C. (2018). The role of collaboration, computer use, learning environments, and supporting strategies in CSCL: A meta-analysis. *Review of Educational Research*, 88(6), 799–843.
- Cheung, A. C. K., & 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>.
- Clark, R. E. (1994). Media will never influence learning. *Educational Technology Research & Development*, 42(2), 21–29. <https://doi.org/10.1007/BF02299088>.
- Clark, R. E., & Sugrue, B. M. (1990). North American disputes about research on learning from media. *International Journal of Educational Research*, 14(6), 507–520. [https://doi.org/10.1016/0883-0355\(90\)90021-Y](https://doi.org/10.1016/0883-0355(90)90021-Y).
- Cohen, J. (1968). Weighted Kappa: Nominal scale Agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin*, 70(4), 213–220.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hoboken, NJ: Taylor and Francis.
- Cooper, H. M. (2010). Research synthesis and meta-analysis: A step-by-step approach. In *Applied social research methods series* (4th ed., Vol. 2). Los Angeles, CA: Sage Publications.
- Cumming, G. (2012). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. New York, NY: Routledge.
- Deci, E. L., & Ryan, R. M. (2010). *Self-determination. The Corsini encyclopedia of psychology*.
- *Dhindsa, H. S., & Shahrizal-Emran. (2011). Using interactive whiteboard technology-rich constructivist learning environment to minimize gender differences in chemistry achievement. *International Journal of Environmental & Science Education*, 6(4), 393–414.
- Dowker, A., Sarkar, A., & Looi, C. Y. (2016). Mathematics anxiety: What have we learned in 60 years? *Frontiers in Psychology*, 7, 508. <https://doi.org/10.3389/fpsyg.2016.00508>.
- *Drijvers, P., Doorman, M., Kirschner, P., Hoogveld, B., & Boon, P. (2014). The effect of online tasks for algebra on student achievement in grade 8. *Technology, Knowledge and Learning*, 19(1–2), 1–18. <https://doi.org/10.1007/s10758-014-9217-5>.
- Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, 56(2), 455–463.
- *Ekici, F. T., & Pekmezci, S. (2015). Using ICT-supported narratives in teaching science and their effects on middle school students. *Turkish Online Journal of Educational Technology*, 14(4), 173–186.
- Ellis, S. M., & Steyn, H. S. (2003). Practical significance (effect sizes) versus or in combination with statistical significance (p-values). *Management Dynamics*, 12(4), 51–53.
- *Faour, M. A., & Ayoubi, Z. (2018). The effect of using virtual laboratory on grade 10 students' conceptual understanding and their attitudes towards physics. *Journal of Education in Science, Environment and Health*, 4(1), 54–68.
- *Frailich, M., Kesner, M., & Hofstein, A. (2009). Enhancing students' understanding of the concept of chemical bonding by using activities provided on an interactive website. *Journal of Research in Science Teaching*, 46(3), 289–310. <https://doi.org/10.1002/tea.20278>.
- Fraillon, J., Ainley, J., Schulz, W., Friedman, T., & Duckworth, D. (2019). *Preparing for life in a digital world: The IEA international computer and information literacy study 2018 international report*. Melbourne, Australia: Australian Council for Educational Research (ACER).
- Fraillon, J., Ainley, J., Schulz, W., Friedman, T., & Gebhardt, E. (2014). *Preparing for life in a digital age: The IEA international computer and information literacy study international report*. Melbourne, Australia: Australian Council for Educational Research (ACER).
- *Funkhouser, C. (2002). The effects of computer-augmented geometry instruction on student performance and attitudes. *Journal of Research on Technology in Education*, 35(2), 163–175.
- *Gelbart, H., Brill, G., & Yarden, A. (2009). The impact of a web-based research simulation in bioinformatics on students' understanding of genetics. *Research in Science Education*, 39(5), 725–751. <https://doi.org/10.1007/s11165-008-9101-1>.
- *Graff, M., Mayer, P., & Lebens, M. (2008). Evaluating a web based intelligent tutoring system for mathematics at German lower secondary schools. *Education and Information Technologies*, 13(3), 221–230. <https://doi.org/10.1007/s10639-008-9062-z>.
- Greerath, G., Hertleif, C., & Siller, H.-St. (2018). Mathematical modelling with digital tools—a quantitative study on mathematizing with dynamic geometry software. *ZDM Mathematics Education*, 50(1), 233–244. <https://doi.org/10.1007/s11858-018-0924-6>.
- *Gregorius, R. M., Santos, R., Dano, J. B., & Gutierrez, J. J. (2010). Can animations effectively substitute for traditional teaching methods? Part I: Preparation and testing of materials. *Chemistry Education: Research and Practice*, 11(4), 253–261. <https://doi.org/10.1039/C0RP90006K>.
- Gros, B. (2007). Digital games in education: The design of games-based learning environments. *Journal of Research on Technology in Education*, 40(1), 23–38.
- *Gunbas, N. (2015). Students' mathematics word problem-solving achievement in a computer-based story. *Journal of Computer Assisted Learning*, 31(1), 78–95. <https://doi.org/10.1111/jcal.12067>.
- *Gürbüz, R., & Birgin, O. (2012). The effect of computer-assisted teaching on remedying misconceptions: The case of the subject “probability”. *Computers & Education*, 58(3), 931–941. <https://doi.org/10.1016/j.compedu.2011.11.005>.
- Hamilton, E. R., Rosenberg, J. M., & Akaoglu, M. (2016). Examining the substitution augmentation modification redefinition (SAMR) model for technology integration. *TechTrends*, 60, 433–441. <https://doi.org/10.1007/s11528-016-0091-y>.
- *Harskamp, E., & Suhre, C. (2007). Schoenfeld's problem solving theory in a student controlled learning environment. *Computers & Education*, 49(3), 822–839. <https://doi.org/10.1016/j.compedu.2005.11.024>.
- Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning*. London, UK: Routledge.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>.
- Hedges, L. V. (1981). Distribution theory for glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, 6(2), 107–128.
- Hedges, L. V., & Pigott, T. D. (2004). The power of statistical tests for moderators in meta-analysis. *Psychological Methods*, 9(4), 426–445.
- *Hegedus, S. J., Dalton, S., & Tapper, J. R. (2015). The impact of technology-enhanced curriculum on learning advanced algebra in US high school classrooms. *Educational Technology Research & Development*, 63(2), 203–228. <https://doi.org/10.1007/s11423-015-9371-z>.
- *Herga, N. R., & Grmek, M. I. (2014). Virtual laboratory as an element of visualization when teaching chemical contents in science class. *The Turkish Online Journal of Educational Technology*, 13(4), 157–165.
- Higgins, K., Huscroft-D'Angelo, J., & Crawford, L. (2019). Effects of technology in mathematics on achievement, motivation, and attitude: A meta-analysis. *Journal of Educational Computing Research*, 57(2), 283–319.

- Hillmayr, D., Reinhold, F., Zierwald, L., & Reiss, K. (2017). *Digitale Medien im mathematisch-naturwissenschaftlichen Unterricht der Sekundarstufe: Einsatzmöglichkeiten, Umsetzung und Wirksamkeit [Teaching mathematics and science with digital media in secondary schools: usability, implementation, and effectiveness]*. Muenster, Germany: Waxmann.
- Hohenwarter, M., Jarvis, D., & Lavicza, Z. (2009). Linking geometry, algebra and mathematics teachers: GeoGebra software and the establishment of the international GeoGebra institute. *International Journal for Technology in Mathematics Education*, 16(2), 83–87.
- Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings* (2nd ed.). Thousand Oaks, CA: Sage Publications.
- *Huppert, J., Lomask, S. M., & Lazarowitz, R. (2002). Computer simulations in the high school: Students' cognitive stages, science process skills and academic achievement in microbiology. *International Journal of Science Education*, 24(8), 803–821. <https://doi.org/10.1080/09500690110049150>.
- Jackson, D., & Turner, R. (2017). Power analysis for random-effects meta-analysis. *Research Synthesis Methods*, 8(3), 290–302. <https://doi.org/10.1002/jrsm.1240>.
- Jeno, L. M., Vandvik, V., Eliassen, S., & Grytnes, J.-A. (2019). Testing the novelty effect of an m-learning tool on internalization and achievement: A self-determination theory approach. *Computers & Education*, 128, 398–413. <https://doi.org/10.1016/j.compedu.2018.10.008>.
- Jimoyiannis, A., & Komis, V. (2001). Computer simulations in physics teaching and learning: A case study on students' understanding of trajectory motion. *Computers & Education*, 36(2), 183–204.
- *Johnson-Glenberg, M. C., Birchfield, D. A., Tolentino, L., & Koziupa, T. (2014). Collaborative embodied learning in mixed reality motion-capture environments: Two science studies. *Journal of Educational Psychology*, 106(1), 86–104. <https://doi.org/10.1037/a0034008>.
- *Jones, B. D. (2002). Students as web site authors: Effects on motivation and achievement. *Journal of Educational Technology Systems*, 31(4), 441–461.
- *Kaheru, S. J., & Kriek, J. (2016). The effect of computer simulations on acquisition of knowledge and cognitive load: A gender perspective. *African Journal of Research in Mathematics Science and Technology Education*, 20(1), 67–79. <https://doi.org/10.1080/10288457.2016.1150558>.
- *Kara, Y., & Yesilyurt, S. (2008). Comparing the impacts of tutorial and edutainment software programs on students' achievements, misconceptions, and attitudes towards biology. *Journal of Science Education and Technology*, 17(1), 32–41. <https://doi.org/10.1007/s10956-007-9077-z>.
- Karich, A. C., Burns, M. K., & Maki, K. E. (2014). Updated meta-analysis of learner control within educational technology. *Review of Educational Research*, 84(3), 392–410. <https://doi.org/10.3102/0034654314526064>.
- *Koedinger, K. R., McLaughlin, E. A., & Heffernan, N. T. (2010). A quasi-experimental evaluation of an on-line formative assessment and tutoring system. *Journal of Educational Computing Research*, 43(4), 489–510. <https://doi.org/10.2190/EC.43.4.d>.
- *Koklu, O., & Topcu, A. (2012). Effect of Cabri-assisted instruction on secondary school students' misconceptions about graphs of quadratic functions. *International Journal of Mathematical Education in Science & Technology*, 43(8), 999–1011. <https://doi.org/10.1080/0020739X.2012.678892>.
- *Kolcak, D. Y., & Mogol, S. (2014). A comparison of the effects of laboratory method and computer simulations to avoid misconceptions in physics education. *Education in Science*, 39(175), 154–171. <https://doi.org/10.15390/EB.2014.2052>.
- Kozma, R. B. (1994). Will media influence learning? Reframing the debate. *Educational Technology Research & Development*, 42(2), 7–19. <https://doi.org/10.1007/BF02299087>.
- *Kramarski, B., & Hirsch, C. (2003). Effects of computer algebra system (CAS) with metacognitive training on mathematical reasoning. *Educational Media International*, 40(3–4), 249–257. <https://doi.org/10.1080/0952398032000113167>.
- Kulik, C.-L. C., & Kulik, J. A. (1991). Effectiveness of computer-based instruction: An updated analysis. *Computers in Human Behavior*, 7(1–2), 75–94.
- *Kwon, S. Y., & Cifuentes, L. (2007). Using computers to individually-generate vs. Collaboratively-generate concept maps. *Educational Technology & Society*, 10(4), 269–280.
- *Kwon, S. Y., & Cifuentes, L. (2009). The comparative effect of individually-constructed vs. collaboratively-constructed computer-based concept maps. *Computers & Education*, 52(2), 365–375. <https://doi.org/10.1016/j.compedu.2008.09.012>.
- *Lalley, J. P., Piotrowski, P. S., & Battaglia, B. (2010). A comparison of V-Frog© to physical frog dissection. *International Journal of Environmental & Science Education*, 5(2), 189–200.
- *Lamb, R. L., & Annetta, L. (2013). The use of online modules and the effect on student outcomes in a high school chemistry class. *Journal of Science Education and Technology*, 22(5), 603–613. <https://doi.org/10.1007/s10956-012-9417-5>.
- Landis, R. J., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159–174.
- *Leong, K. E. (2013). Impact of geometer's sketchpad on students achievement in graph functions. *Malaysian Online Journal of Educational Technology*, 1(2), 19–33.
- Lesh, R., Post, T., & Behr, M. (1987). Dienes revisited: Multiple embodiments in computer environments. In I. Wirsup, & R. Streit (Eds.), *Development in school mathematics education around the world* (pp. 647–680). Reston, VA: National Council of Teachers of Mathematics.
- Lichti, M., & Roth, J. (2018). How to foster functional thinking in learning environments using computer-based simulations or real materials. *Journal for STEM Education Research*, 1, 148–172. <https://doi.org/10.1007/s41979-018-0007-1>.
- *Lin, Y.-T., Chang, C.-H., Hou, H.-T., & Wu, K.-C. (2016). Exploring the effects of employing google docs in collaborative concept mapping on achievement, concept representation, and attitudes. *Interactive Learning Environments*, 24(7), 1552–1573. <https://doi.org/10.1080/10494820.2015.1041398>.
- *Lin, C.-y., Cheng, Y.-j., & Chang, Y.-t. (2002). The use of internet-based learning in biology. *Innovations in Education & Teaching International*, 39(3), 237–424.
- Lipsey, M. W., Puzio, K., Yun, C., Hebert, M. A., Steinka-Fry, K., Cole, M. W., et al. (2012). *Translating the statistical representation of the effects of education interventions into more readily interpretable forms*. U.S. Department of Education.
- Lipsey, M. W., & Wilson, D. B. (2001). Practical meta-analysis. In *Applied social research methods series* (Vol. 49). Thousand Oaks, CA: Sage Publications.
- *Looi, C.-K., & Lim, K.-S. (2009). From bar diagrams to letter-symbolic algebra: A technology-enabled bridging. *Journal of Computer Assisted Learning*, 25(4), 358–374. <https://doi.org/10.1111/j.1365-2729.2009.00313.x>.
- Lortie-Forgues, H., Tian, J., & Siegler, R. S. (2015). Why is learning fraction and decimal arithmetic so difficult? *Developmental Review*, 38, 201–221. <https://doi.org/10.1016/j.dr.2015.07.008>.
- *Lundeberg, M., Bergland, M., & Klyczek, K. (2003). Windows to the world: Perspectives on case-based multimedia web projects in science. *The Journal of Interactive Online Learning*, 1(3), 1–16.
- Ma, W., Adesope, O. O., Nesbit, J. C., & Liu, Q. (2014). Intelligent tutoring systems and learning outcomes: A meta-analysis. *Journal of Educational Psychology*, 106(4), 901–918. <https://doi.org/10.1037/a0037123>.
- *Marbach-Ad, G., Rotbain, Y., & Stavry, R. (2008). Using computer animation and illustration activities to improve high school students' achievement in molecular genetics. *Journal of Research in Science Teaching*, 45(3), 273–292. <https://doi.org/10.1002/tea.20222>.
- Mayer, R. E. (2014). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 31–48). New York, NY: Cambridge University Press.
- *McLaughlin, J. S., & Munsell, D. S. (2012). Evolving on-line pedagogy: Developing research-based multimedia learning tools for the high school and undergraduate biology "classroom". *International Journal of Online Pedagogy and Course Design*, 2(1), 1–20. <https://doi.org/10.4018/ijopcd.2012010101>.
- *Mihindo, W. J., Wachanga, S. W., & Anditi, Z. O. (2017). Effects of computer-based simulations teaching approach on students' achievement in the learning of chemistry among secondary school students in Nakuru sub county, Kenya. *Journal of Education and Practice*, 8(5), 65–75.
- Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A new framework for teacher knowledge. *Teachers College Record*, 108(6), 1017–1054.
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, 6(7). <https://doi.org/10.1371/journal.pmed.1000097>.
- Moreno, R. (2006). Optimizing learning from animations by minimizing cognitive load: Cognitive and affective consequences of signaling and segmentation methods. *Applied Cognitive Psychology*, 21, 1–17.
- Moreno, R., & Mayer, R. (). Interactive multimodal learning environments: Special issue on interactive learning environments: Contemporary issues and trends. *Educational Psychology Review*, 19(3), 309–326. <https://doi.org/10.1007/s10648-007-9047-2>.
- *Moyer-Packenham, P., & Suh, J. (2012). Learning mathematics with technology: The influence of virtual manipulatives on different achievement groups. *Journal of Computers in Mathematics and Science Teaching*, 31(1), 39–59.

- *Mustafa, M., & Trudel, L. (2013). The impact of cognitive tools on the development of the inquiry skills of high school students in physics. *International Journal of Advanced Computer Science and Applications*, 4(9), 124–129. <https://doi.org/10.14569/IJACSA.2013.040920>.
- Nattland, A., & Kerres, M. (2009). Computerbasierte Methoden im Unterricht [Computer-based Methods in Class]. In K.-H. Arnold, U. Sandfuch, & J. Wiechmann (Eds.), *Handbuch Unterricht* (2nd ed., pp. 317–324). Bad Heilbrunn, Germany: Julius Klinkhardt.
- *Nugraini, S. H., Choo, K. A., Hin, H. S., & Hoon, T. S. (2013). Impact of e-AV biology website for learning about renewable energy. *The Turkish Online Journal of Educational Technology*, 12(2), 376–386.
- Obersteiner, A., Hoof, J. V., Verschaffel, L., & Dooren, W. V. (2016). Who can escape the natural number bias in rational number tasks? A study involving students and experts. *British Journal of Psychology*, 107(3), 537–555. <https://doi.org/10.1111/bjop.12161>.
- OECD. (2015). *Students, computers and learning: Making the connection*. Paris: Organisation for Economic Co-operation and Development (OECD). <https://doi.org/10.1787/9789264239555-en>.
- OECD. (2016). *PISA 2015: Results in focus. PISA in focus, No. 67*. Paris: Organisation for Economic Co-operation and Development (OECD). <https://doi.org/10.1787/aa9237e6-en>.
- OECD. (2016). *PISA 2015 results (volume I): Excellence and equity in education*. Paris: Organisation for Economic Co-operation and Development (OECD). <https://doi.org/10.1787/9789264266490-en>.
- OECD. (2019). *PISA 2018 results (volume I): What students know and can do*. Paris: Organisation for Economic Co-operation and Development (OECD). <https://doi.org/10.1787/5f07c754-en>.
- OECD. (2019). *Gross domestic spending on R&D (indicator)*. <https://doi.org/10.1787/d8b068b4-en>. (Accessed 11 December 2019).
- *Onder, R., & Aydin, H. (2016). The effect of the use of smart board in the biology class on the academic achievement of student. *Journal on School Educational Technology*, 12(1), 18–29.
- *Osman, K., & Lee, T. T. (2013). Impact of interactive multimedia module with pedagogical agents on students' understanding and motivation in the learning of electrochemistry. *International Journal of Science and Mathematics Education*, 12, 395–421.
- *Owusu, K. A., Monney, K. A., Appiah, J. Y., & Wilmot, E. M. (2010). Effects of computer-assisted instruction on performance of senior high school biology students in Ghana. *Computers & Education*, 55(2), 904–910. <https://doi.org/10.1016/j.compedu.2010.04.001>.
- Ozdemir, M., Sahin, C., Arcagok, S., & Demir, M. K. (2018). The effect of augmented reality applications in the learning process: A meta-analysis study. *European Journal of Educational Research*, 74, 165–186.
- *Özmen, H. (2008). The influence of computer-assisted instruction on students' conceptual understanding of chemical bonding and attitude toward chemistry: A case for Turkey. *Computers & Education*, 51(1), 423–438. <https://doi.org/10.1016/j.compedu.2007.06.002>.
- *Özmen, H., Demircioğlu, H., & Demircioğlu, G. (2009). The effects of conceptual change texts accompanied with animations on overcoming 11th grade students' alternative conceptions of chemical bonding. *Computers & Education*, 52(3), 681–695. <https://doi.org/10.1016/j.compedu.2008.11.017>.
- *Özyurt, Ö., Özyurt, H., Güven, B., & Baki, A. (2014). The effects of UZWEBMAT on the probability unit achievement of Turkish eleventh grade students and the reasons for such effects. *Computers & Education*, 75, 1–18. <https://doi.org/10.1016/j.compedu.2014.02.005>.
- *, Perry, D. R., & Steck, A. K. (2015). Increasing student engagement, self-efficacy, and meta-cognitive self-regulation in the high school geometry classroom: Do iPads help? *Computers in the Schools*, 32(2), 122–143. <https://doi.org/10.1080/07380569.2015.1036650>.
- *Poehnl, S., & Bogner, F. X. (2013). Cognitive load and alternative conceptions in learning genetics: Effects from provoking confusion. *The Journal of Educational Research*, 106(3), 183–196. <https://doi.org/10.1080/00220671.2012.687790>.
- *Pol, H., Harskamp, E., & Suhre, C. (2005). Solving physics problems with the help of computer-assisted instruction. *International Journal of Science Education*, 27(4), 451–469. <https://doi.org/10.1080/0950069042000266164>.
- *Puhke, M., Perse, M., & Sorgo, A. (2012). Comparison between a real field trip and a virtual field trip in a nature preserve: Knowledge gained in biology and ecology. *Journal of Baltic Science Education*, 11(2), 164–174.
- *Rahman, J. A., Ismail, A. H., & Nasir, M. (2014). Development and evaluation of the effectiveness of computer-assisted physics instruction. *International Education Studies*, 7(13), 14–21. <https://doi.org/10.5539/ies.v7n13p14>.
- Reinhold, F., Hoch, S., Werner, B., Richter-Gebert, J., & Reiss, K. (2020). Learning Fractions with and without Educational Technology: What Matters for High-Achieving and Low-Achieving Students? *Learning and Instruction*, 65, 101264. <https://doi.org/10.1016/j.learninstruc.2019.101264>.
- *Roschelle, J., Shechtman, N., Tatar, D., Hegedus, S., Hopkins, B., Empson, S., ... Gallagher, L. P. (2010). Integration of technology, curriculum, and professional development for advancing middle school mathematics. *American Educational Research Journal*, 47(4), 833–878. <https://doi.org/10.3102/0002831210367426>.
- Rosenthal, R. (1979). The "file drawer problem" and tolerance for null results. *Psychological Bulletin*, 86(3), 638–641. <https://doi.org/10.1037/0033-2909.86.3.638>.
- Rosnow, R. L., & Rosenthal, R. (2003). Effect sizes for experimenting psychologists. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 57(3), 221–237. <https://doi.org/10.1037/h0087427>.
- Sailer, M., & Homner, L. (2019). The gamification of learning: A meta-analysis. *Educational Psychology Review*, 32(1), 77–112. <https://doi.org/10.1007/s10648-019-09498-w>.
- Schiepe-Tiska, A., Roczen, N., Müller, K., Prenzel, M., & Osborne, J. (2017). Science-related outcomes: Attitudes, motivation, value beliefs, strategies. In S. Kuger, E. Klieme, N. Jude, & D. Kaplan (Eds.), *Methodology of educational measurement and assessment. Assessing contexts of learning: An international perspective* (pp. 301–329). Cham, Switzerland: Springer.
- *Sengel, E., & Özden, Y. (2010). The effects of computer simulated experiments on high school students' understanding of the displacement and velocity concepts. *European Journal of Educational Research*, 39, 191–211.
- *Shadaan, P., & Leong, K. E. (2013). Effectiveness of using geogebra on students' understanding in learning circles. *Malaysian Online Journal of Educational Technology*, 1(4), 1–11.
- Shadish, W. R., & Haddock, K. C. (2009). Combining estimates of effect size. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (pp. 257–277). New York, NY: Russell Sage Foundation.
- *She, H.-C., & Lee, C.-Q. (2008). SCCR digital learning system for scientific conceptual change and scientific reasoning. *Computers & Education*, 51(2), 724–742. <https://doi.org/10.1016/j.compedu.2007.07.009>.
- *Shieh, R. S. (2012). The impact of Technology-Enabled Active Learning (TEAL) implementation on student performance and teachers' teaching in a high school context. *Computers & Education*, 59(3), 206–214. <https://doi.org/10.1016/j.compedu.2012.03.024>.
- *Shim, K.-C., Park, J.-S., Kim, H.-S., & Kim, J.-H. (2003). Application of virtual reality technology in biology education. *Journal of Biological Education*, 37(2), 71–74.
- *Shirvani, H. (2010). The effects of using computer technology with lower-performing students: Technology and student mathematics achievement. *The International Journal of Learning*, 17(1), 144–154.
- Sink, C. A., & Stroh, H. R. (2006). Practical significance: the use of effect sizes in school counseling research. *Professional School Counseling*, 9(5), 401–411.
- Slavin, R. E., Cheung, A., Holmes, G., Madden, N. A., & Chamberlain, A. (2013). Effects of a data-driven district reform model on state assessment outcomes. *American Educational Research Journal*, 50(2), 371–396. <https://doi.org/10.3102/0002831212466909>.
- Slavin, R., & Smith, D. (2009). The relationship between sample sizes and effect sizes in systematic reviews in education. *Educational Evaluation and Policy Analysis*, 31(4), 500–506.
- *Soliman, M. M., & Hilal, A. J. (2016). Investigating the effects of computer-assisted instruction on achievement and attitudes towards mathematics among seventh-grade students in Kuwait. *International Journal for Technology in Mathematics Education*, 23(4), 145–159.
- *Starbek, P., Starčič Erjavec, M., & Pečklaj, C. (2010). Teaching genetics with multimedia results in better acquisition of knowledge and improvement in comprehension. *Journal of Computer Assisted Learning*, 26(3), 214–224. <https://doi.org/10.1111/j.1365-2729.2009.00344.x>.
- Steenbergen-Hu, S., & Cooper, H. (2013). A meta-analysis of the effectiveness of intelligent tutoring systems on K–12 students' mathematical learning. *Journal of Educational Psychology*, 105(4), 970–987.
- Steenbergen-Hu, S., & Cooper, H. (2014). A meta-analysis of the effectiveness of intelligent tutoring systems on college students' academic learning. *Journal of Educational Psychology*, 106(2), 331–347. <https://doi.org/10.1037/a0034752>.

- *Stieff, M. (2011). Improving representational competence using molecular simulations embedded in inquiry activities. *Journal of Research in Science Teaching*, 48(10), 1137–1158. <https://doi.org/10.1002/tea.20438>.
- Sung, Y. T., Chang, K. E., & Liu, T. C. (2016). The effects of integrating mobile devices with teaching and learning on students' learning performance: A meta-analysis and research synthesis. *Computers & Education*, 94, 252–275.
- Sung, Y.-T., Yang, J.-M., & Lee, H.-Y. (2017). The effects of mobile-computer-supported collaborative learning: Meta-analysis and critical synthesis. *Review of Educational Research*, 87(4), 768–805.
- *Taasobshirazi, G., Zuiker, S. J., Anderson, K. T., & Hickey, D. T. (2006). Enhancing inquiry, understanding, and achievement in an astronomy multimedia learning environment. *Journal of Science Education and Technology*, 15(5–6), 383–395. <https://doi.org/10.1007/s10956-006-9028-0>.
- *Tatli, Z., & Ayas, A. (2013). Effect of a virtual chemistry laboratory on students' achievement. *Educational Technology & Society*, 16(1), 159–170.
- *Tienken, C., & Wilson, M. (2007). The impact of computer assisted instruction on seventh-grade students' mathematics achievement. *Planning and Changing*, 38(3–4), 181–190.
- *Tunaboylu, C., & Demir, E. (2017). The effect of teaching supported by interactive whiteboard on students' mathematical achievements in lower secondary education. *Journal of Education and Learning*, 6(1), 81–94.
- *Turk, H. S., & Akyuz, D. (2016). The effects of using dynamic geometry on eighth grade students' achievement and attitude towards triangles. *International Journal for Technology in Mathematics Education*, 23(3), 95–102.
- *Ülen, S., Cagran, B., Slavinec, M., & Gerlič, I. (2014). Designing and evaluating the effectiveness of physlet-based learning materials in supporting conceptual learning in secondary school physics. *Journal of Science Education and Technology*, 23(5), 658–667. <https://doi.org/10.1007/s10956-014-9492-x>.
- Van der Kleij, F. M., Feskens, R. C., & Eggen, T. J. (2015). Effects of feedback in a computer-based learning environment on students' learning outcomes: A meta-analysis. *Review of Educational Research*, 85(4), 475–511.
- **Vrtačnik, M., Sajovec, M., Dolničar, D., Pučko-Razdevsek, C., Glazar, A., & Brouwer, N. Z. (2000). An interactive multimedia tutorial teaching unit and its effects on student perception and understanding of chemical concepts. *Westminster Studies in Education*, 23(1), 91–105. <https://doi.org/10.1080/0140672000230109>.
- *Wang, T.-H. (2011). Implementation of Web-based dynamic assessment in facilitating junior high school students to learn mathematics. *Computers & Education*, 56(4), 1062–1071. <https://doi.org/10.1016/j.compedu.2010.09.014>.
- Wouters, P., van Nimwegen, C., van Oostendorp, H., & van der Spek, E. D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology*, 105(2), 249–265. <https://doi.org/10.1037/a0031311>.
- *Yang, K.-Y., & Heh, J.-S. (2007). The impact of internet virtual physics laboratory instruction on the achievement in physics, science process skills and computer attitudes of 10th-grade students. *Journal of Science Education and Technology*, 16(5), 451–461. <https://doi.org/10.1007/s10956-007-9062-6>.
- *Yenmez, A. A., Özpınar, I., & Gökçe, S. (2017). Use of WebQuests in mathematics instruction: Academic achievement, teacher and student opinions. *Universal Journal of Educational Research*, 5(9), 1554–1570.
- **Yildirim, Z., Ozden, Y., & Aksu, M. (2001). Comparison of hypermedia learning and traditional instruction on knowledge acquisition and retention. *The Journal of Educational Research*, 94(4), 207–214.
- *Yuan, Y., Lee, C.-Y., & Wang, C.-H. (2010). A comparison study of polyominoes explorations in a physical and virtual manipulative environment. *Journal of Computer Assisted Learning*, 26(4), 307–316. <https://doi.org/10.1111/j.1365-2729.2010.00352.x>.
- *Zulnaldi, H., & Zakaria, E. (2012). The effect of using GeoGebra on conceptual and procedural knowledge of high school mathematics students. *Asian Social Science*, 8(11), 102–106. <https://doi.org/10.5539/ass.v8n11p102>.