

RESEARCH

Open Access



Effectiveness of digital educational game and game design in STEM learning: a meta-analytic review

Yang Gui^{1,2}, Zhihui Cai^{1,2*}, Yajiao Yang^{1,2}, Lingyuan Kong^{1,2}, Xitao Fan³ and Robert H. Tai⁴

Abstract

Digital educational games exhibit substantial promise in advancing STEM education. Nevertheless, the empirical evidence on both the efficacy of digital game-based learning and its designs in STEM education is characterized by notable inconsistencies. Therefore, the current study aimed to investigate (1) the general effect of digital game-based STEM learning over STEM learning without digital game, and (2) the enhancement effect of added game-design elements against base game versions in STEM learning. Two meta-analyses were conducted in this study. Based on the 136 effect sizes extracted from 86 studies, the first meta-analysis revealed a medium to large general effect of digital game-based STEM learning over conventional STEM learning ($g = 0.624$, 95% CI [0.457, 0.790]). In addition, digital game-based STEM learning appeared to be differentially effective for different learning outcome, different types of game, and different subject. A total of 44 primary studies and 81 effect sizes were identified in the second meta-analysis. The results revealed a small to medium enhancement effect of added game-design elements over base game versions ($g = 0.301$, 95% CI [0.163, 0.438]). Furthermore, our results indicated that the game-design elements added for content learning were more effective than those added for gaming experience. Possible explanations for these findings, as well as the limitations and directions for future research were discussed.

Keywords Digital game-based learning, Game design, STEM education, Academic achievement, Meta-analysis

Introduction

Science, technology, engineering, and mathematics (STEM) education has become an increasingly important education issue around the world and has received extensive attention from educators and other stakeholders (Kayan-Fadlelmula et al., 2022; The White House, 2018).

STEM education aims to train new talents with twenty-first century skills such as computational, critical, and creative thinking (Li et al., 2016; Wahono et al., 2020). In addition, STEM education plays a unique role in addressing real-world issues such as energy, the environment and health (Martín-Páez et al., 2019; Struyf et al., 2019). Therefore, many countries regard STEM education as a national strategy to lead the reform and development of basic education (Dou, 2019). However, currently, STEM education faces some issues. On one hand, traditional STEM classroom can hardly attract students' interest (Gao et al., 2020). On the other hand, the traditional classroom environment is difficult to provide effective practical activities to meet the educational needs of cultivating students' complex problem-solving ability due to limited time and resources (Klopfer & Thompson,

*Correspondence:

Zhihui Cai
caizhihui922@ccnu.edu.cn; caizhihui922@163.com

¹ Key Laboratory of Adolescent Cyberpsychology and Behavior (CCNU), Ministry of Education, Wuhan 430079, China

² School of Psychology, Central China Normal University, Wuhan, Hubei, China

³ Faculty of Education, The Chinese University of Hong Kong, Hong Kong, China

⁴ School of Education and Human Development, University of Virginia, Charlottesville, USA

2020). Therefore, new instructional methods are urgently needed to improve STEM learning.

Digital game-based learning (DGBL), which was discussed as one of the twenty-first century global pedagogical approaches (Kukulska-Hulme et al., 2021), has unique advantages in enhancing STEM education compared with other pedagogical strategies (Ishak et al., 2021). Digital educational games provide an engaging learning environment that allows learners to interact with game mechanics in a virtual world, which provides the learners with a meaningful gaming experience and also greatly enhances their learning motivation (Ball et al., 2020; Ishak et al., 2021). At the same time, digital educational game can serve as an effective learning environment, providing players with ample opportunities for simulation, real-world questions, and rich instructional support. In such an environment, learners can practice problem solving skills, develop critical thinking and foster STEM literacy (Kayan-Fadlilmula et al., 2022; Klopfer & Thompson, 2020).

Despite the great potential of digital educational games, there is no consensus among researchers on the effectiveness of DGBL in STEM education. On one hand, some researchers showed that, compared to traditional instructions, DGBL could support learners to improve learning motivation, understand STEM concepts and develop practical skills (Halpern et al., 2012; Johnson & Mayer, 2010; Masek et al., 2017; Wu & Anderson, 2015). On the other hand, some researchers believed that DGBL may not have significant advantages over traditional methods for STEM learning (Renken & Nunez, 2013; Riopel et al., 2019). Moreover, poorly designed educational games may produce extra cognitive load, which may make the learning worse compared with traditional instructional approaches (Wang, 2020). Therefore, several synthesis studies have pooled the findings of previous studies involving DGBL in STEM disciplines. For example, previous reviews have examined the effects of game-based science learning (Riopel et al., 2019; Tsai & Tsai, 2020), examined the effectiveness of game-based math learning (Byun & Joung, 2018; Tokac et al., 2019), or examined the effects of digital game-based STEM education on student knowledge gains (Wang et al., 2022). These synthesis studies, however, are limited to a single subject or single learning outcome of STEM fields, without providing an overall understanding about the effects of DGBL in STEM education.

In addition to the general question about the overall effectiveness of game-based STEM learning (e.g., comparison between game-based STEM learning vs. traditional STEM learning) as discussed above, some researchers also called for attention to the question of whether some additional gaming and learning mechanics

added in an educational game would enhance the effectiveness of game-based STEM learning (Proulx et al., 2017; Tsai & Tsai, 2020). To examine this question, a study design would typically involve two conditions: one group using the basic (or base) version of a digital learning game, while the other group using an enhanced version of the same digital learning game with added game-design element(s). Thus, the current study, through systematic reviews and meta-analyses of the relevant empirical studies, would attempt to answer two questions: (1) What is the overall effect of digital game-based STEM learning compared to traditional STEM learning? (2) What is the enhancement effect of added game-design element over a base game version in game-based STEM learning?

Literature review

Effectiveness of digital game-based learning in STEM education

Mayer (2014) defines educational games as digital games designed to promote students' academic performance. This kind of digital game contains rich learning and gaming mechanisms, providing learners with an engaging and positive learning environment (Lameras et al., 2017; Maheu-Cadotte et al., 2018). For STEM education, digital educational games are viewed as providing interactive and interesting learning environments that help to develop students' knowledge and skills related to STEM. Why can digital educational games enhance STEM learning? The situated learning theory proposed by Lave and Wenger (1991) provides a possible explanation. From the perspective of situated learning, digital educational games can serve as a learning environment for constructing new knowledge, in which learners can learn and practice skills through interactions within the game and with other players. This game-based STEM learning environment can be viewed as a virtual community of practice that provides learners with learning context, guidelines supporting exploration, and opportunities to collaborate (Klopfer & Thompson, 2020). At the same time, digital educational games contain rich learning mechanisms or elements (e.g., pedagogical agents, self-explanation strategies, and adaptation) that act like cognitive apprenticeships to support learners from novice to expert. For example, Grivokostopoulou et al. (2020) proposed a learning approach in simulation-based game, using embodied pedagogical agents to guide students to explore in virtual communities and to achieve learning goals.

As the field of DGBL evolves, Mayer (2015) advocated an evidence-based approach and proposed a new research framework, calling on researchers to conduct research in three areas: (1) value-added research, (2)

cognitive consequence research, and (3) media comparison research. The value-added research compares the learning outcomes of using a games with added game-design elements against the use of base game without the added design elements. Cognitive consequences research compares cognitive skill gains between groups who play off-the-shelf games for extended time periods and those who engage in control activities. Media comparison research examines whether DGBL promotes more learning when compared to conventional instructional approach.

Many “media comparison research” studies were conducted to explore the effect of digital educational games in STEM learning relative to some conventional learning approaches. However, there is disagreement about the effectiveness of STEM digital educational games. On one hand, many researchers affirm the potential of digital games for STEM learning. The positive effects of game-based STEM learning were shown in many empirical studies (e.g., Kao et al., 2017; Khamparia & Pandey, 2018; Soflano et al., 2015). For example, Kao et al. (2017) showed that game-based learning significantly enhanced scientific problem-solving performance and creativity compared to traditional STEM learning. On the other hand, insignificant or negative effects were also common among empirical studies (e.g., Beserra et al., 2014; Freeman & Higgins, 2016; Sadler et al., 2015) of game-based learning in STEM education. In addition, from the cognitive load perspective, poorly designed games or unnecessary game mechanics may increase extraneous cognitive load (i.e., cognitive load that are not related to learning), resulting in less efficient learning. For example, Schrader and Bastiaens (2012) compared the learning effect in immersive DGBL vs. low-immersion hypertext environments, and showed that DGBL resulted in a higher cognitive load, which reduced the retention and transfer of physical knowledge. In summary, these conflicting findings suggest that the effectiveness of game-based STEM learning is not yet certain, and there could be underlying factors and variables that could influence the effectiveness of STEM digital educational games.

In addition to “media comparison research”, more researchers are now turning their attention to the effectiveness of game design in game-based STEM education, and this direction is aligned with “value-added research” as described in Mayer (2015). As discussed in Clark et al. (2016), the positive effects of digital educational games on learning could largely depend on game design. Plass et al. (2015) described the elements of educational game in an integrated design framework. Game design elements consist of six parts: game mechanics, visual aesthetics, narrative, incentives, musical score, and knowledge/skills, with game mechanics and knowledge/

skills being the most crucial. Games were designed and used to facilitate learners’ cognitive processing in the process of learning the content (knowledge/skills) (Mayer, 2014). As a result, game designers need to consider how the game content should be presented and how the learning mechanism should be designed to facilitate learners’ cognitive development.

Adams and Clark (2014) compared the differences between basic games and the games with added self-explanatory mechanisms. It was found that middle school students had higher extraneous cognitive load and worse physical test scores when playing games with added self-explaining mechanics. Based on these, the current study also intended to synthesize “value-added research” studies that compared different game mechanic design conditions.

Previous meta-analysis studies on DGBL in STEM education

Up to now, there have been several meta-analyses that attempted to integrate existing research on DGBL related to STEM disciplines, and the results from these meta-analytic reviews were inconsistent. Byun and Joung (2018) investigated the effect of digital educational games on K-12 students’ mathematics performance. This meta-analysis including 17 empirical studies found that game-based math learning produced a small to medium effect size ($d=0.37$). In addition, Tokac et al. (2019) also examined the effect of math video games on the academic performance of PreK-12 students (pre-kindergarten through 12th grade). The meta-analysis included 24 studies and showed that DGBL had a very small effect size on math learning ($d=0.13$). Riopel et al. (2019) focuses on the effect of digital educational games on students’ scientific knowledge compared with traditional instructions. The meta-analysis pooled 79 studies and found that digital educational games produced small to medium effect sizes on science learning ($d=0.31-0.41$). The meta-analysis conducted by Tsai and Tsai (2020) included 26 empirical studies investigating the effectiveness of gameplay design and game mechanics’ design for enhancing science learning. This meta-analysis compiled the findings from empirical studies that involved the comparison between using base game version and using the same game with added gameplay design feature or added game-mechanic design feature. The results reported a medium effect size for added gameplay design feature ($k=14, g=0.646$) over the base version of the game, and a small effect size for added game-mechanism design ($k=12, g=0.270$) over the base version of the game.

The meta-analysis studies described above generally indicated that digital educational games had a small to medium effect on science and math learning. Furthermore, these meta-analyses reviewed the effects of DGBL

on learning in science and in mathematics separately. However, as Kelley and Knowles (2016) discussed, STEM education should be treated as a more integrated education model that uses multidisciplinary (science, technology, engineering and math) thinking and knowledge to solve real-world problems. Therefore, it is necessary to examine STEM education as a whole to develop better understanding about the application of digital educational games in STEM education. Moreover, the generalizability and robustness of the findings is also an issue due to the small numbers of studies included in these meta-analyses.

Very recently at the completion of this manuscript, Yu et al. (2022) conducted a systematic review aimed to exploring the application of augmented reality (AR) in STEM education through digital game-based learning (DGBL). The review included a total of 46 articles published between 2010 and 2020, providing valuable insights into the potential use of AR games in STEM education. However, it is important to note that this article only provides a qualitative overview of AR in STEM education and does not address the effectiveness of AR learning games compared to other instructional approaches, nor does it explore other types of games, such as virtual reality and tablet games, in STEM education. Arzmann et al. (2023) also conducted a meta-analysis that examined the effectiveness of game-based learning compared to traditional classroom teaching in STEM education. The meta-analysis revealed a moderate positive effect of game-based learning on cognitive ($g=0.67$), motivational ($g=0.51$), and behavioral outcomes ($g=0.93$) relative to conventional teaching methodologies. However, the definition of game-based learning used in this meta-analysis is too broad, including traditional games such as board games and card games played in non-digital environments, thereby precluding specific inquiry into the effectiveness of DGBL. In addition, the meta-analysis exclusively considered the impact of student characteristics on the effectiveness of game-based learning without due consideration of other pertinent factors, including game features. Therefore, it is imperative to undertake more comprehensive meta-analysis to ascertain the comparative effectiveness of DGBL vs. traditional teaching in STEM education, encompassing all relevant influencing factors. Such investigations will facilitate evidence-based practices and promote optimal educational outcomes for students.

In addition, Wang et al. (2022) also published a new similar meta-analysis study about the effects of digital game-based STEM education on student learning gains. This meta-analysis, which included 33 empirical studies involving STEM subjects from 2010 to 2020, found that digital educational games showed a moderate overall

effect size ($d=0.677$) compared to traditional instructing approaches. This new review study provided very useful information as a summary of empirical studies on digital game-based STEM education, but this study also left a couple of important questions unanswered, as discussed below.

First, this meta-analysis only examined the effectiveness of game-based STEM education relative to other educational approaches; in other words, this meta-analysis only synthesized the primary studies of “media comparison research” (Mayer, 2015). However, as the research framework by Mayer (2015) indicated, research on DGBL should not only focus on the general effect of DGBL over other STEM education approaches (i.e., media comparison research studies), but should also examine the effects of different game-design elements in enhancing STEM learning (i.e., value-added research). The large number of empirical studies for value-added research were absent in this new meta-analysis, thus leaving an obvious vacuum in our understanding about the current situation of DGBL.

Second, compared to previous similar review studies, this new meta-analysis focusing on STEM disciplines included somewhat limited number of empirical studies in STEM sub-areas, thus may not fully reflect the current status of research on game-based STEM education. For example, in terms of mathematics learning, the number ($k=24$) of empirical studies included in a previous meta-analysis (Tokac et al., 2019) is twice as much as that in this new meta-analysis ($k=12$). In terms of science learning, the number ($k=79$) of empirical studies included in a previous meta-analysis (Riopel et al., 2019) is much larger than that included in this new meta-analysis ($k=18$).

Third, methodologically, most of the included studies reported more than one effect size in Wang et al. (2022)'s meta-analysis, and the multiple effect sizes were statistically non-independent. However, the traditional meta-analytic methods used in this new meta-analysis ignored the dependence of effect sizes within a study, which could have led to potentially biased estimates (i.e., inflated Type I errors and smaller confidence intervals of estimates) (Becker, 2000).

In summary, previous research, including several meta-analysis studies, has provided good evidence that DGBL could promote STEM learning when compared against conventional STEM education approaches. However, our understanding about the effect of DGBL in STEM education is still limited. First, previous reviews focused on “media comparison research” studies, but the “value-added research” studies on the role of game design in DGBL in STEM education were not adequately covered or addressed. As indicated by the research framework proposed by Mayer (2015), and as discussed in Clark et.

al. (2016), the positive effects of digital educational games on learning could depend on game design. With the recent focus on the role of game design in game-based STEM learning, Klopfer and Thompson (2020) called for linking game design with learning outcomes and for using learning theory as a guide to examine the effects of game design elements. Some researchers went beyond the “media comparison research,” and paid attention to the role of game design in game-based STEM learning (Clark et al., 2016; Tsai & Tsai, 2018, 2020). However, at this time, we know relatively little about the role of game design in game-based STEM learning, and about what game design factors may affect the effectiveness of game-based STEM learning.

Second, previous reviews focused on knowledge gains as outcomes in game-based STEM education, but the question about the effect of game-based approach on promoting cognitive skills was not addressed, although there was the general belief that STEM games have good potential in promoting twenty-first century skills (Kayan-Fadlelmula et al., 2022; Klopfer & Thompson, 2020). With these important issues relevant to game-based STEM education, there is the great need to expand and deepen our understanding of the impact of digital educational games and their design elements on STEM learning.

Potential moderators

Through the review of previous empirical studies and relevant meta-analyses, we identified some game characteristics and study features that may have contributed to the inconsistent findings across the studies investigating the overall effect and also enhancement effect of game-based STEM learning. These were described below.

Game-related characteristics

Game type: Researchers examined the effect of game type on the effectiveness of game-based learning (Mao et al., 2022; Tsai & Tsai, 2018). Ke (2015) divided games into seven categories: Role-playing, Strategy, Simulation, Construction, Adventure, Action, and Puzzle game (see

Table 1). Previous studies, however, have not reached consensus on the effects of educational game types on learning. Romero et al. (2015) argued that strategy games could foster critical thinking better than role-playing and competitive games. In contrast, Mao et al. (2022) found that role-playing games were the best for developing thinking skills.

Game time: Game time has been the focus of previous studies (e.g., Riopel et al., 2019; Tsai & Tsai, 2018; Wouters et al., 2013). There were two different views on whether the effect of DGBL depended on the length of the game time. One common-sense view was that, multiple sessions in digital educational games would provide students with more opportunities to learn for achieving higher level learning than a single intervention (Clark et al., 2016). For example, Wouters et al. (2013) found that multiple interventions in digital educational games were more effective than a single intervention. A different view, however, pointed to the novelty effect in new technology-based learning (e.g., an educational game). Learners were temporarily interested in learning due to their curiosity about new technologies, but this interest would weaken with time (Poppenk et al., 2010). For example, Riopel et al. (2019) found that the longer the game time, the worse the science learning gains. Therefore, it should be worthwhile to examine the potential effect of game time on DGBL.

Level of realism: The level of realism in a game is regarded as one of the important game attributes. Wilson et al. (2009) defined the level of realism as the similarity between the game scene and the real environment. Based on previous review studies (Clark et al., 2016; Riopel et al., 2019), game realism could be divided into three levels: schematic, cartoon-like, and photorealistic. In schematic games, the game environment is mainly composed of lines, geometry, or text (e.g., Hodges et al., 2018). The cartoon-like game has cartoon characters, and the game environment is mainly 2D (e.g., Vanbecelaere et al., 2020). Photorealistic game has game character similar to human, and the game environment is close to the real

Table 1 Specific description of the game type

Game type	Description	Example
Role-playing	Playing one or more characters and moves in the virtual world	Giannakos (2013)
Strategy	Making strategic deployment through the application of game rules and plan	Roodt and Ryklief (2019)
Simulation	Training in a realistic game environment	Raman et al. (2014)
Construction	Accessing, using, and managing resources in an open game environment	Sung et al. (2018)
Adventure	Focusing on episodic and exploratory interactions such as exploring the unknown and solving puzzles	Fiorella et al. (2018)
Action	Making quick movements and quick reactions in the game	Weng et al. (2015)
Puzzle	Solving puzzles through logical thinking	Vanbecelaere et al. (2020)

environment (e.g. C. -Y. Chen et al., 2020). It is not clear how game-based STEM learning achievement is affected by the level of realism of a game. Some empirical studies showed that learners could achieve excellent STEM academic performances in a highly realistic game environment (Fiorella et al., 2018; Jong, 2015; Okutsu et al., 2013). However, some reviews found that less realistic games appeared to enhance game-based science learning more than photorealistic games (Riopel et al., 2019). Therefore, this study will examine the potential moderating effect of the level of realism.

Game mechanism: The game mechanism reflects the complex relationship between learning and play in game-based learning (Mayer, 2015). As discussed by Arnab et al. (2015), the game mechanism consists of the learning mechanism and the gaming mechanism. Learning mechanisms (e.g., concept maps, feedback) refer to game elements related to the learning content, and the addition and absence of which can change the learning process (e.g., Hwang et al., 2013). In contrast, gaming mechanics (e.g., points, leaderboards) are game elements that are not related to the learning content, but instead, are designed to increase entertainment and gaming experience (e.g., Hsu & Wang, 2018). Some previous empirical studies showed that digital educational games with added learning mechanisms could significantly promote STEM knowledge or skills (Hwang et al., 2013; Khamparia & Pandey, 2018; Sung & Hwang, 2013). Some studies also indicated that games with some added gaming mechanics could be beneficial for students learning (Hsiao et al., 2014; Hsu & Wang, 2018; Nelson et al., 2014). Tsai and Tsai (2020), however, reviewed 12 studies on science learning and did not find beneficial effects of added gaming mechanics or learning mechanics. In general, it remains unclear how game mechanism could moderate the effects of digital educational games on STEM learning. Thus, we will examine the potential moderating effects of game mechanisms in this study.

Study features

Educational level: Whether digital educational games promote academic performance for learners at all education levels has been a topic of interest. Previous review studies involving different STEM fields have examined the moderating role of education level, but have not reached a consistent conclusion. Riopel et al. (2019) found that high school students benefited more from game-based science learning than elementary and college students. However, the meta-analysis by Tsai and Tsai (2020) showed that there was no significant difference in science academic performance among students at different education levels. In view of these

inconsistent results, this study will investigate whether there are differences in game-based STEM learning effects among students at different education levels.

Subject: STEM education includes science, technology, engineering and mathematics. The effect of digital educational games may be different among different STEM disciplines. For example, Cheng et al. (2015) argued that digital educational games were particularly suitable for science learning because DGBL provided a simulation-based environment, where students could experience relevant phenomena that could not be experienced in traditional classroom settings. Some meta-analyses also showed that DGBL was more beneficial for science learning than for other fields (Talan et al., 2020; Wouters & van Oostendorp, 2013). Other meta-analytical reviews, however, found that digital educational games appeared to be more beneficial for students to acquire mathematical knowledge and skills (C. -H. Chen et al., 2020; Wouters et al., 2013). Furthermore, some reviews found that digital educational games did not seem to be suitable for engineering education or computer science compared to other subjects (Talan et al., 2020; Wouters et al., 2013). Based on the above findings, the current meta-analysis considers different STEM disciplines as a potential moderator.

Learning outcome: Supporting learners to achieve excellent learning outcomes has always been a core goal of digital educational games. Consistent with the previous studies (e.g., Riopel et al., 2019; Tsai & Tsai, 2020), we focus on the cognitive outcomes of learning. Wouters et al. (2009) divided cognitive outcomes into knowledge and cognitive skills. Knowledge, including declarative knowledge and procedural knowledge, was expressed as learning gain (Hooshyar et al., 2021), retention (Guo & Goh, 2016), or recall (Hsiao et al., 2014). Cognitive skills involved the use of rules to solve problems or make decisions, and were measured in terms of transfer (Aladé et al., 2016) or problem-solving (Sarvehana, 2019). Digital educational games could enhance STEM learning such as STEM concepts and problem-solving skills (NRC, 2011). However, some empirical studies suggest that the positive effects of digital educational games seemed to vary by learning outcomes. For example, Huang et al. (2020) found that digital educational games promoted problem-solving skills more than STEM knowledge gains. Furthermore, previous meta-analyses involving STEM fields have only found positive effects of digital educational games on knowledge acquisition, and the effects on cognitive skills have not been examined (e.g., Riopel et al., 2019; Tsai & Tsai, 2020). Therefore, we consider the learning outcomes as a potential moderator for the effect of DGBL.

Aim of the present study

Digital educational games have great potential to enhance STEM learning. However, there are inconsistent research findings about the effectiveness of game-based STEM learning, and there is also a lack of systematic review to shed light on the inconsistent findings from these empirical studies. As of now, most studies on game-based STEM learning generally fell into two categories: (1) to compare game-based STEM learning with traditional STEM instruction/learning, and (2) to examine the enhancement effect of added game-design elements through comparison with the base game version. Therefore, the current study would focus on these two types of studies by systemically synthesizing the findings of these two types of studies, and synthesizing the moderating effects of some game features and study features (reviewed above) on the results from different studies on game-based STEM learning.

Building on the findings of the previous meta-analysis, the current study would extend and refine our understanding of the effects of digital educational games and design on STEM learning. First, the current study would provide a more exhaustive quantitative review of the empirical research studies related to game-based STEM education, for the purpose of adequately reflecting the current state of affairs in this active research area. Second, the current review would extend STEM learning outcomes from knowledge gains to cognitive skills, which was not adequately addressed in previous reviews. Third, guided by the research framework of Mayer (2015), the current review would move beyond the sole focus on media comparison research studies, and the value-added research studies about the effects of game design factors on STEM

learning would be synthesized. Fourth, the current quantitative review would use a cutting-edge three-level meta-analytic model so as to avoid the analytical pitfalls and biases as a result of the non-independent multiple effect sizes within one study.

Practically, the current study would perform two meta-analyses, one to focus on the effect of game-based STEM learning in comparison with traditional STEM instruction/learning (media comparison research; Mayer, 2015), and the other to focus on the enhancement effect of added game-design elements in STEM learning in comparison with a base game version (value-added research; Mayer, 2015). In each of these two meta-analyses, potential moderator variables would be examined for their roles in contributing to the inconsistent findings across the studies as observed in the research literature, the research framework is visualized in Fig. 1. More specifically, these two meta-analyses would answer the following questions:

1. Meta-analysis 1: Based on the studies that compared game-based STEM learning with traditional STEM learning:
 - RQ1a: What is the general effect of using digital games in STEM education?
 - RQ1b: Are the effects of game-based learning across different studies as analyzed in RQ1a moderated by game features and study features (i.e., moderator variables)
2. Meta-analysis 2: Based on the studies that compared two game-based STEM learning conditions: one involving a base game version, and the other involv-

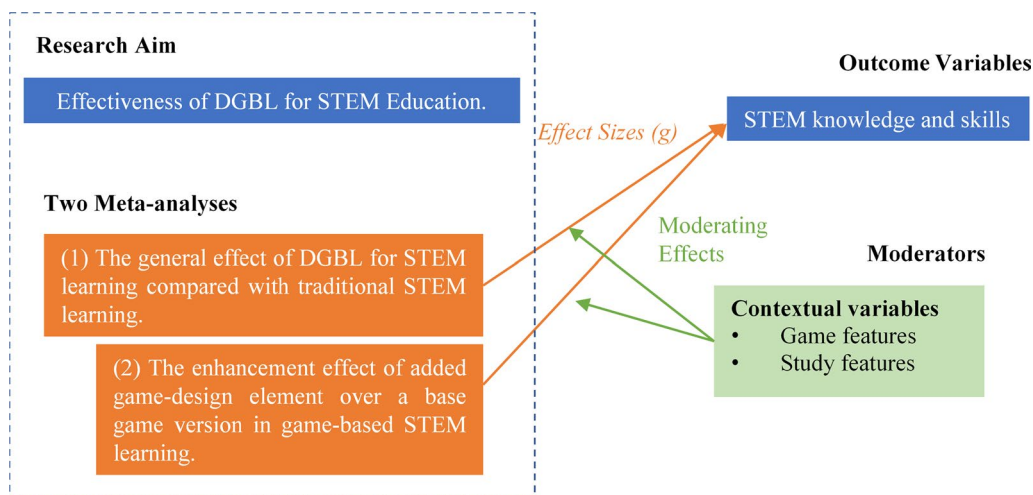


Fig. 1 Research framework

ing the same game but with some added game-design elements:

RQ2a: What is the enhancement effect of added game-design element over a base game version in game-based STEM learning?

RQ2b: Are the enhancement effects of some added game-design elements across different studies moderated by game features and study features (i.e., moderator variables)?

Methods

Literature retrieval and screening

The current meta-analysis strictly followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) criterion to identify relevant empirical research studies (Ziegler et al., 2011). We used three sets of keywords, *games* (“game-based learning” OR “serious game*” OR “educational game*” OR “simulation game*” OR “electronic game*” OR “digital game*” “computer game*” OR “video game*”), AND *STEM* (“science or biology” or “physics” or “chemistry” or “math*” or “technology” or “engineering” or “STEM”), AND *learning outcome* (“academic achievement” OR “academic performance” OR “student performance” OR “learning outcome*” OR “learning effect*” OR “learning performance”) to screen the literature in the databases: *Web*

of Science, *EBSCOhost*, and *ScienceDirect*.. In addition, *Google Scholar*, *ResearchGate*, and *library document delivery* were also used to find unpublished studies. The deadline for the literature search was August 2021. Finally, a total of 2802 articles were initially identified and imported into EndNoteX9 for further examination. After duplicated articles removed, the title and the abstract of each of the remaining 2343 entries were examined by the research team for its relevance and appropriateness for the current meta-analysis. After initial screening based on each article’s title and abstract by checking whether a study was an experimental or quasi-experimental study about the effect of DGBL in STEM education, we narrowed down the list to 467 studies for further full-text reading. The full text screening process was conducted by two authors independently first, and a very limited number of discrepancies between the two authors were later resolved by discussion in the research team. To sift these studies, we developed and used the inclusion/exclusion criteria shown in Table 2. Eventually, 123 studies were identified to have met the inclusion criteria, thus included in one of the two meta-analyses. The flow of screening studies is shown in Fig. 2. The inter-rater reliability coefficient (Krippendorff’s alpha) for the final screening results was greater than 0.95.

Coding of studies

As described under “**Potential moderators**”, several salient characteristics (i.e., game features and study features) were considered potential moderators that could have influenced the findings across the primary studies. These moderator variables were coded into different levels, and the coding details of these moderator variables are presented in Table 3. To ensure the reliability of the coding, three researchers independently coded 25 randomly selected primary studies (20% of all studies). The coding consistency coefficient among the three coders was calculated and the coding results were highly reliable ($\kappa=0.81$). Differences arising in the process of coding were negotiated among the three coders by consulting the original literature. Once the coding procedures established, differences among the coders resolved, and the coding reliability established, then the remaining studies were randomly divided into three groups, with each coder independently coding one group of the studies.

Quality assessment and influential effect sizes detection

We conducted quality assessment for included empirical studies using the Medical Education Study Quality Instrument (MERSQI) (Reed et al., 2007). This 10-item instrument was designed to assess the quality of empirical research in six domains including study design, sampling, type of data, validity of evaluation instrument, data

Table 2 Inclusion/exclusion criteria

Inclusion/exclusion criteria	
1	The learning environment must involve digital educational games. Studies using game elements in non-game environments (e.g., studies on gamification) were excluded
2	Research topics were STEM education, including science, physics, chemistry, biology, mathematics, technology and engineering education
3	The research type should be experimental or quasi-experimental design. Studies without a control/comparison group were excluded
4	The comparison types between the experimental group and the control group were digital educational games and non-educational games (for inclusion in the first meta-analysis), or basic digital educational games and digital educational games with added game mechanics (for inclusion in the second meta-analysis). Other studies that compared game genres and did not isolate the effect of games were excluded
5	There were sufficient statistical data in the research report to allow calculation or estimation of effect size between the comparison groups (e.g., sample size, mean value, standard deviation, <i>T</i> value, <i>F</i> value, etc.)
6	The study participants were students, excluding those with learning disabilities
7	The language of published studies should be English

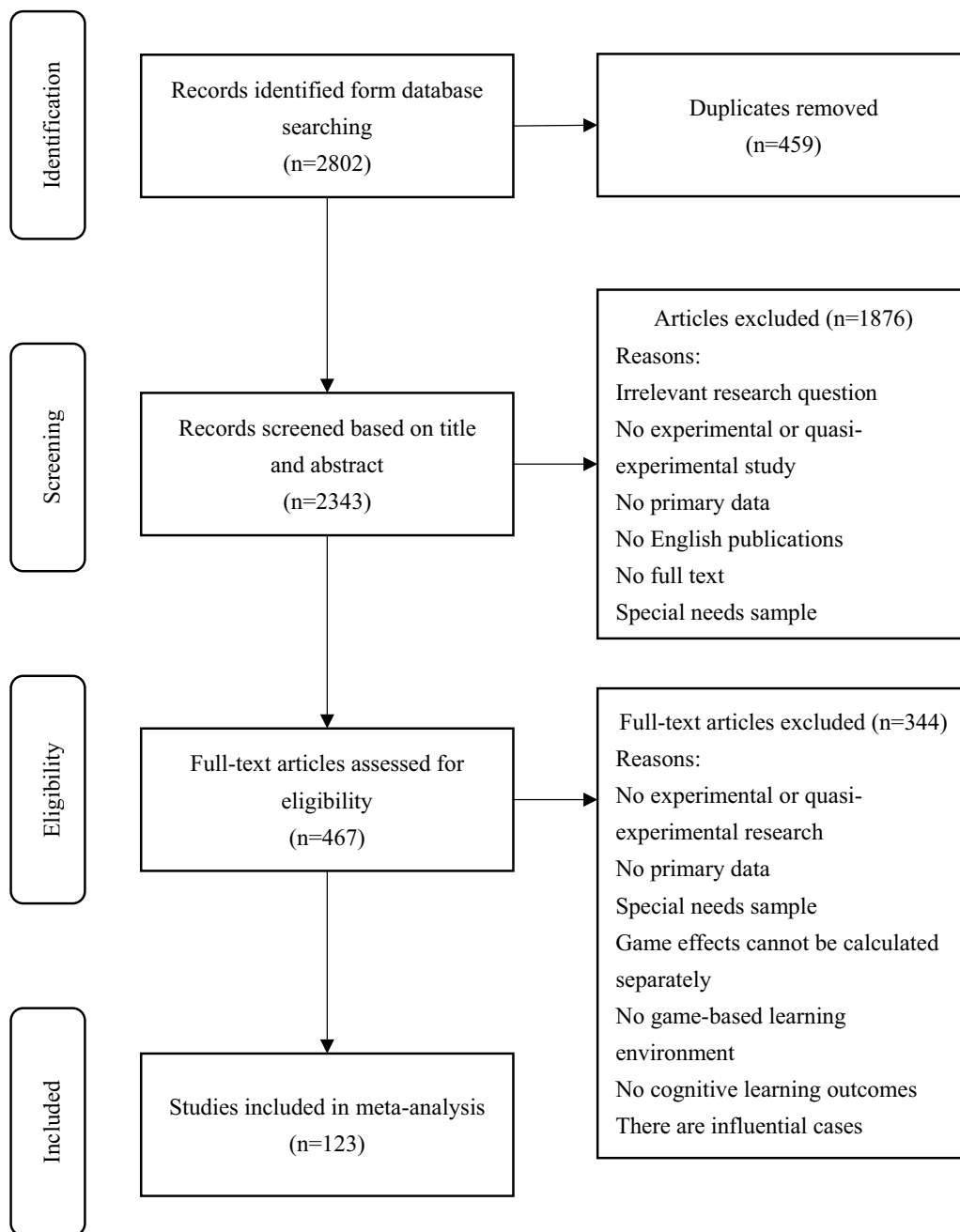


Fig. 2 PRISMA diagram

analysis and outcomes. The full score of MERSQI is 18, with each item scoring from 0 to 3. The quality of empirical studies included in the current meta-analysis was considered adequate when the mean on the MERSQI was greater than 9 (Smith & Learman, 2017).

Most researchers agree that the outliers and influential case diagnostics should be examined before conducting a meta-analysis, because previous studies have

showed that they could affect the validity and robustness of the conclusions from a meta-analysis (Viechtbauer & Cheung, 2010). In the current study, we used influential case diagnostics (e.g., Cook's distances, DFBETAS, and studentized deleted residuals) to detect outliers (Viechtbauer & Cheung, 2010). When an outlier was identified, it was removed in subsequent data analyses, as this would

Table 3 Coding of potential moderators

Potential moderators	Description
Game features	
Game types	
Role-playing	Playing one or more characters and moves in the virtual world
Strategy	Making strategic deployment through the application of game rules and plan
Simulation	Training in a realistic game environment
Construction	Accessing, using, and managing resources in an open game environment
Adventure	Focusing on episodic and exploratory interactions such as exploring the unknown and solving puzzles
Action	Making quick movements and quick reactions in the game
Puzzle	Solving puzzles through logical thinking
Game time	
Up to 1 day	
Between 1 and 7 days	
Over 7 days	
Level of realism	
Schematic	
Cartoon-like	
Photorealistic	
Added game design element	
For learning	If added element focused on learning content
For gaming	If added element focused on gaming experience
Study features	
Education level	
Primary school	Including primary and pre-school education
Junior high school	
Senior high school	
College/University	Including undergraduate and postgraduate education
Subject	
Science	
Math	
Engineering	
Computer	
Learning outcome	
Knowledge	
Cognitive skills	

increase the precision of the estimated average effect sizes (Viechtbauer, 2010).

Statistical analyses

We used R 4.0.5 (R Core Team, 2021) to analyze data extracted from primary studies. Some studies contained multiple effect sizes, thus leading to the issue of nested

data, which, as well-known in research literature, would lead to inflated correlation between variables (Borenstein et al., 2009) and biased estimates in meta-analysis. To handle this issue, we performed a multilevel meta-analysis (Cheung, 2014). In the multilevel meta-analysis, the sources of variation are divided into sampling variance, intra-study variance, and inter-study variance, which would statistically address the issue of non-independence of the multiple effect sizes within one study.

Data extracted from a primary study, including mean, standard deviation, and sample size, were used to calculate the effect size. If mean and standard deviation were not available in a study, other statistics (t , F , or χ^2) were used to estimate an effect size (Glass, 1981). Then, all the obtained effect sizes were converted to Hedges' g (Hedges, 1981) to minimize the impact of small sample studies.

As described previously, in this study, two meta-analyses were performed. In both meta-analyses, three-level random-effects meta-analysis model was used for statistical analysis. More specifically, in each of the two meta-analyses, the random-effects model was used to estimate the overall effect of the study, in which model parameters were obtained through Restricted Maximum Likelihood Estimation (REML), and t tests were used to calculate regression coefficients and corresponding confidence intervals (Viechtbauer, 2010). Subsequently, the sampling variation (level 1) in the study was calculated, and the significance of variation within studies (level 2) and variation between studies (level 3) was determined by a one-tailed log-likelihood ratio test. Finally, the omnibus F test was used to assess whether the moderating effect was significant (Gao et al., 2017).

Finally, publication bias is an existing and widely discussed research problem in meta-analysis. It indicates that studies with larger effect sizes or significant results are more likely to be published than those with smaller effect sizes or insignificant results (Kuppens et al., 2013). In the current meta-analyses, Funnel Plot (Macaskill et al., 2001) and Trim-and-Fill Method (Duval & Tweedie, 2000) were used to detect publication bias. The symmetrical distribution of the funnel plot and the small number of studies suppressed (i.e., $L_0^+ \leq 2$) would suggest a lack of evidence for publication bias. Furthermore, researchers tampering with data in changing a non-significant result into a significant result (p -hacking) could lead to the presence of publication bias (Simonsohn et al., 2014). p -curve, as a method for assessing p -hacking, was used to assess the publication bias due to p -hacking, with the p -curve skewed to the right indicating that p -hacking was impossible for the given data (Simonsohn et al., 2014).

Results

Quality assessment and influential effect sizes detection

The quality appraisal score of the MERSQI was 13.38 (SD=1.02), which was much higher than the prescribed threshold (Table 4) for adequate quality of the included studies, suggesting that the quality of the studies included in the current meta-analyses was very good.

The results of influential effect sizes detections are not presented in the paper for space consideration. The graphs of the influential case diagnostics are available

in Additional file 1, which showed that 4 outliers were flagged in the first and second meta-analyses, respectively, and these outliers varied from 2.652 to 6.618. To avoid the undue impact of these outliers on the results, the outliers were deleted in the subsequent analyses.

Descriptive characteristics

The current meta-analytic data set included 123 studies yielding 217 effect sizes from a cumulative total 11,714 participants. For the first meta-analysis (i.e., studies that

Table 4 MERSQI scores

Domain	MERSQI item	Item	No. studies (%)	Mean (SD)	
				Item	Domain
Study design	1. Study design			2.59 (0.49)	2.59 (0.49)
	Single group post-test only	1	0		
	Single group pre–post-test	1.5	0		
	Nonrandomized, 2 groups	2	50 (40.7)		
Sampling	Randomized controlled trial	3	73 (59.3)		
	2. No. of institutions studies			0.64 (0.33)	2.13 (0.34)
	1	0.5	104 (84.6)		
	2	1	4 (3.3)		
	> 2	1.5	15 (12.1)		
	3. Response rate, %			1.5 (0.04)	
	< 50 or not reported	0.5	0 (0)		
	50–74	1	1 (0.8)		
	≥ 75	1.5	122 (99.2)		
	Type of data	4. Type of data			2.76 (0.65)
Assessment by subjects		1	15 (12.2)		
Objective measurement		3	108 (87.8)		
Validity of instrument	5. Internal structure			0.4 (0.49)	1.4 (0.49)
	Not reported	0	74 (60.2)		
	Reported	1	49 (39.8)		
	6. Content			0.99 (0.09)	
	Not reported	0	1 (0.8)		
	Reported	1	122 (99.2)		
	7. Relationships to other variables			0.01 (0.09)	
	Not reported	0	122 (99.2)		
	Reported	1	1 (0.8)		
	Data analysis	8. Appropriateness of analysis			1 (0)
Inappropriate		0			
Appropriate		1	123 (100)		
9. Complexity of analysis				2 (0)	
Descriptive analysis only		1			
Beyond descriptive analysis	2	123 (100)			
Outcomes	10. Outcomes			1.5 (0)	1.5 (0)
	Satisfaction, attitudes, opinions	1			
	Knowledge, skills	1.5	123 (100)		
	Behaviors	2			
	Patient/health care outcome	3			

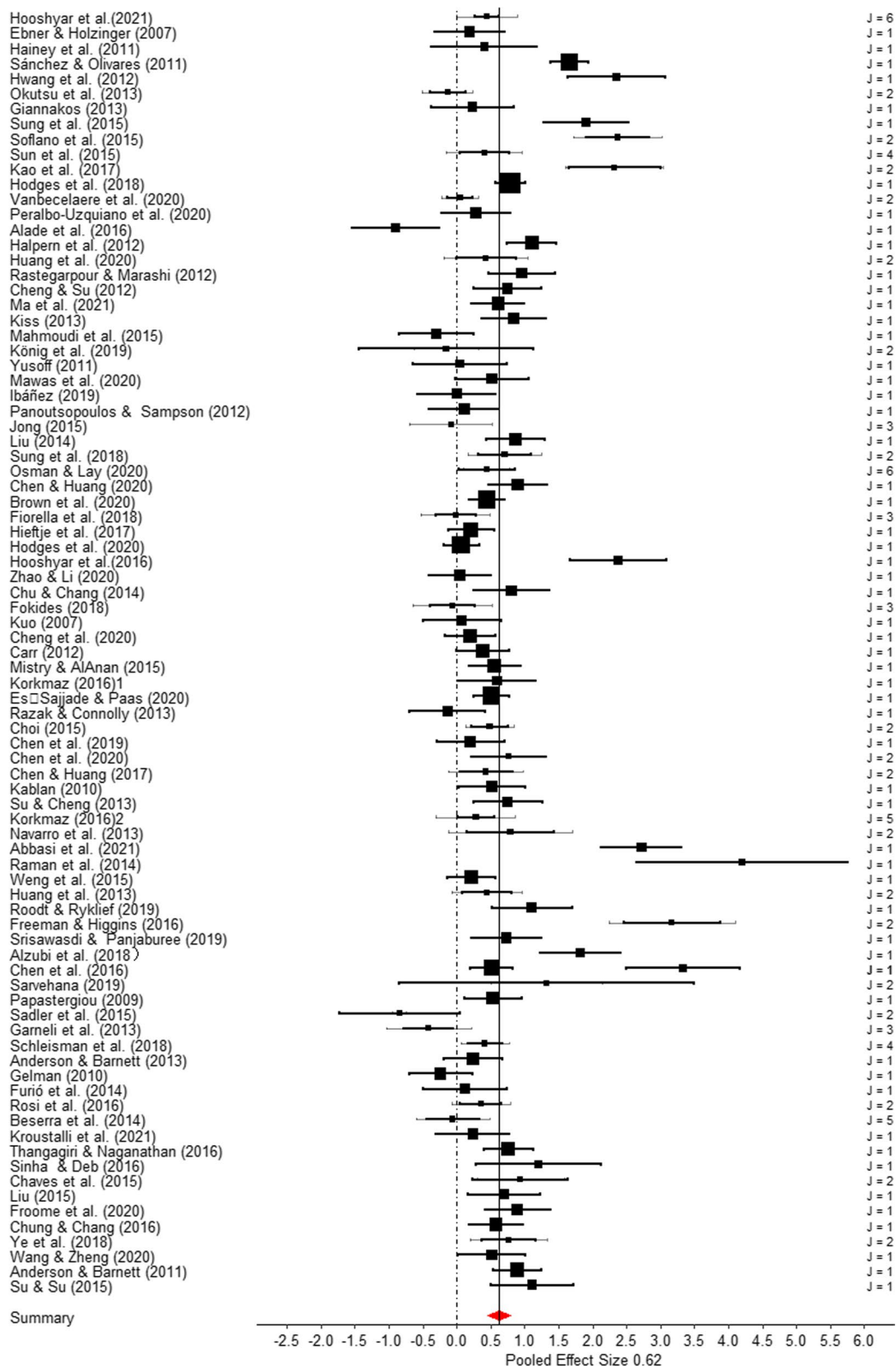


Fig. 3 Forest plot of meta-analysis 1 (game-based STEM learning vs. alternative activities in STEM learning)

compared game-based STEM learning with traditional STEM learning), the number of included studies was 86 with 136 effect sizes. For the second meta-analysis (i.e., studies in which a base game version was compared with an enhanced game version containing an added game mechanics element), the number of studies was 44 with 81 effect sizes. The included studies were published between 2005 and 2020, with the majority (97%) published after 2010. Most studies were published journal articles (89%) and involved the pre–post-test design (86%). The study samples included students of primary school and below (40%), junior high school students (22%), senior high school students (13%), and college and university students (25%). The included studies were conducted mainly in Europe (24%) and Asia (50%), and the studies involved STEM areas of science (50%), mathematics (25%) and computer courses (20%). The types of games used in the studies were mainly puzzle (38%), role-playing (27%), and simulation (14%) games. Half of the games were cartoon-like, 24% were photorealistic and 13% were schematic.

Meta-analysis 1: general effect of game-based STEM learning (RQ1a)

The first meta-analysis focused on the general effect of game-based STEM learning when compared with traditional STEM learning without using digital games. As described previously, 86 studies that compared digital game-based STEM learning with traditional STEM learning without the use of digital games were used in this meta-analysis, and a total of with 136 effect sizes from these primary studies were quantitatively synthesized. These effects sizes are shown graphically in Fig. 3 in the form of forest plot. The overall effect of digital game-based STEM learning over traditional STEM learning was shown to be larger than a medium effect ($g=0.624$, $t=7.403$, $p<0.001$, 95% CI [0.457, 0.790]), indicating that students using digital educational games in STEM learning significantly outperformed their counterparts who used alternative learning activities in STEM learning. The one-tailed log likelihood ratio test for the variation of the effect sizes suggested significant variations both within and between studies ($Q(135)=2095.093$, $p<0.001$, $I^2=38.965\%$, $I_3^2=54.479\%$), indicating statistical heterogeneity among the effect sizes that warrant follow-up analysis for potential moderators (see Table 5, upper panel).

Moderator analysis for meta-analysis 1 (RQ1b)

The three-level mixed-effects model identified game type as a significant moderator ($F(6, 119)=2.229$, $p<0.05$). Specifically, studies involving strategy games showed extremely large effect sizes ($g=1.841$), while those

involving role-playing ($g=0.586$), puzzle ($g=0.551$) and adventure ($g=0.701$) games showed effect sizes that were medium and above, Studies involving action ($g=0.394$) and simulation ($g=0.464$) games showed effect sizes between small and medium. As too few studies involved construction games, its effect size should not be trusted or interpreted.

There were significant differences in effect sizes from the studies that involved different STEM subject areas ($F(3, 132)=5.7$, $p<0.01$). Studies that used game-based learning in computer courses had very large average effect size ($g=1.077$), and the studies using game-based learning in science courses showed above medium average effect size ($g=0.674$). Studies involving game-based learning in mathematics courses ($g=0.179$) and in engineering courses ($g=0.271$) showed small average effect sizes.

The types of learning outcomes also appeared to be a significant moderator of the effect sizes from different studies ($F(2, 133)=4.273$, $p<0.05$), with the studies involving cognitive skills as measured outcome showing much larger average effect size ($g=0.91$; a large effect) than the studies that measured knowledge as the learning outcome ($g=0.538$; a medium effect). Other game or study features (game time, game realism, participants' educational level did not turn out to be statistically significant moderators for effect sizes across the studies (see Table 6).

Meta-analysis 2: enhancement effect of added game-design elements in STEM learning (RQ2a)

The second meta-analysis focused on the enhancement effect of some added game-design mechanics. As described above, 44 studies that compared STEM learning condition of using a base game version and the condition of using the same game but with some added game-design elements were used in this meta-analysis, and a total of with 81 effect sizes from these primary studies were obtained and quantitatively synthesized. These effects sizes are shown graphically in Fig. 4 in the form of forest plot. The three-level meta-analysis revealed that games with added game-design mechanics showed enhancement effect over the base game version condition for improving students' STEM learning ($g=0.301$, $t=4.342$, $p<0.001$, 95% CI [0.163, 0.438]). Furthermore, the three-level random effect model showed that both within-study variance ($p<0.001$) and between-study heterogeneity among the effect sizes were statistically significant ($Q(80)=357.037$, $p<0.001$, $I^2=36.682\%$, $I_3^2=43.160\%$) (see Table 5, lower panel), indicating the need for moderator analysis to identify what game or study features could have contributed to the inconsistent effect sizes across the studies.

Table 5 Results for the overall effects of two meta-analyses

Meta-analysis	No. St	No. ES	Mean g (SE)	95% CI	t value	Q total	% var. level 1	Level 2 variance	% var. level 2	Level 3 variance	% var. level 3
#1: Overall effect of game-based STEM learning over STEM learning with alternative activities											
	86	136	0.624 (0.084)	[0.457, 0.790]	7.403***	2095.093***	6.556	0.256***	38.965	0.359***	54.479
#2: Enhancement effect of added game-design element over base game version											
	44	81	0.301 (0.069)	[0.163, 0.438]	4.342***	357.037***	20.159	0.094***	36.682	0.11	43.160

No. St: number of studies; no. ES: number of effect sizes; mean g: mean effect size (g); SE: standard error; CI: confidence interval; Sig: significance; Q_{total}: homogeneity test of all effect sizes; Var.: variance; Level 1 variance: sampling variance of observed effect sizes; level 2 variance: variance between effect sizes extracted from the same study (heterogeneity of within-study); level 3 variance: variance between studies (heterogeneity of between-study)

***p < 0.001

Table 6 Results of moderator analyses for meta-analysis 1 (game-based STEM learning vs. alternative activities in STEM learning)

Moderator variables	No. St	No. ES	Mean g	SE	t	p	95% CI lower	95% CI upper	Omnibus test F (df1, df2)	p	Level 2 variance	Level 3 variance
Game type	82	126							$F(6, 119)=2.229$	0.045	0.235	0.346
Role playing	18	29	0.586	0.178	3.29	0.001	0.233	0.939				
Strategy	5	6	1.841	0.355	5.193	<0.001	1.139	2.543				
Action	6	10	0.394	0.31	1.274	0.205	-0.219	1.007				
Construction	1	2	0.703	0.709	0.992	0.323	-0.7	2.106				
Adventure	5	13	0.701	0.317	2.212	0.029	0.074	1.329				
Simulation	15	21	0.464	0.201	2.309	0.023	0.066	0.862				
Puzzle	32	45	0.551	0.136	4.04	<0.001	0.281	0.821				
Game time	65	106							$F(2, 103)=0.888$	0.415	0.295	0.312
Up to 1 day	21	33	0.792	0.165	4.784	<0.001	0.463	1.12				
Between 1 and 7 days	11	20	0.444	0.233	1.905	0.06	-0.018	0.905				
Over 7 days	33	53	0.573	0.132	4.332	<0.001	0.311	0.835				
Level of realism	74	117							$F(2, 113)=0.580$	0.562	0.247	0.466
Cartoon-like	39	60	0.545	0.135	4.045	<0.001	0.278	0.813				
Photorealistic	25	36	0.78	0.172	4.545	<0.001	0.44	1.119				
Schematic	10	21	0.66	0.262	2.524	0.013	0.142	1.179				
Educational level	86	136							$F(3, 132)=1.379$	0.252	0.256	0.354
Elementary school	35	54	0.543	0.132	4.118	<0.001	0.282	0.804				
Middle school	13	22	0.461	0.212	2.175	0.031	0.042	0.88				
High school	14	28	0.534	0.199	2.685	0.008	0.141	0.927				
College/University	24	32	0.907	0.164	5.516	<0.001	0.582	1.232				
Subject	86	136							$F(3, 132)=5.700$	0.001	0.253	0.278
Science	40	61	0.674	0.114	5.895	<0.001	0.448	0.9				
Math	21	34	0.179	0.156	1.145	0.254	-0.13	0.487				
Engineering	5	8	0.271	0.316	0.858	0.392	-0.354	0.896				
Computer	20	33	1.077	0.165	6.526	<0.001	0.75	1.403				
Learning outcome	86	136							$F(2, 133)=4.273$	0.016	0.276	0.284
Knowledge	65	90	0.538	0.092	5.839	<0.001	0.356	0.721				
Cognitive skills	26	40	0.91	0.137	6.641	<0.001	0.639	1.181				
Mixed	3	6	-0.015	0.394	-0.039	0.969	-0.795	0.764				

No. St: number of studies; no. ES: number of effect sizes; g: mean effect size (g); SE: standard error; CI: confidence interval; F (df1, df2): omnibus test of regression coefficient in the model; p: p-value of the omnibus test; level 2 variance: variance between effect sizes extracted from the same study (heterogeneity of within-study); level 3 variance: variance between studies (heterogeneity of between-study)

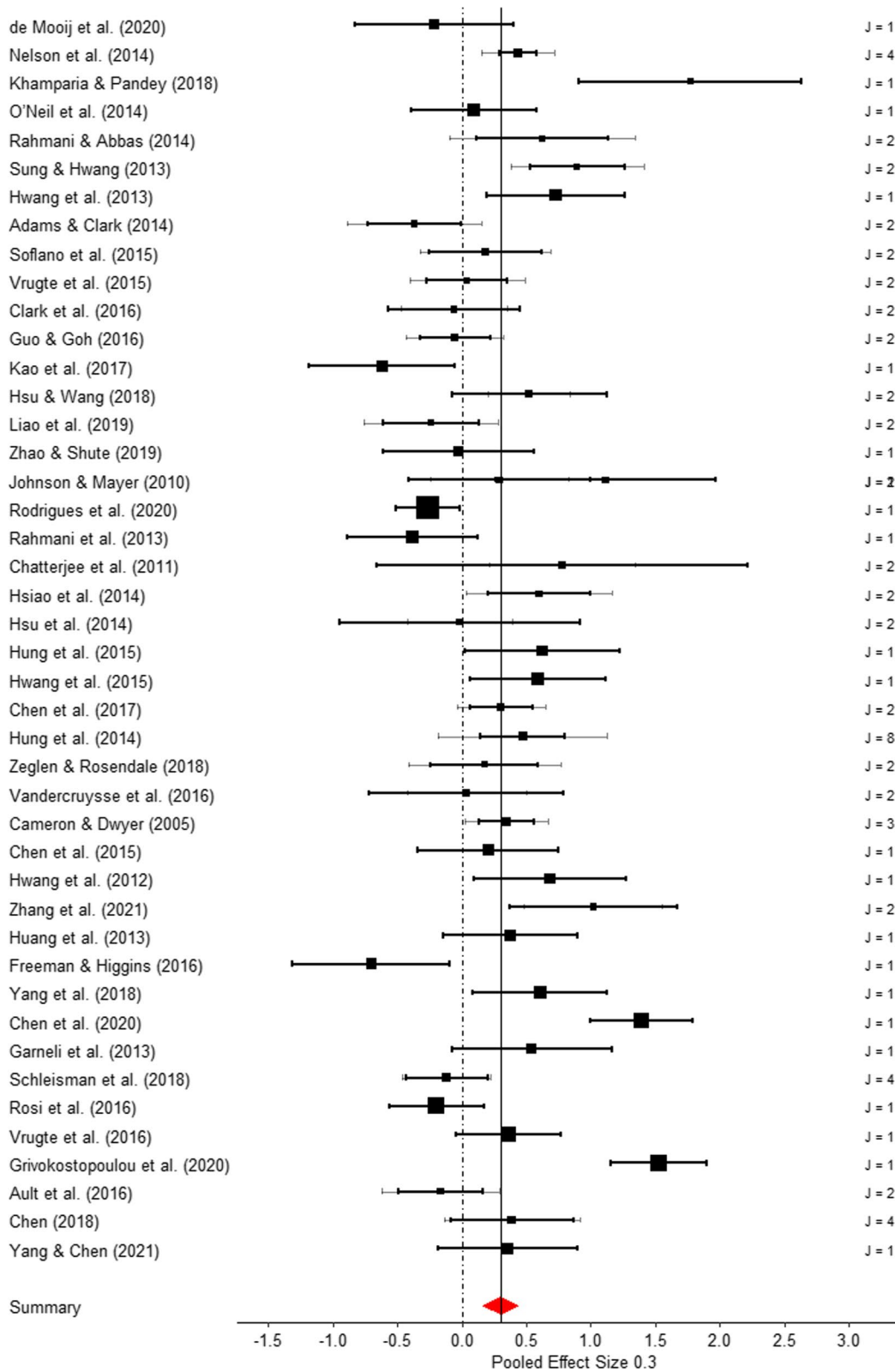


Fig. 4 Forest plot of meta-analysis 2 (enhancement effect of added game-design element over base game version)

Table 7 Results of moderator analyses for meta-analysis 2 (enhancement effect of added game-design element over base game version)

Moderator variables	No. St	No. ES	Mean g	se	t	p	95% CI lower	95% CI upper	Omnibus test F (df1, df2)	p	Level 2 variance	Level 3 variance
Game type	44	81							$F(5, 75) = 2.230$	0.06	0.087	0.092
Role playing	17	32	0.397	0.103	3.849	<0.001	0.192	0.603				
Strategy	1	1	-0.618	0.512	-1.207	0.231	-1.638	0.402				
Action	1	2	-0.165	0.404	-0.41	0.683	-0.97	0.639				
Adventure	5	7	-0.031	0.202	-0.153	0.879	-0.432	0.371				
Simulation	3	3	0.779	0.276	2.824	0.006	0.229	1.328				
Puzzle	17	36	0.287	0.104	2.769	0.007	0.08	0.493				
Game time	39	68							$F(2, 65) = 0.342$	0.711	0.126	0.098
Up to 1 day	22	39	0.282	0.099	2.841	0.006	0.084	0.48				
Between 1 and 7 days	8	15	0.142	0.174	0.818	0.417	-0.205	0.489				
Over 7 days	9	14	0.329	0.162	2.033	0.046	0.006	0.652				
Level of realism	40	74							$F(2, 71) = 0.031$	0.97	0.072	0.15
Cartoon-like	26	49	0.295	0.096	3.076	0.003	0.104	0.487				
Photorealistic	7	13	0.284	0.181	1.566	0.122	-0.078	0.646				
Schematic	7	12	0.342	0.185	1.852	0.068	-0.026	0.71				
Educational level	44	81							$F(3, 77) = 1.423$	0.242	0.096	0.099
Elementary school	16	28	0.403	0.117	3.444	<0.001	0.17	0.636				
Middle school	16	28	0.19	0.112	1.704	0.092	-0.032	0.412				
High school	4	8	0.016	0.227	0.07	0.944	-0.436	0.468				
College/University	8	17	0.442	0.145	3.047	0.003	0.153	0.731				
Subject	44	81							$F(3, 77) = 2.374$	0.077	0.088	0.098
Science	24	53	0.295	0.087	3.394	0.001	0.122	0.468				
Math	12	15	0.259	0.138	1.882	0.064	-0.015	0.533				
Engineering	1	1	1.528	0.471	3.247	0.002	0.591	2.465				
Computer	7	12	0.223	0.166	1.345	0.182	-0.107	0.553				
Learning outcome	44	81							$F(1, 79) = 6.055$	0.016	0.088	0.099
Knowledge	28	52	0.428	0.085	5.059	<0.001	0.26	0.596				
Cognitive skills	16	29	0.091	0.108	0.849	0.398	-0.123	0.305				
Added game design element	44	81							$F(1, 79) = 4.315$	0.041	0.085	0.112
For gaming	25	40	0.175	0.091	1.917	0.059	-0.007	0.357				
For learning	25	41	0.432	0.093	4.629	<0.001	0.246	0.617				

No. St: number of studies; no. ES: number of effect sizes; g: mean effect size (g); SE: standard error; CI: confidence interval; F (df1, df2): Omnibus test of regression coefficient in the model; p: p-value of the omnibus test; level 2 variance: variance between effect sizes extracted from the same study (heterogeneity of within-study); level 3 variance: variance between studies (heterogeneity of between-study)

Moderator analysis for meta-analysis 2 (RQ2b)

The results of moderating effect analysis showed that learning outcome ($F(1, 79) = 6.055, p < 0.016$) was a likely moderator for the observed inconsistent findings about the enhancement effect of added game-design element over base game for STEM learning. More specifically, with regard to the enhancement effect of added game-design element in STEM learning, in the studies that measured knowledge as learning outcome, the average effect size of such enhancement effect was 0.428 (a medium effect size), compared with a very small average effect size of 0.091 from the studies that measured cognitive skill as the learning outcome.

In this study, two types of added game-design element were coded: game element added for improving gaming experience (For Gaming), or game element added for learning content (For Learning). It turned out that, the enhancement effect of the added game-design element differed depending on which type of added game-design element: for gaming vs. for learning ($F(1, 79) = 4.315, p < 0.05$). In the studies with added game-design element for gaming, the average enhancement effect is small ($g = 0.175$), in contrast to the average enhancement effect of almost medium effect size ($g = 0.432$) from the studies with added game-design element for learning. Other game features or study features (publication type, game type, level of realism, sample's education level, or STEM subject content area) were statistically not significant (see Table 7).

Publication bias

For meta-analysis 1: The funnel plot appeared to be symmetrical, and Trim-and-fill method showed that the L_0^+ value was 0, far lower than the recommended threshold ($L_0^+ = 2$) for publication bias. In addition, the right-skewness test of the p -curve analysis was statistically significant ($Z = -24.664, p < 0.001$), suggesting no evidence for p -hacking (Additional file 1).

For meta-analysis 2: The funnel plot appeared to be symmetrical, and the Trim-and-fill method showed a L_0^+ value of 0, suggesting no evidence for publication bias. In addition, the p -curve analysis showed that the curve had statistically significant right-skewness ($Z = -8.465, p < 0.001$), indicating no evidence for p -hacking (see Additional file 1).

Discussion

Effect of DGBL in STEM learning and enhancement effect of game-design elements

DGBL, as a pedagogical strategy for STEM education, may promote learning motivation, enhance STEM knowledge and develop problem-solving skills (Klopfer & Thompson, 2020). While game-based STEM learning

may benefit student STEM learning, previous empirical studies have not reached consistent conclusions on the effectiveness of DGBLs. To fill the research gaps, the current study, following the research framework proposed by Mayer (2015, 2019), comprehensively and systematically reviewed two types of previous empirical studies on game-based STEM learning:

- 1) Studies that compared game-based STEM learning with traditional STEM learning with other alternatives. In the research framework proposed by Mayer (2015), this research direction was defined as "media comparison research". The purpose of "media comparison research" is to assess the effect of game-based STEM learning over traditional STEM learning. Our first meta-analysis quantitatively summarized the findings across these studies of "media comparison research" in the literature;
- 2) Studies that compared two different conditions of game-based STEM learning: one condition involved the use of a base game version, and the other condition involved the use of the same game but with some added game-design elements. In the research framework proposed by Mayer (2015), this research direction was defined as "value-added research". The purpose of this research direction is to assess the enhancement effect of the added game-design elements over the base game. Our second meta-analysis quantitatively summarized the findings across the studies of "value-added research" in the literature.

The findings from the first meta-analysis for the studies of "media comparison research" revealed that students with game-based STEM learning significantly outperformed their counterparts with traditional STEM learning, and the overall effect was a medium to large effect size of $g = 0.624$ across all the studies. Consistent with previous empirical and also meta-analysis studies (Brown et al., 2020; C. -L.D. Chen et al., 2016, C. -H. Chen et al., 2016; Halpern et al., 2012), this finding provides strong support that DGBL is likely an effective pedagogical strategy for promoting STEM education. The findings also lend support to the cognitive theory of game-based learning connecting playing game and cognitive development (Mayer, 2020).

The findings from the second meta-analysis for the studies of "value-added research" showed an overall small to medium effect size ($g = 0.301$), indicating that across all the studies that examined the enhancement effect of added game-design elements in game-based STEM learning, students using games with added game-design elements outperformed

their counterparts who used the base game. This result is consistent with the findings of a previous review of game mechanic design (Tsai & Tsai, 2018, 2020), in which it was suggested that adding some game mechanics in education game could improve the effect of digital game-based learning. Further, the review findings of this study responded well to the call by Mayer (2019) for more empirical evidence on the effectiveness of DGBL, and on the value of game mechanics for enhancing learning.

Factors contributing to inconsistent findings across the studies

The statistically significant heterogeneity tests for the effect sizes across the studies in both meta-analyses in this study suggested that the empirical findings about game-based STEM learning varied across the studies in the literature. To gain a better understanding of the underlying factors contributing to such inconsistent findings as observed in the research literature about DGBL, we investigated the potential moderating variables (including game characteristics and study features) in both meta-analyses conducted in this study.

Moderating variables in meta-analysis 1

In the first meta-analysis on “media comparison research” studies, which compared game-based STEM learning vs. STEM learning without the use of digital games, three game/study features (game type, subject area, and learning outcome) were shown to have statistically contributed to the variations of effect sizes across the studies, while other three game/study features (education level, game time, and level of realism) were not such statistical contributors.

For game type, seven types of games were coded for moderator analysis, and these seven types of games showed average effect sizes ranging from 0.394 to 1.841, with strategy games showing largest effect size ($g=1.841$). This finding suggests that strategy games could be more effective in the STEM learning environment than other types of games. The report by NRC (2011) discussed that an effective STEM learning environment should include handling cognitively demanding tasks, focusing on complex problem solving, and engaging in sound planning. In contrast to other types of games, a strategy game tends to involve complex tasks with an emphasis on tactics and on long-term planning (Gerber & Scott, 2011). Learners need to carefully evaluate and analyze problems, and actively develop critical thinking and problem-solving skills for success in a strategy game. Compared with strategy games, other game types, such as simulation, puzzle

solving and role-playing, do not require learners to invest as much cognitive resources to solve problems. The review by Romero et. al. (2015) also pointed out that players in strategy games showed higher critical problem-solving ability than those playing other types of games. Therefore, a well-designed strategy game may provide students with a more effective STEM learning environment, thereby helping students achieve better learning benefits. It is worth noting that there have been relatively few empirical studies focusing on strategy games (and on some other types such as action, and building games), and more research is needed to focus on these games in the future.

With regard to STEM subject area, the results of current meta-analysis indicated that the positive effect of game-based learning was much larger from the studies involving science and computer learning as STEM areas ($g=0.674$ and 1.077 for science and computer learning, respectively), but the effect sizes were much smaller from the studies involving math and engineering learning ($g=0.179$ and 0.271 for math and engineering learning, respectively). These findings are consistent with some previous reviews (Talan et al., 2020; Wouters & van Oostendorp, 2013) suggesting that digital educational games are powerful tools to support science learning. Tsai and Tsai (2020) pointed out that the effectiveness of game-based science learning came from a well-designed game environment and strong instructional support. For computer science, DGBL could have advantages over other learning methods in multiple areas, such as learner-centricity, interactivity, and immediate feedback (Kazimoglu, 2020). These advantages not only could motivate students to immerse themselves in tedious programming practices, but could also support the implementation of basic programming constructs to ease the difficulty of learning to do programming. Some empirical studies on computer programming also showed that digital educational games could enhance learning more than other instructions (Chaves et al., 2015; Freeman & Higgins, 2016).

Although game-based STEM learning significantly and positively enhanced STEM knowledge and cognitive skills, the studies involving cognitive skills as learning outcome had large effect size ($g=0.91$) than those involving STEM knowledge as learning outcome ($g=0.538$). The report by NRC (2011) discussed that DGBL for STEM field involving high cognitive tasks and complex problems focused more on enhancing cognitive skills such as creative thinking and problem solving. Similarly, Klopfer and Thompson (2020) also discussed that game-based learning environment that provides students with simulation opportunities, problem-solving strategies and

various instructional support could effectively improve students' higher-order cognitive abilities. Therefore, digital educational games in the STEM field may promote students' cognitive skills more than their knowledge acquisition (C. -Y. Chen et al., 2020; Huang et al., 2020).

Several game/study features (game time, game realism, education level) did not turn out to be statistically significant moderators for the effect sizes across the studies, despite some descriptive differences in the average effect sizes under different conditions. As these game/study features were statistically non-significant in the moderator analysis, we refrained from any further discussion on these variables.

Moderating variables in meta-analysis 2

In the second meta-analysis on "value-added research" studies, which compared two conditions of game-based STEM learning: using base game version vs. using the same game but with added game-design elements. The results of the moderator analysis showed that two game/study features (game-design element, learning outcome) were shown to have statistically contributed to the variations of effect sizes across the studies, while other game/study features (game type, game time, level of realism, educational level, and STEM subject area) were not such statistical contributors.

With regard to the statistically significant moderator of "game-design element" in Table 7, in the "value-added research" studies, the primary focus was on the enhancement effect of added game-design elements over the same game but without such added game-design elements. Previously in Table 5 ("Results for the Overall Effects of Two Meta-Analyses"), it was already shown that the added game-design elements had positive enhancement effect ($g=0.301$) on STEM learning across all the "value-added research" studies. The moderator analysis results shown in Table 7 further informed us that the two types of added game-design elements (design elements for gaming experience, design elements for content learning) showed different enhancement effects, with game-design elements intended for content learning showing larger enhancement effect ($g=0.432$; a medium effect size) than those intended for gaming experience ($g=0.175$; a small effect size).

The cognitive theory of game-based learning could provide a possible explanation for this finding (Mayer, 2020), which stipulates that the effectiveness of DGBL stems from its powerful learning mechanism that minimizes irrelevant cognitive processing and promotes generative processing related to learning tasks. On the other hand, digital educational games with excessive gaming elements or mechanisms may distract learners' attention from the learning tasks, and result in extra cognitive

processing that may overburden the limited cognitive processing capacity of students. Some previous studies also showed that digital educational games with some gaming mechanic elements (e.g., narrative, multiplayer) did not perform better than the games without these gaming mechanic elements (Tsai & Tsai, 2020; Wouters et al., 2013).

With regard to learning outcome, the results of moderating effect analysis showed that the enhancement effect of added game-design elements showed larger effect ($g=0.428$) in studies with STEM knowledge as outcome, but very small enhancement effect in studies with cognitive skills as outcome ($g=0.091$). This suggests game learning with added game mechanic elements significantly enhanced STEM knowledge acquisition, but did little for cognitive skills. Previously in our first meta-analysis for "media comparison research" studies, we observed that game-based STEM learning, when compared with traditional STEM learning, could be more effective for developing cognitive skills than for STEM knowledge acquisition. Here, in the second meta-analysis, however, the comparison was between game-based STEM learning *with* or *without* added game-design elements. Thus, the foci of these two meta-analyses, as well as the findings of the moderator analyses, were entirely different, thus not to be compared or interpreted together. It is very likely that what game-design elements to add to an educational game must match what is already in the game to have meaningful enhancement effect over the original game version.

As shown in Table 7, five game/study features (game type, game time, level of realism, educational level, and STEM subject area) were not statistically significant moderators for the effect sizes across these "value-added research" studies investigating the enhancement effect of added game-design elements. Although there appeared to be some obvious descriptive differences in the average effect sizes across different levels of some of these game/study features (e.g., game type, subject), it should be noted that the numbers of included studies/effect sizes for some levels were too small for meaningful statistical comparisons. For example, under "game type", the numbers of studies/effect sizes for Strategy games, Action games, Adventure games, and Simulation games were all small. Similarly, under STEM "Subject", the number of studies/effect sizes for "Engineering" was also too small. The large variations in the effect sizes across the levels appeared to have been contributed by these levels with very small number of studies/effect sizes (e.g., under "Subject", Engineering ($n=1$) had effect size of 1.528, while others had 0.223–0.295; under "Game Type", the four game types with small n had widely varying effect

sizes, while two game types with much larger n had mean effect sizes not much different from each other). Without the levels with very small n conditions, the descriptive differences for these moderator variables would be much smaller. It was likely that this could be a reason for the statistically non-significant tests for these moderators, despite some descriptive differences being obvious. As these game/study features were statistically non-significant in the moderator analysis, we refrained from any further discussion on these variables.

Limitations and future research directions

There are some deficiencies in the current research. First, due to the limited information available from the included studies, our meta-analyses reported only cognitive outcomes. Wouters et al. (2009) discussed that the effects of DGBL are not only manifested in cognition, but also in motivation, emotion, and motor skills. Future research in DGBL should consider and measure other types of meaningful outcomes to provide more comprehensive evidence for the effectiveness of DGBL.

Second, limited by the information available from the included studies, in our second meta-analysis on the “value-added research” studies that investigated the enhancement effect of game-design mechanic elements, we were only able to roughly divide the added game mechanic elements into those for gaming experience and those for learning. In fact, game-design mechanics include other types, such as narrative, collaboration, multiplayer, feedback, concept maps, and coaching, etc. (Plass et al., 2020). Wouters et al. (2013) believed that not all game-design mechanic elements were necessary for DGBL. Future research in DGBL may examine the effect of some specific game-design mechanic elements.

Finally, this study focused primarily on the moderating effects of game characteristics and study features from a cognitive perspective. However, some sociocultural attributes such as cultural differences may have an impact on game-based STEM learning (Wahono et al., 2020). Future research studies in DGBL may consider including sociocultural attributes in the studies, so that the potential effects of sociocultural attributes on game-based STEM learning can be better understood.

Findings and suggestions

Built upon previous studies on DGBL effectiveness in STEM education, the current study enhanced our understanding of the positive effect of game-based STEM learning on knowledge gains. In addition to extending and updating the findings of previous reviews on the general effect of DGBL on knowledge gains in STEM disciplines, the current study further suggested that STEM digital games were more effective for developing

cognitive skills than for facilitating knowledge acquisition. This finding provides empirical support for the claim that game-based STEM education is particularly suitable for developing higher-order skills. Thus, we suggest that future research efforts should focus more on the facilitative effects of digital educational games on students' cognitive skills. National Research Council (2011) indicated that DGBL is critical for improving learners' cognitive skills in STEM education. Previous research has also shown that educational games could promote cognitive skills (C. -Y. Chen et al., 2020; Huang et al., 2020). Therefore, educators could increase the use of DGBL in STEM education that involve high levels of cognitive tasks and complex problems to promote students' cognitive skills.

In addition, we also found that educational games appeared to be more effective for science and computer learning, and they were suitable for students of any educational level. On the other hand, game-play time and the level of realism of game screen did not appear to have impact on the effect of game-based STEM learning. These findings suggest that educational practitioners can consider more use of educational games in science and computing classrooms without too much concern about the potential impact of playtime and the level of realism of games on student' academic performance. Furthermore, current work suggests that strategy digital games may be more effective than other types of digital games in STEM learning. Plass et al. (2020) argued that, as an important gameplay design, game type may have a significant impact on the effectiveness of DGBL. Consistent with our finding, Romero et al. (2015) argued that strategy games could foster learning better than role-playing and competitive games. Therefore, research should focus on the effect of game genre on the effectiveness of DGBL for STEM education in the future work.

The current study also sheds lights on some new aspects about feedback effectiveness that were not known previously. The new findings about the enhancement effect of game-design elements in our study revealed that, in general, enhanced game design outperformed the basic game version in enhancing students' STEM academic performance. This empirical finding strongly suggests that more attention should be given to the game design elements for game-based STEM education. This finding also provides empirical support for an integrated design framework in game-based learning from a cognitive perspective (e.g., Plass et al., 2015), indicating that the application of appropriate game design elements could be useful for cognitive engagement. More importantly, the current synthesis also revealed that the added elements related to learning mechanism were more effective in promoting students' cognitive performance than

those related to gaming mechanism. Therefore, for the purpose of enhancing the effectiveness of digital games in STEM education, game designers should pay more attention to the design of learning mechanisms.

In summary, this meta-analytic review strengthened that digital games as instructional medium for STEM education are effective in promoting not only student's knowledge gains but also their cognitive skills. More importantly, the findings suggest that game design mechanism could play an important role in terms of effectiveness of game-based STEM education. Hence, the focus of future research on game-based STEM learning should shift from the proof-of-concept research (i.e., whether STEM games are effective) to the value-added research (i.e., how to design games to make them more effective), so that future educational games can be more effective in STEM education.

Conclusions

This meta-analysis study makes a valuable contribution in enhancing our understanding about the effectiveness of DGBL about STEM learning and the enhancement effect of added game-design elements in STEM learning. We conclude from this study that students with DGBL significantly outperformed their counterparts with traditional learning in STEM education. Furthermore, we also conclude from this study that students using games with added game-design elements outperformed their counterparts who used the base game. To facilitate learning in digital game-based STEM learning, designers of games are encouraged to embed appropriate game elements (e.g., pedagogical agents, self-explanation strategies, concept maps, feedback, and adaptation) into the process of game-based learning. In addition, moderator analysis reveals that the inconsistent findings about the effectiveness of DGBL in STEM education are accounted for by game type, learning subject as well as learning outcome, and the inconsistent enhancement effect of game-design elements are accounted for by types of added game-design elements.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40594-023-00424-9>.

Additional file 1: Detailed results for Influential effect sizes detection and publication bias test.

Acknowledgements

Authors would like to thank the reviewers and editors, whose comments and feedback helped us to improve the original manuscript.

Author contributions

ZC led the study design, helped statistical analysis, led the interpretation of results, and drafted the manuscript. YG reviewed studies, extracted the data, performed the statistical analyses and draft the manuscript. YY and LK reviewed studies and extracted the data. XF participated in interpretation of results and helped to draft the manuscript. RHT contributed to manuscript revision. ZC, YG and XF read and approved the final manuscript.

Funding

This research was supported by a grant from the National Natural Science Foundation of China (No. 62107018) to Zhihui Cai. This research was also supported by Grant CCNU22QN020 and CCNU22QN013 from the Fundamental Research Funds for the Central Universities.

Availability of supporting data

Additional data to this article can be found online.

Declarations

Competing interests

The authors declare that they have no competing interests.

Received: 17 May 2022 Accepted: 1 May 2023

Published online: 19 May 2023

References

- Adams, D. M., & Clark, D. B. (2014). Integrating self-explanation functionality into a complex game environment: Keeping gaming in motion. *Computers & Education*, 73, 149–159. <https://doi.org/10.1016/j.compedu.2014.01.002>
- Aladé, F., Lauricella, A. R., Beaudoin-Ryan, L., & Wartella, E. (2016). Measuring with Touchscreen technology and preschoolers' STEM learning. *Computers in Human Behavior*, 62, 433–441. <https://doi.org/10.1016/j.chb.2016.03.080>
- Arnab, S., Lim, T., Carvalho, M. B., Bellotti, F., de Freitas, S., Louchart, S., Suttie, N., Berta, R., & De Gloria, A. (2015). Mapping learning and game mechanics for serious games analysis. *British Journal of Educational Technology*, 46(2), 391–411. <https://doi.org/10.1111/bjet.12113>
- Arztmann, M., Hornstra, L., Jeuring, J., & Kester, L. (2023). Effects of games in STEM education: A meta-analysis on the moderating role of student background characteristics. *Studies in Science Education*, 59(1), 109–145. <https://doi.org/10.1080/03057267.2022.2057732>
- Ball, C., Huang, K.-T., Cotten, S. R., & Rikard, R. V. (2020). Gaming the SySTEM: The relationship between video games and the digital and STEM divides. *Games and Culture*, 15(5), 501–528. <https://doi.org/10.1177/1555412018812513>
- Becker, B. J. (2000). Multivariate meta-analysis. In S. D. Brown & H. E. A. Tinsley (Eds.), *Handbook of applied multivariate statistics and mathematical modeling* (pp. 499–525). Academic Press. <https://doi.org/10.1016/B978-012691360-6/50018-5>
- Beserra, V., Nussbaum, M., Zeni, R., Rodriguez, W., & Wurman, G. (2014). Practising arithmetic using educational video games with an interpersonal computer. *Educational Technology & Society*, 17(3), 343–358.
- Borenstein, M., Hedges, L. V., Higgins, J., & Rothstein, H. R. (2009). *Introduction to meta analysis*. Wiley.
- Brown, F. L., Farag, A. I., Hussein Abd Alla, F., Radford, K., Miller, L., Neijenhuijs, K., Stubbé, H., de Hoop, T., Abdullatif Abbadi, A., Turner, J. S., Jetten, A., & Jordans, M. J. D. (2020). Can't wait to learn: A quasi-experimental mixed-methods evaluation of a digital game-based learning programme for out-of-school children in Sudan. *Journal of Development Effectiveness*. <https://doi.org/10.1080/19439342.2020.1829000>
- Byun, J., & Joung, E. (2018). Digital game-based learning for K-12 mathematics education: A meta-analysis. *School Science and Mathematics*, 118(3–4), 113–126. <https://doi.org/10.1111/ssm.12271>
- Chaves, R. O., von Wangenheim, C. G., Costa Furtado, J. C., Bezerra Oliveira, S. R., Santos, A., & Favero, E. L. (2015). Experimental evaluation of a serious

- game for teaching software process modeling. *IEEE Transactions on Education*, 58(4), 289–296. <https://doi.org/10.1109/te.2015.2411573>
- Chen, C.-L.D., Chang, C.-Y., & Yeh, T.-K. (2016). The effects of game-based learning and anticipation of a test on the learning outcomes of 10th grade geology students. *Eurasia Journal of Mathematics, Science and Technology Education*, 12(5), 1379–1388. <https://doi.org/10.12973/eurasia.2016.1519a>
- Chen, C.-Y., Huang, H.-J., Lien, C.-J., & Lu, Y.-L. (2020). Effects of multi-genre digital game-based instruction on students' conceptual understanding, argumentation skills, and learning experiences. *IEEE Access*, 8, 110643–110655. <https://doi.org/10.1109/access.2020.3000659>
- Chen, C.-H., Liu, G.-Z., & Hwang, G.-J. (2016). Interaction between gaming and multistage guiding strategies on students' field trip mobile learning performance and motivation. *British Journal of Educational Technology*, 47(6), 1032–1050. <https://doi.org/10.1111/bjet.12270>
- Chen, C.-H., Shih, C.-C., & Law, V. (2020). The effects of competition in digital game-based learning (DGBL): A meta-analysis. *Educational Technology Research and Development*, 68(4), 1855–1873. <https://doi.org/10.1007/s11423-020-09794-1>
- Cheng, M.-T., Chen, J.-H., Chu, S.-J., & Chen, S.-Y. (2015). The use of serious games in science education: A review of selected empirical research from 2002 to 2013. *Journal of Computers in Education*, 2(3), 353–375. <https://doi.org/10.1007/s40692-015-0039-9>
- Cheung, M. W. L. (2014). Modeling dependent effect sizes with three-level meta-analyses: A structural equation modeling approach. *Psychological Methods*, 19(2), 211–229. <https://doi.org/10.1037/a0032968>
- Clark, D. B., Tanner-Smith, E. E., & Killingsworth, S. S. (2016). Digital games, design, and learning: A systematic review and meta-analysis. *Review of Educational Research*, 86(1), 79–122. <https://doi.org/10.3102/0034654315582065>
- Dou, L. (2019). A cold thinking on the development of STEM education. In *2019 international joint conference on information, media and engineering (IJCIME)* (pp. 88–92). <https://doi.org/10.1109/IJCIME49369.2019.00027>
- 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. <https://doi.org/10.1111/j.0006-341X.2000.00455.x>
- Fiorella, L., Kuhlmann, S., & Vogel-Walcutt, J. J. (2018). Effects of playing an educational math game that incorporates learning by teaching. *Journal of Educational Computing Research*, 57(6), 1495–1512. <https://doi.org/10.1177/0735633118797133>
- Freeman, B., & Higgins, K. (2016). A randomised controlled trial of a digital learning game in the context of a design-based research project. *International Journal of Technology Enhanced Learning*, 8(3–4), 297–317. <https://doi.org/10.1504/ijtel.2016.082316>
- Gao, F., Li, L., & Sun, Y.-Y. (2020). A systematic review of mobile game-based learning in STEM education. *Educational Technology Research and Development*, 68(4), 1791–1827. <https://doi.org/10.1007/s11423-020-09787-0>
- Gao, S., Assink, M., Cipriani, A., & Lin, K. (2017). Associations between rejection sensitivity and mental health outcomes: A meta-analytic review. *Clinical Psychology Review*, 57, 59–74. <https://doi.org/10.1016/j.cpr.2017.08.007>
- Gerber, S., & Scott, L. (2011). Gamers and gaming context: Relationships to critical thinking. *Journal of Educational Technology*, 42(5), 842–849. <https://doi.org/10.1111/j.1467-8535.2010.01106.x>
- Giannakos, M. N. (2013). Enjoy and learn with educational games: Examining factors affecting learning performance. *Computers & Education*, 68, 429–439. <https://doi.org/10.1016/j.compedu.2013.06.005>
- Glass, G. (1981). *Meta-analysis in social research*. Sage. [https://doi.org/10.1016/0149-7189\(82\)90011-8](https://doi.org/10.1016/0149-7189(82)90011-8)
- Grivokostopoulou, F., Kostas, K., & Perikos, I. (2020). The effectiveness of embodied pedagogical agents and their impact on students learning in virtual worlds. *Applied Sciences*, 10(5), 1739. <https://doi.org/10.3390/app10051739>
- Guo, Y. R., & Goh, D.H.-L. (2016). Evaluation of affective embodied agents in an information literacy game. *Computers & Education*, 103, 59–75. <https://doi.org/10.1016/j.compedu.2016.09.013>
- Halpern, D. F., Millis, K., Graesser, A. C., Butler, H., Forsyth, C., & Cai, Z. (2012). Operation ARA: A computerized learning game that teaches critical thinking and scientific reasoning. *Thinking Skills and Creativity*, 7(2), 93–100. <https://doi.org/10.1016/j.tsc.2012.03.006>
- Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, 6(2), 107–128. <https://doi.org/10.2307/1164588>
- Hodges, G. W., Wang, L., Lee, J., Cohen, A., & Jang, Y. (2018). An exploratory study of blending the virtual world and the laboratory experience in secondary chemistry classrooms. *Computers & Education*, 122, 179–193. <https://doi.org/10.1016/j.compedu.2018.03.003>
- Hooshyar, D., Malva, L., Yang, Y., Pedaste, M., Wang, M., & Lim, H. (2021). An adaptive educational computer game: Effects on students' knowledge and learning attitude in computational thinking. *Computers in Human Behavior*, 114, 106575. <https://doi.org/10.1016/j.chb.2020.106575>
- Hsiao, H.-S., Chang, C.-S., Lin, C.-Y., Chang, C.-C., & Chen, J.-C. (2014). The influence of collaborative learning games within different devices on student's learning performance and behaviours. *Australasian Journal of Educational Technology*, 30(6), 652–669. <https://doi.org/10.14742/ajet.347>
- Hsu, C.-C., & Wang, T.-I. (2018). Applying game mechanics and student-generated questions to an online puzzle-based game learning system to promote algorithmic thinking skills. *Computers & Education*, 121, 73–88. <https://doi.org/10.1016/j.compedu.2018.02.002>
- Huang, S.-Y., Kuo, Y.-H., & Chen, H.-C. (2020). Applying digital escape rooms infused with science teaching in elementary school: Learning performance, learning motivation, and problem-solving ability. *Thinking Skills and Creativity*, 37, 100681. <https://doi.org/10.1016/j.tsc.2020.100681>
- Hwang, G.-J., Yang, L.-H., & Wang, S.-Y. (2013). A concept map-embedded educational computer game for improving students' learning performance in natural science courses. *Computers & Education*, 69, 121–130. <https://doi.org/10.1016/j.compedu.2013.07.008>
- Ishak, S. A., Din, R., & Hasran, U. A. (2021). Defining digital game-based learning for science, technology, engineering, and mathematics: A new perspective on design and developmental research. *Journal of Medical Internet Research*, 23(2), e20537. <https://doi.org/10.2196/20537>
- Johnson, C. I., & Mayer, R. E. (2010). Applying the self-explanation principle to multimedia learning in a computer-based game-like environment. *Computers in Human Behavior*, 26(6), 1246–1252. <https://doi.org/10.1016/j.chb.2010.03.025>
- Jong, M. S. Y. (2015). Does online game-based learning work in formal education at school? A case study of VISOLE. *The Curriculum Journal*, 26(2), 249–267. <https://doi.org/10.1080/09585176.2015.1018915>
- Kao, G.Y.-M., Chiang, C.-H., & Sun, C.-T. (2017). Customizing scaffolds for game-based learning in physics: Impacts on knowledge acquisition and game design creativity. *Computers & Education*, 113, 294–312. <https://doi.org/10.1016/j.compedu.2017.05.022>
- Kayan-Fadllemula, F., Sellami, A., Abdelkader, N., & Umer, S. (2022). A systematic review of STEM education research in the GCC countries: Trends, gaps and barriers. *International Journal of STEM Education*, 9, 2. <https://doi.org/10.1186/s40594-021-00319-7>
- Kazimoglu, C. (2020). Enhancing confidence in using computational thinking skills via playing a serious game: A case study to increase motivation in learning computer programming. *IEEE Access*, 8, 221831–221851. <https://doi.org/10.1109/access.2020.3043278>
- Ke, F. (2015). Designing and integrating purposeful learning in game play: A systematic review. *Educational Technology Research and Development*, 64(2), 219–244. <https://doi.org/10.1007/s11423-015-9418-1>
- Kelley, T. R., & Knowles, J. G. (2016). A conceptual framework for integrated STEM education. *International Journal of STEM Education*, 3(1), 11. <https://doi.org/10.1186/s40594-016-0046-z>
- Khamparia, A., & Pandey, B. (2018). Effects of visual map embedded approach on students learning performance using Briggs-Myers learning style in word puzzle gaming course. *Computers & Electrical Engineering*, 66, 531–540. <https://doi.org/10.1016/j.compeleceng.2017.12.041>
- Klopfer, E., & Thompson, M. (2020). Game-based learning in science, technology, engineering, and mathematics. In J. L. Plass, R. E. Mayer, & B. D. Homer (Eds.), *Handbook of GAME-BASED LEARNING* (p. 387). MIT Press.
- Kukulska-Hulme, A., Bossu, C., Coughlan, T., Ferguson, R., FitzGerald, E., & Gaved, M. (2021). *Innovating pedagogy 2021: Open university innovation report 9*. The Open University.
- Kuppens, S., Laurent, L., Heyvaert, M., & Onghena, P. (2013). Associations between parental psychological control and relational aggression in children and adolescents: A multilevel and sequential meta-analysis.

- Developmental Psychology*, 49(9), 1697–1712. <https://doi.org/10.1037/a0030740>
- Lameras, P., Arnab, S., Dunwell, I., Stewart, C., Clarke, S., & Petridis, P. (2017). Essential features of serious games design in higher education: Linking learning attributes to game mechanics. *British Journal of Educational Technology*, 48(4), 972–994. <https://doi.org/10.1111/bjet.12467>
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press.
- Li, Y., Huang, Z., Jiang, M., & Chang, T.-W. (2016). The effect on pupils' science performance and problem-solving ability through lego: An engineering design-based modeling approach. *Educational Technology & Society*, 19(3), 143–156.
- Macaskill, P., Walter, S. D., & Irwig, L. (2001). A comparison of methods to detect publication bias in meta-analysis. *Statistics in Medicine*, 20(4), 641–654. <https://doi.org/10.1002/sim.698>
- Maheu-Cadotte, M.-A., Cossette, S., Dube, V., Fontaine, G., Mailhot, T., Lavoie, P., Cournoyer, A., Balli, F., & Mathieu-Dupuis, G. (2018). Effectiveness of serious games and impact of design elements on engagement and educational outcomes in healthcare professionals and students: A systematic review and meta analysis protocol. *British Medical Journal Open*, 8(3), e019871. <https://doi.org/10.1136/bmjopen-2017-019871>
- Mao, W., Cui, Y., Chiu, M. M., & Lei, H. (2022). Effects of game-based learning on students' critical thinking: A meta-analysis. *Journal of Educational Computing Research*, 59(8), 1682–1708. <https://doi.org/10.1177/07356331211007098>
- Martín-Páez, T., Aguilera, D., Perales-Palacios, F. J., & Vílchez-González, J. M. (2019). What are we talking about when we talk about STEM education? A review of literature. *Science Education*, 103(4), 799–822. <https://doi.org/10.1002/sce.21522>
- Masek, M., Boston, J., Lam, C. P., & Corcoran, S. (2017). Improving mastery of fractions by blending video games into the math classroom. *Journal of Computer Assisted Learning*, 33(5), 486–499. <https://doi.org/10.1111/jcal.12194>
- Mayer, R. E. (2014). *Computer games for learning: An evidence-based approach*. MIT Press.
- Mayer, R. E. (2015). On the need for research evidence to guide the design of computer games for learning. *Educational Psychologist*, 50(4), 349–353. <https://doi.org/10.1080/00461520.2015.1133307>
- Mayer, R. E. (2019). Computer games in education. *Annual Review of Psychology*, 70(1), 531–549. <https://doi.org/10.1146/annurev-psych-010418-102744>
- Mayer, R. E. (2020). Cognitive foundations of game-based learning. In J. L. Plass, R. E. Mayer, & B. D. Homer (Eds.), *Handbook of game-based learning* (pp. 83–110). The MIT Press.
- National Research Council. (2011). *Successful K-12 STEM education: Identifying effective approaches in science, technology, engineering, and mathematics*. National Academies Press.
- Nelson, B. C., Kim, Y., Foshee, C., & Slack, K. (2014). Visual signaling in virtual world-based assessments: The SAVE science project. *Information Sciences*, 264, 32–40. <https://doi.org/10.1016/j.ins.2013.09.011>
- Okutsu, M., DeLaurentis, D., Brophy, S., & Lambert, J. (2013). Teaching an aerospace engineering design course via virtual worlds: A comparative assessment of learning outcomes. *Computers & Education*, 60(1), 288–298. <https://doi.org/10.1016/j.compedu.2012.07.012>
- Plass, J. L., Homer, B. D., & Kinzer, C. K. (2015). Foundations of game-based learning. *Educational Psychologist*, 50(4), 258–283. <https://doi.org/10.1080/00461520.2015.1122533>
- Plass, J. L., Mayer, R. E., & Homer, B. D. (2020). *Handbook of game-based learning*. MIT Press.
- Poppenk, J., Kohler, S., & Moscovitch, M. (2010). Revisiting the novelty effect: When familiarity, not novelty, enhances memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36(5), 1321–1330. <https://doi.org/10.1037/a0019900>
- Proulx, J.-N., Romero, M., & Arnab, S. (2017). Learning mechanics and game mechanics under the perspective of self-determination theory to foster motivation in digital game based learning. *Simulation & Gaming*, 48(1), 81–97. <https://doi.org/10.1177/1046878116674399>
- R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing Vienna. <https://www.R-project.org/>.
- Raman, R., Lal, A., & Achuthan, K. (2014). Serious games based approach to cyber security concept learning: Indian context. In *2014 international conference on green computing communication and electrical engineering (ICGCCCE)* (pp. 1–5), Coimbatore, India. <https://doi.org/10.1109/icgcccce.2014.6921392>
- Reed, D. A., Cook, D. A., Beckman, T. J., Levine, R. B., Kern, D. E., & Wright, S. M. (2007). Association between funding and quality of published medical education research. *JAMA—Journal of the American Medical Association*, 298(9), 1002–1009. <https://doi.org/10.1001/jama.298.9.1002>
- Renken, M. D., & Nunez, N. (2013). Computer simulations and clear observations do not guarantee conceptual understanding. *Learning and Instruction*, 23, 10–23. <https://doi.org/10.1016/j.learninstruc.2012.08.006>
- Riopel, M., Nenciovici, L., Potvin, P., Chastenay, P., Charland, P., Sarrasin, J. B., & Masson, S. (2019). Impact of serious games on science learning achievement compared with more conventional instruction: An overview and a meta-analysis. *Studies in Science Education*, 55(2), 169–214. <https://doi.org/10.1080/03057267.2019.1722420>
- Romero, M., Usart, M., & Ott, M. (2015). Can serious games contribute to developing and sustaining 21st century skills? *Games and Culture*, 10(2), 148–177. <https://doi.org/10.1177/1555412014548919>
- Roodt, S., & Ryklief, Y. (2019). Using digital game-based learning to improve the academic efficiency of vocational education students. *International Journal of Game-Based Learning*, 9(4), 45–69. <https://doi.org/10.4018/ijgbl.2019100104>
- Sadler, T. D., Romine, W. L., Menon, D., Ferdig, R. E., & Annetta, L. (2015). Learning biology through innovative curricula: A comparison of game- and nongame-based approaches. *Science Education*, 99(4), 696–720. <https://doi.org/10.1002/sce.21171>
- Sarvehana, Z. S. J. (2019). Study on effectiveness of Hana game application on cognitive problem solving skill, attention and academic achievement linguistics and mathematics on first grade student. In *2019 international serious games symposium* (pp. 109–118). IEEE. <https://doi.org/10.1109/ISGS49501.2019.9047023>
- Schrader, C., & Bastiaens, T. (2012). Learning in educational computer games for novices: The impact of support provision types on virtual presence, cognitive load, and learning outcomes. *International Review of Research in Open and Distributed Learning*, 13(3), 206–227. <https://doi.org/10.19173/irrodl.v13i3.1166>
- Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014). P-curve: A key to the file-drawer. *Journal of Experimental Psychology-General*, 143(2), 534–547. <https://doi.org/10.1037/a0033242>
- Smith, R. P., & Learman, L. A. (2017). A plea for MERSQI the medical education research study quality instrument. *Obstetrics and Gynecology*, 130(4), 686–690. <https://doi.org/10.1097/aog.0000000000002091>
- Soflano, M., Connolly, T. M., & Hainey, T. (2015). An application of adaptive games-based learning based on learning style to teach SQL. *Computers & Education*, 86, 192–211. <https://doi.org/10.1016/j.compedu.2015.03.015>
- Struyf, A., De Loof, H., Boeve-de Pauw, J., & Van Petegem, P. (2019). Students' engagement in different STEM learning environments: Integrated STEM education as promising practice? *International Journal of Science Education*, 41(10), 1387–1407. <https://doi.org/10.1080/09500693.2019.1607983>
- Sung, H.-Y., & Hwang, G.-J. (2013). A collaborative game-based learning approach to improving students' learning performance in science courses. *Computers & Education*, 63, 43–51. <https://doi.org/10.1016/j.compedu.2012.11.019>
- Sung, H.-Y., Hwang, G.-J., Wu, P.-H., & Lin, D.-Q. (2018). Facilitating deep-strategy behaviors and positive learning performances in science inquiry activities with a 3D experiential gaming approach. *Interactive Learning Environments*, 26(8), 1053–1073. <https://doi.org/10.1080/10494820.2018.1437049>
- Talan, T., Dogan, Y., & Batdi, V. (2020). Efficiency of digital and non-digital educational games: A comparative meta-analysis and a meta-thematic analysis. *Journal of Research on Technology in Education*, 52(4), 474–514. <https://doi.org/10.1080/15391523.2020.1743798>
- The White House. (2018). *Charting a course for success: America's strategy for STEM education*. <https://www.whitehouse.gov/wp-content/uploads/2018/12/STEM-Education-Strategic-Plan-2018.pdf>
- Tokac, U., Novak, E., & Thompson, C. G. (2019). Effects of game-based learning on students' mathematics achievement: A meta-analysis. *Journal of Computer Assisted Learning*, 35(3), 407–420. <https://doi.org/10.1111/jcal.12347>
- Tsai, Y.-L., & Tsai, C.-C. (2018). Digital game-based second-language vocabulary learning and conditions of research designs: A meta-analysis study.

- Computers & Education*, 125, 345–357. <https://doi.org/10.1016/j.compedu.2018.06.020>
- Tsai, Y.-L., & Tsai, C.-C. (2020). A meta-analysis of research on digital game-based science learning. *Journal of Computer Assisted Learning*, 36(3), 280–294. <https://doi.org/10.1111/jcal.12430>
- Vanbecelaere, S., Van den Berghe, K., Cornillie, F., Sasanguie, D., Reynvoet, B., & Depaepe, F. (2020). The effects of two digital educational games on cognitive and non-cognitive math and reading outcomes. *Computers & Education*, 143, 103680. <https://doi.org/10.1016/j.compedu.2019.103680>
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, 36(3), 1–48. <https://doi.org/10.18637/jss.v036.i03>
- Viechtbauer, W., & Cheung, M. W. (2010). Outlier and influence diagnostics for meta-analysis. *Research Synthesis Methods*, 1(2), 112–125. <https://doi.org/10.1002/jrsm.11>
- Wahono, B., Lin, P.-L., & Chang, C.-Y. (2020). Evidence of STEM enactment effectiveness in Asian student learning outcomes. *International Journal of STEM Education*, 7(1), 1–18. <https://doi.org/10.1186/s40594-020-00236-1>
- Wang, L.-H., Chen, B., Hwang, G.-J., Guan, J.-Q., & Wang, Y.-Q. (2022). Effects of digital game-based STEM education on students' learning achievement: A meta-analysis. *International Journal of STEM Education*, 9(1), 1–13. <https://doi.org/10.1186/s40594-022-00344-0>
- Wang, Y.-H. (2020). Exploring the effects of designing a role-playing game with single and peer mode for campus learning. *Educational Technology Research and Development*, 68(3), 1275–1299. <https://doi.org/10.1007/s11423-019-09726-8>
- Weng, H. J., Chang, D. F., & Shyu, H. Y. (2015). Testing the effect of mnemonic strategy embedded in digital game. *JCIC Express Letters*, 9(3), 827–833.
- Wilson, K. A., Bedwell, W. L., Lazzara, E. H., Salas, E., Burke, C. S., Estock, J. L., Orvis, K. L., & Conkey, C. (2009). Relationships between game attributes and learning outcomes review and research proposals. *Simulation & Gaming*, 40(2), 217–266. <https://doi.org/10.1177/1046878108321866>
- Wouters, P., van der Spek, E. D., & van Oostendorp, H. (2009). Current practices in serious game research: a review from a learning outcomes perspective. *Games-based learning advancements for multi-sensory human computer interfaces* (pp. 232–250). IGI Publishing. <https://doi.org/10.4018/978-1-60566-360-9.ch014>
- Wouters, P., van Nimwegen, C., van Oostendorp, H., & van der Spek, E. (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>
- Wouters, P., & van Oostendorp, H. (2013). A meta-analytic review of the role of instructional support in game-based learning. *Computers & Education*, 60(1), 412–425. <https://doi.org/10.1016/j.compedu.2012.07.018>
- Wu, Y.-T., & Anderson, O. R. (2015). Technology-enhanced stem (science, technology, engineering, and mathematics) education. *Journal of Computers in Education*, 2(3), 245–249. <https://doi.org/10.1007/s40692-015-0041-2>
- Yu, J., Denham, A. R., & Searight, E. (2022). A systematic review of augmented reality game-based learning in STEM education. *Educational Technology Research and Development*, 70(4), 1169–1194. <https://doi.org/10.1007/s11423-022-10122-y>
- Ziegler, A., Antes, G., & Koenig, I. R. (2011). Preferred report items for systematic reviews and meta-analysis: The PRISMA statement. *Deutsche Medizinische Wochenschrift*, 136(8), E9–E15. <https://doi.org/10.1055/s-0031-1272978>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

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

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

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
