The Effect of Digital Game-Based Learning Interventions on Cognitive, Metacognitive, and Affective-Motivational Learning Outcomes in School: A Meta-Analysis

Article *in* Review of Educational Research · May 2023

DOI: 10.3102/00346543231167795

4 authors:

Nathalie Barz
Saarland University

Laura Dörrenbächer-Ulrich
Saarland University

Manuela Benick
Saarland University

Franziska Perels
Saarland University

The full article can be retrieved with the following citation:

Barz, N., Benick, M., Dörrenbächer-Ulrich, L., & Perels, F. (2023). The Effect of Digital Game-Based Learning Interventions on Cognitive, Metacognitive, and Affective-Motivational Learning Outcomes in School: A Meta-Analysis. Review of Educational Research, 0(0). Copyright © 2023 SAGE. https://doi.org/10.3102/00346543231167795

The Effect of Digital Game-Based Learning Interventions on Cognitive, Metacognitive, and Affective-Motivational Learning Outcomes in School: A Meta-Analysis.

Abstract

Digital game-based learning (DGBL) interventions can be superior to traditional instruction methods for learning, but previous meta-analyses covered a huge period and included a variety of different target groups, limiting the results' transfer on specific target groups. Therefore, the aim of this meta-analysis is a theory-based examination of DGBL interventions' effects on different learning outcomes (cognitive, metacognitive, affective-motivational) in the school context, using studies published between 2015 and 2020 and meta-analytic techniques (including moderator analyses) to examine the effectiveness of DGBL interventions compared to traditional instruction methods. Results from random-effects models revealed a significant medium effect for overall learning (g = .54) and cognitive learning outcomes (g = .67). Also found were a small effect for affective-motivational learning outcomes (g = .32) and no significant effect for metacognitive learning outcomes. Additionally, there was no evidence of publication bias. Further meta-regression models did not reveal evidence of moderating personal, environmental, or confounding factors. The findings partially support the positive impact of DGBL interventions in school, and the study addresses its practical implications.

Keywords: game-based learning, learning outcomes, meta-analysis, school context, systematic review

The Effect of Digital Game-Based Learning Interventions on Cognitive, Metacognitive, and Affective-Motivational Learning Outcomes in School: A Meta-Analysis

Due to the COVID-19 pandemic, learning with digital media is more relevant than ever. But learners may experience mere e-learning environments as boring because those environments often lack engaging elements (Connolly & Stansfield, 2009). One solution could be the application of digital game-based learning (DGBL), because games can be highly motivating (Chang et al., 2017).

Prensky (2007) strongly influenced the term 'digital game-based learning', defining it as 'any marriage of educational content and computer games' (p. 145). Over time, DGBL definitions focused more on the aim to promote teaching and learning processes. Al-Azawi et al. (2016) argue that we can speak of DGBL when digital games are specifically 'designed and used for teaching and learning' (p. 132). According to Malliarakis et al. (2018), the goal of DGBL is to learn during gameplay and promote desired learning outcomes, which also summon up the terms 'serious game' or 'game with a purpose'. Another characteristic of DGBL that Cojocariu and Boghian (2014) highlighted is the connection of educational content with new learning technologies, such as mobile devices (e.g., tablets, smartphones). Moreover, the use of new learning technologies could foster cognitive change and enrich the learning process with entertainment, leading to improved learning (Khan et al., 2017). The extensive definition by Erhel and Jamet (2013) summarized the previous views of DGBL, defining it as:

> a competitive activity in which students are set educational goals intended to promote knowledge acquisition. The games may either be designed to promote learning or the development of cognitive skills, or else take the form of simulations allowing learners to practice their skills in a virtual environment. (p. 156)

This definition combines the digital environments and the objective of promoting learning processes or specific skills, but also encompasses simulations as DGBL interventions, providing the basis for the present meta-analysis.

A growing body of research indicates DGBL interventions' positive effects in comparison with traditional instruction methods. Increased learning gains (e.g., higher test scores) have appeared after DGBL interventions, in such different domains as language learning (Franciosi, 2017) or STEM education (McLaren et al., 2017) in comparison with traditional instruction methods.

The comparison with traditional instruction methods is of relevance to determine the advantages of DGBL interventions and deduce recommendations for scientific and educational practice. Digital games offer an innovative approach to educate pupils in the classroom because well-designed games can adapt to the learners' needs (Plass & Pawar, 2020) and therefore, can contribute to the successful handling of heterogeneity in the classroom. Furthermore, they provide a secure learning environment where pupils are allowed to make mistakes and try again. This 'graceful failure' (Plass et al., 2015, p. 261) leads to mastery experiences, which increase pupils' self-efficacy and motivation. Digital games support a learner-centered pedagogy and enable pupils to actively make their own learning experiences within the game.

According to the self-determination theory (Ryan & Deci, 2000), implementing game mechanics to satisfy the basic psychological needs for competence, autonomy, and relatedness (Ryan & Rigby, 2020) can make digital games highly motivating. Chang et al. (2017) conducted a study with 103 university students in an education course, comparing game-based learning and traditional learning methods regarding flow experiences. The experimental group played a digital game about the carbon footprint; the control condition was learning with webpages. The results show that participants who played the game had a significantly stronger experience of flow during the experiment, represented through higher enjoyment, engagement, and control.

As a group, that Admiraal et al. (2011) described as 'bored and disengaged' (p. 1185), pupils could especially benefit from motivation through DGBL interventions, and evidence exists for DGBL interventions' positive effect on them (López-Fernández et al., 2021). In an Iranian elementary school, Partovi & Razavi (2019) examined the influence of DGBL on pupils' academic achievement motivation. After three months of learning with a digital game, the experimental group had significantly higher academic achievement motivation than the control group.

These findings support the assumption that DGBL interventions could benefit pupils, but they just capture a snapshot of learning with digital games. Previous meta-analyses show that digital games outmatch traditional instruction methods regarding learning gains for young children and young adults (Clark et al., 2016). In the context of elementary school to university, there is also evidence for DGBL interventions' positive impact on cognitive, affective, and behavioral outcomes (Lamb et al., 2018). Prior meta-analyses agree that DGBL can have positive effects on learning, but they only included studies until 2015. In the last five years, which the current meta-analysis includes, rapid technological development occurred, offering new technological opportunities for DGBL interventions. Games can be developed at lower costs and are more accessible for research and teachers. Due to the ongoing digital transformation, the number of digital devices in school has increased. Those digital devices provide more memory space, better graphic boards, and main storage, which facilitated the implementation of DGBL interventions in school in the recent years. In game development, more and more complex game mechanics and detailed textures can be applied. This makes it hard to compare modern DGBL interventions with DGBL interventions from ten years ago. Therefore, there is renewed demand to analyze the impact of DGBL interventions in the last years. To make DGBL interventions comparable and to merge with the latest meta-analyses, only studies between the years 2015 and 2020 are analyzed in the current meta-analysis. Moreover, the prior meta-analyses examined a wide range of target groups, making transferring the findings to special subgroups difficult. As mentioned, pupils could greatly benefit from DGBL interventions, which is why they are the focus of this meta-analysis. Accordingly, the objective of this meta-analysis is to analyze the effect of DGBL interventions in comparison with that of traditional instruction methods within the years 2015–2020 purely in the school context. To specify the impact of DGBL interventions, a theoretical model, namely the *Integrated Design Framework for Playful Learning* (Plass et al., 2015), supported deducing different categories of learning outcomes.

Integrated Design Framework for Playful Learning

As shown in Figure 1, the *Integrated Design Framework for Playful Learning* (Plass et al., 2015) combines different theoretical foundations of gameplay that lead to specific game-design elements. The model describes four areas of theoretical foundations: affective, motivational, cognitive, and sociocultural foundations.

Affective foundations of DGBL include the evocation of emotions to promote learning. The so-called emotional design deals with the question of how to induce emotion with different game-design elements, such as narrative or music (Plass et al., 2015). Games can evoke various emotions with, for example, empathic characters that increase cognitive processing and thus, facilitate learning processes (Plass et al., 2020). One emotional foundation of DGBL represents the integrative model by Loderer et al. (2020), which is based on Pekrun's (2006) content-value theory of achievement emotions. It assumes that emotions depend on the evaluation of internal (e.g., value) and external (e.g., musical score) stimuli and the emotional transmission from other people or game elements (e.g., peers, visual aesthetic; Loderer et al., 2020).

Motivational foundations of DGBL emphasize the characteristics of games to motivate or engage players to play for enjoyment. Ryan and Rigby (2020) refer to motivation as the 'core element in game-based learning' (p. 153). One motivation theory they include is the expectancy-value theory (Wigfield & Eccles, 2000), which assumes that the expectation of a benefit from their actions motivates players. In addition, self-determination theory (Ryan & Deci, 2000) focuses on intrinsic motivation, which means that a person pursues something because of interest or fulfillment resulting from the action itself. According to this theory, intrinsic motivation depends on three different psychological needs, namely competence, autonomy, and relatedness. Competence describes the feeling of effectiveness and mastery of a task (Ryan & Deci, 2000). In DGBL, feedback or level-ups satisfy the need for competence because players experience growth and self-efficacy, 'people's beliefs about their capabilities to exercise control over events that affect their lives' (Bandura, 1989, p. 1175). Digital games offer the possibility for a retry after a failure, and they enable 'graceful failure' (Plass et al., 2015, p. 261), which positively affects self-efficacy. Moreover, games can be adapted to the players' skills or abilities, and therefore, lead to mastery experiences that can also increase self-efficacy (Bandura, 1997). The need for autonomy refers to the desire for self-determined and volitional activities. Digital games provide virtual worlds with many possibilities and the freedom to personalize and choose activities that increase satisfaction of autonomy needs (Rigby & Ryan, 2011). The last factor, relatedness, includes the need to help others, collaborating to reach common goals. Multiplayer games or adoption of special team roles to contribute to joint goal attainment could meet the demand for relatedness.

Cognitive foundations of DGBL refer to cognitive characteristics for learning—for example, information processing and explaining how players could learn from digital games. The cognitive theory of game-based learning combines the cognitive load theory (Sweller, 2011) as well as the cognitive theory of multimedia learning (Mayer, 2014) and makes three assumptions: 1) Players use two separate channels to process visual and verbal information which interact with each other. 2) Players have a limited capacity to process information in the respective channels. 3) Players must actively process information by paying attention, structuring information, and integrating new knowledge to be able to learn (Mayer, 2014). Games may contain pictures or verbal narrations that players receive through eyes or ears. The working memory further processes the filtered visual or verbal information into verbal and pictorial models. Prior knowledge enriches created verbal and visual representations, and long-term memory stores them (Mayer, 2020). Information processing can occur in three different ways. Extraneous processing includes cognitive processing unrelated to the learning goal, which is caused by distraction from extraneous stimuli (e.g., unnecessary background animations). Essential processing refers to the cognitive processing that assimilating information in the working memory requires, and generative processing manages the organization and integration of information in long-term memory (Mayer, 2020).

Sociocultural foundations of DGBL take account of 'interactions among players, the construction of collective knowledge, and the application of this knowledge in the context of cultural norms' (Plass et al., 2020, p. 17). Games enable interactions and relationships with other players and provide a connection to the game's community, which could exceed the game itself (Steinkuehler & Tsaasan, 2020). Social context can promote learning, for example, by increasing motivation due to the feeling of belonging to the community (Plass et al., 2015). Social interactions, like collaborative or competitive play, could influence how much effort a player invests in a game and, therefore, represents an important factor of DGBL (Plass et al., 2020).

To summarize, different perspectives on digital games are represented in their different theoretical foundations and the selected game-design elements depend on the chosen theoretical foundation and the intended learning goals. To assess the broad variety of learning goals, the present meta-analysis examines different learning outcomes based on this model and extends them to the category of metacognitive outcomes.

Different game-design elements for digital game-based learning are suitable to promoting specific learning goals and aspiring to 'achieve the intended interactions with the learning content in a playful, motivating way' (Plass et al., 2020, p. 11). Game-design elements in the model are incentive systems, aesthetic design, game mechanics, narrative, and sound design, which depend on the content and skills that the game should convey. The incentive system offers rewards to provide feedback and influence player behavior (Kinzer et al., 2012). Incentives could trigger extrinsic motivation with extrinsic rewards that are not necessary for the core gameplay (e.g., experience points) or increase intrinsic motivation with extrinsic rewards that could result in intrinsic motivation (e.g., a new tool for exploration; Tam & Pawar, 2020). Aesthetic design includes the visual design of the game, the avatars, and characters, as well as the visual signaling to provide cues and feedback.

Game mechanics describe repeated activities the player performs to enable gameplay. They comprise two categories: assessment and learning mechanics. Assessment mechanics, based on test theory approaches, contain diagnostical goals (e.g., apply rules to solve a problem; Plass et al., 2011), whereas learning mechanics are based on learning theories and contain learning goals (e.g., acquire knowledge through interaction with other characters; Plass et al., 2020). A narrative contains the storyline or dialogue with characters. It could provide a context for learning processes and often connects different game elements with each other (Dickey, 2020). The game's sound design represents all auditory stimuli (e.g., character sounds, action sounds) and could support cueing (Pawar et al., 2020). The implemented game-design elements could induce different types of engagement, including cognitive, behavioral, affective, and social engagement (Domagk et al., 2010). For example, using a narrative that captivates the player could achieve an affective engagement that causes the player to finish the game.

Based on the presented model, we deduced different learning outcomes to represent the engagement types that DGBL interventions could evoke. The next chapter describes the analyzed learning outcomes in detail.

Definition of Learning Outcomes

To enable a more precise insight into the effectiveness of digital game-based learning interventions, the present meta-analysis examines learning outcomes in specific categories. The previous theoretical model classifies learning outcomes as cognitive and affective-motivational, reflecting the types of engagement that different game elements could induce. DGBL interventions often neglect behavioral and social engagement because their main goal is to acquire knowledge or motivate people to learn. Therefore, this meta-analysis does not include behavioral and social outcomes. However, the model disregards another important category of learning outcomes, namely, metacognitive learning outcomes, which the present meta-analysis considers as an extension of the theoretical model. Cognitive learning outcomes contain conceptual or domain-specific knowledge and the ability to remember, understand and recall this knowledge (Post et al., 2019). Affective-motivational learning outcomes include attitudes, beliefs, emotions, values, and interests (Allen & Friedman, 2010). Metacognitive learning outcomes comprise two components: metacognitive knowledge and metacognitive skills. Metacognitive knowledge includes declarative (e.g., what learning strategy would be appropriate) and procedural (e.g., how to use a learning strategy) knowledge of one's own knowledge and describes higher order knowledge, whereas metacognitive skills involve planning, monitoring, and reflection strategies to regulate learning processes (Veenman et al., 2006).

Evidence From Current Meta-Analyses

Investigating the influence of digital game-based learning is a relevant topic with which meta-analyses have dealt before. Vogel et al. (2006) compared the effects of DGBL interventions, —namely, serious games and simulations—regarding cognitive gains with traditional teaching methods, between 1986 and 2003. Their analyses included 32 studies with different target groups and revealed that serious games and simulations lead to higher cognitive gains than traditional methods. Sitzmann (2011) analyzed the effect of simulation games on trainees in studies from 1976 to 2009. As dependent variables, self-efficacy, declarative and procedural knowledge were analyzed in a sample of 65 studies. The results showed that the participants who used simulation games for learning had higher procedural and declarative knowledge as well as better knowledge recall than participants utilizing traditional learning. Furthermore, in eight studies, self-efficacy was 20% higher when learning with a simulation game in comparison to the control group.

In 2013, Wouters et al. examined the effectiveness of DGBL interventions, assessed by using serious games for knowledge acquisition and learning retention, between the years 1990 and 2012. They did not set an age restriction for the studies and consequently examined a sample of 39 studies with a broad target population. Their findings indicate greater learning gains for serious game conditions compared to traditional instruction methods (d = .29), but they also reveal no difference in motivation between the two conditions. Serious games seem no more motivating than traditional methods.

Clark et al. (2016) investigated the effect of DGBL interventions on learning, assessed through the use of digital games in the age group from 6 to 25 years, by synthesizing studies between the years 2000 and 2012. They included 69 studies in their analysis and, in accordance with the results of previous meta-analyses, they found evidence that digital games outmatch control groups regarding learning.

Another meta-analysis that Lamb et al. (2018) conducted focused on the effect of simulations and serious games on students' cognition, affect and learning from elementary school to university. They analyzed 46 studies from the period 2002–2015. The results of their meta-analysis indicate a medium effect for cognitive and affective outcomes regarding learning and a small effect for behavioral outcomes for participants using simulation games and serious games.

All these meta-analyses share the conclusion that DGBL has positive effects on learning, but they examined studies covering a huge period and, therefore, do not do justice to the rapid technological development and new technological opportunities. Furthermore, they focused on very heterogeneous target groups, making it hard to transfer the findings to specific target groups. Therefore, the present meta-analysis examines studies that cover a shorter and more up-to-date period and focuses on a more specific target group, namely pupils, because they are predestined to benefit from DGBL interventions due to their motivational needs. Additionally, to extend the previous meta-analyses, all learning outcomes derive from a theoretical model, considering metacognitive learning outcomes in the analysis that have not been examined before. Based on the previous findings, we assume positive effects on learning in general (Hypothesis 1.1) as well as on cognitive (Hypothesis 1.2), metacognitive (Hypothesis 1.3) and affective-motivational learning outcomes (Hypothesis 1.4).

Impact of the Moderating Role of Personal, Environmental, and Confounding Factors

In addition to the examination of the main effects of DGBL on different learning outcomes, this meta-analysis also scrutinized the impact of personal factors, learning environment factors, and confounding factors to pursue the aim to deduce recommendations for scientific and educational practice how to design and implement digital games. The variables, which were considered in the different categories, were deduced from previous meta-analyses on the one hand and based on explorative considerations on the other hand. The following sections describe the moderator variables.

Personal Factors

Age. For age as a moderator in DGBL interventions, findings show no differences between the age groups from preschool to college students (Vogel et al., 2006). The comparison of different age groups (children, preparatory education, students, adults) revealed also no differences regarding DGBL interventions in the meta-analysis by Wouters and colleagues (2013). DGBL interventions led to better learning gains in all age groups (except adults) when compared to the control group. We also assume that DGBL interventions are effective for all participants, and no differences between the age groups would appear in the present meta-analysis (Hypothesis 2.1).

Gender. Vogel et al. (2006) found evidence for gender differences regarding cognitive gains for DGBL interventions. Female participants 'showed significant cognitive gains favoring the interactive simulation and game method' (p. 234). We also assume, due to playing behaviors that differ from males' (Veltri et al., 2014) that DGBL interventions are more effective for female players (Hypothesis 2.2).

Learning Environment Factors

Additional Non-Game Instruction. Many studies use additional non-game instruction (e.g., discussion, preparative lessons) but the findings are mixed regarding additional non-game instruction as a moderator for DGBL. Whereas Sitzmann (2011) and Wouters et al. (2013) found evidence that DGBL interventions combined with additional non-game instruction led to greater learning gains, Clark et al. (2016) did not find differences between DGBL interventions with and without additional non-game instruction. Therefore, additional non-game instruction is further analyzed exploratively in the current meta-analysis (Research question 1).

Type of DGBL Intervention. Regarding the type of DGBL intervention, using a serious game or an interactive simulation made no difference; both were likewise superior to traditional methods (Vogel et al., 2006). The results from Lamb et al. (2018) indicated differences between serious games and simulations, in that the former had a greater effect on learning than simulations. With the unclear impact of the DGBL intervention type, this variable is exploratively examined as a moderator (Research question 2).

Number of Sessions. The assumption that one game session may not be sufficient to increase learning was based on the well-investigated advantage of distributed learning over massed learning (Cepeda et al., 2006). Wouters et al. (2013) substantiated this hypothesis for serious games, reporting in their meta-analysis that serious games with multiple sessions led to higher learning gains than traditional learning methods. They explained this result by stating that participants need time to understand the games' control. Clark et al. (2016) found further evidence of the superiority of multiple game sessions over single sessions. The effects of digital games on learning were smaller when games were only played once rather than in distributed sessions. Therefore, we include the number of sessions in the meta-analysis and assume the superiority of multiple game sessions over single game sessions in DGBL (Hypothesis 2.3).

Playing Mode. The impact of the playing mode in DGBL interventions is still unclear. Vogel et al. (2006) revealed that both, single players and players in groups, outperformed the control group in terms of learning, but they did not further analyze the difference between the two playing modes. Wouters et al. (2013) found evidence for greater learning gains favoring gaming in groups. Participants in single-player and multiplayer games learned more than the control group, but the effect was larger for multiplayer games. Clark et al. (2016) also analyzed player modes and first reported the largest learning outcomes for single players. However, this effect disappeared when other game characteristics (e.g., story relevance, visual realism) were controlled. Single-player games then showed no significant difference from multiplayer games. As a result of these inconsistent findings, this study exploratively analyzed the influence of player mode in DGBL interventions (Research question 3).

Competition. Closely associated with the playing mode are competitive game designs. In their examination of player modes, Clark et al. (2016) also considered competitive and noncompetitive single and multiplayer configurations. According to their findings, single-player games without competition did not lead to larger effects than team competition games, but both noncompetitive single-player and competitive multiplayer games led to greater learning outcomes than competitive single-player games. Based on this evidence, the effect of the mere presence of competition in games remains unclear because it was confounded with playing mode. But Clark et al. (2016) provided the first insights onto possible differences between noncompetitive and competitive game designs. Evidence exists that competitive game designs could lead to higher intrinsic motivation and therefore result in greater learning outcomes (Cagiltay et al., 2015). Chen and Chiu (2016) argue that teamwork leads to an effective exchange of information and the creation of new ideas. Competitive teams could increase interest as well as engagement and decrease pressure to perform (Chen, 2019), making them superior to noncompetitive learning. In their meta-analysis concerning competition in DGBL environments, Chen et al. (2020) found a positive effect of competitive DGBL environments, dependent on subject and game type. Thus, we assume that competitive DGBL interventions lead to greater learning gains than noncompetitive DGBL interventions (Hypothesis 2.4).

Dimensionality. 2D and 3D games are effective for learning (Ak & Kutlu, 2017), but the dimensionality of game environments could also influence the learning effect. Three-dimensional games lead to the greatest effect on learning, according to the findings from Lamb et al. (2018), outmatching two-dimensional and mixed environments regarding cognition and affect. Three-dimensional learning environments seem to lead to greater cognitive and affective activation and facilitate the transfer of the learning content to the real world, which is also three-dimensional (Lamb et al., 2018). We assume that three-dimensional DGBL interventions lead to greater learning gains than two-dimensional or mixed DGBL interventions (Hypothesis 2.5).

Visual Realism. Not only the dimensionality but also the level of visual realism could influence learning with DGBL interventions. The impact of visual realism is still unclear, given the inconsistent findings for this variable. Vogel et al. (2006) concluded in their meta-analysis that the level of realism has no influence on learning, whereas Wouters et al. (2013) found evidence that schematic serious games are more effective than traditional instruction methods. Realistic or cartoonlike serious games did not lead to better learning than traditional instruction methods. The results from Clark et al. (2016), who also described schematic games as superior to realistic games, support these findings. Controlled for visual and narrative game characteristics, the difference between schematic and realistic games was diminished to marginal significance. Based on the previous studies and taking into account that highly realistic learning environments could lead to a higher cognitive load and, therefore, more complex information processing (Mayer, 2014; Nelson & Kim, 2020), we assume that schematic DGBL interventions lead to greater learning gains than realistic or cartoonlike DGBL interventions (Hypothesis 2.6).

Learning Domain. Another interesting factor to consider regarding DGBL interventions is the learning domain. Wouters et al. (2013) described DGBL interventions as exceptionally more effective in verbal domains than in scientific or mathematical domains. On the contrary, Karakoç et al. (2020) found no differences regarding learning domains, but their findings are restricted to Turkish publications and difficult to generalize. Due to the unclear impact of different learning domains, this moderator was analyzed exploratively (Research question 4).

Narration. Whether to include a narrative in a game is a controversial topic. One party supports the integration of a strong narration into game environments because it causes immersion and greater intrinsic motivation. Narrations can evoke emotions and promote the relatedness with the game by solving a common problem (Dickey, 2020; Rowe et al., 2011). The other side argues that a strong narrative distracts learners from the actual learning content, because of higher extraneous cognitive load that decreases learning (Novak, 2015). In their meta-analysis, Wouters et al. (2013) could not find evidence for a learning difference between games with or without a narrative, but both were superior to traditional instruction methods. Clark et al. (2016) compared relevant to irrelevant narratives and found significantly larger effects for games that used irrelevant narratives, but when controlled for visual characteristics and story depth, this difference vanished. They also found a negative relationship between narratives and learning, but their results must be handled with caution because they only included a small number of studies with relevant narratives ($n = 5$). There is strong evidence that narratives support problem-solving in games (Lester et al., 2014) and are effective for learning (Jackson et al., 2018). Furthermore, they can foster emotional engagement (Bowers et al., 2013) that could also support deeper information processing. Based on these explanations, we assume that DGBL interventions with a narrative lead to greater learning gains than DGBL interventions without a narrative (Hypothesis 2.7).

Avatar. Avatars, which represent the players, could also influence the learning process. Players identify with their avatar, which causes an emotional bond and increases motivation (van Reijmersdal et al., 2013). The identification means that the player 'imagines being that character and replaces his/her personal identity […] with the role of the character' (Cohen, 2001, p. 251). Thus, the player merges with the avatar during gameplay. This can occur because the avatar is perceived as similar to oneself or because the avatar possesses characteristics the player wishes for (Klimmt et al., 2010). Hence, the identification with the avatar could cause the intrinsic desire to continue with the game, which could lead to greater learning gains in the context of educational games (Tam & Pawar, 2020). Wrzesien et al. (2015) also found that avatars that resemble the player result in better achievement in emotional strategy learning. Therefore, the mere presence of an avatar could make a difference in the effect on learning in DGBL interventions (Hypothesis 2.8).

Digital Agent. Whereas avatars represent characters the player controls, digital agents are so-called nonplayer characters (NPCs) that the player does not control. They can 'create rich face-to-face interactions' (Johnson & Lester, 2016, p. 26) and offer scaffolding and support (e.g., hints, feedback). Digital agents also play a socio-emotional role, for example, as a peer, which could improve learning motivation (Lester et al., 2020). A digital agent significantly improved girls' motivation and self-efficacy for STEM (van der Meij et al., 2015), and Liew and Tan (2016) revealed digital agents mirroring the player's personality are advantageous. Furthermore, digital agents that show emotions (Theng & Aung, 2012) and empathic abilities can also increase player motivation (Chen et al., 2012). The impact of digital agents in games focuses especially on motivational and emotional levels and can increase learning motivation and outcomes (Martha & Santoso, 2019). Therefore, the use of digital agents could make a difference in the effect on learning in DGBL interventions (Hypothesis 2.9).

Confounding Factors

Year of Publication. Sitzmann (2011) did not find a shift in learning for simulation games between 1976 and 2009. But because of the rapid development of digital technology and the difficulty of generalizing to DGBL interventions other than simulations, whether learning has changed in recent years should be controlled. We therefore examine exploratively whether the year of publication influences the effects on DGBL (Research question 5).

Country of Data Collection. The country of data collection should also be considered when analyzing confounding factors. In some countries (e.g., Taiwan), digital learning environments are already well-established, in others (e.g., Germany) digital learning environments are relatively new in the classroom. This may lead to different effects, depending on the location of data collection, which we include exploratively in the analysis (Research question 6).

Objectives

Based on the previous explanations, the objective of this meta-analysis is to analyze the effect of DGBL interventions, assessed through different learning outcomes, in the school context in comparison with traditional instruction methods. Because the synthesis is based on published studies, we examined whether and to what extent publication bias confounded the results. Furthermore, we examined the impact of different personal factors, learning environment factors, and confounding factors to replicate previous findings and enhance the insight into DGBL intervention effectiveness to provide recommendations for the design of future digital learning environments.

Hypotheses

Based on the previous theoretical descriptions, we assume that digital game-based learning has positive effects regarding learning outcomes in the school context, which leads to the following hypotheses:

H 1.1 DGBL interventions in school lead to higher level overall learning, assessed through learning outcomes, than traditional instruction methods.

H 1.2 DGBL interventions in school are associated with higher level cognitive learning outcomes than traditional instruction methods.

H 1.3 DGBL interventions in school are associated with higher level metacognitive learning outcomes than traditional instruction methods.

H 1.4 DGBL interventions in school are associated with higher level affective-motivational learning outcomes than traditional instruction methods.

Moderator Hypotheses

In addition to the hypotheses regarding learning outcomes, we examined the influence of personal, environmental, and confounding factors by using moderator analyses. In a case of sufficient robust evidence from the literature, we deduced hypotheses for the corresponding moderators. If the evidence from the literature was unclear or insufficient, exploratory research questions (RQ) were deduced.

The deduced hypotheses for the corresponding moderators are:

H 2.1 DGBL interventions do not have a different impact in different age groups.

H 2.2 The effect of DGBL interventions is larger for female than for male participants.

H 2.3 DGBL interventions with multiple sessions lead to higher level learning gains than DGBL interventions with only one session.

H 2.4 Competitive DGBL interventions lead to higher level learning outcomes than noncompetitive DGBL interventions.

H 2.5 Three-dimensional DGBL interventions lead to greater effects on learning outcomes than two-dimensional or mixed DGBL interventions.

H 2.6 Schematic DGBL interventions lead to higher level learning gains than realistic or cartoonlike DGBL interventions.

H 2.7 DGBL interventions with a narrative lead to higher level learning gains than DGBL interventions without a narrative.

H 2.8 DGBL interventions that use avatars lead to higher level learning gains than DGBL interventions without avatars.

H 2.9 DGBL interventions that include digital agents are more effective for learning than DGBL interventions without digital agents.

The deduced exploratory research questions for the corresponding moderators are:

RQ 1 What is the difference regarding learning between DGBL interventions with and without additional non-game instruction?

RQ 2 What are the differences regarding learning between different DGBL intervention types?

RQ 3 How does the playing mode influence the effect of DGBL interventions regarding learning?

RQ 4 What influence does the learning domain have on the effect of DGBL interventions regarding learning?

RQ 5 What influence does the year of publication have?

RQ 6 What influence does the country of data collection have?

Method

Eligibility and Exclusion Criteria

To determine whether a study was eligible for synthesis, eligibility and inclusion criteria were defined before the search process started. The following passages describe the criteria in detail.

Digital Game-Based Learning

Eligible studies had to focus on DGBL interventions. To satisfy this criterion, according to the definition of Erhel and Jamet (2013), a complete digital game or simulation had to be implemented as an experimental condition on a computer or mobile device. Studies with games that were not digital (e.g., board games, role play in classroom, card games) or just used one single game element (gamification) were excluded from the sample (e.g., Sala & Gobet, 2017). Furthermore, games that used embodiment (e.g., *Microsoft Kinect*), augmented or virtual reality were also excluded from the sample (e.g., Lai et al., 2019) because of the different quality of immersion, which would be difficult to compare with other digital games.

Period

Due to the fast development of technology and a wider range of possibilities for designing digital games, only studies from 2015 to 2020 were suitable. Studies that were published before this period were excluded from the analyses because previous meta-analyses had already considered them.

Sample Size

To ensure that the effect size Cohen's *d*, needed to calculate Hedges' *g*, is distributed normally within the studies, only studies with more than 10 participants in total were selected. All studies with an insufficient number of participants were excluded from the analyses.

Target Population

Eligible studies had to provide sufficient information about the participants. For the synthesis, only studies with pupils from grades 1 to 13 (depending on the country where the study was conducted) and between ages 6 and 18 were eligible. Studies with university students, third-level education or employees were excluded (e.g., Cohen, 2016) due to assuming that adults show different learning patterns than adolescents and children (Kuhn & Pease, 2006). Studies focusing on special subpopulations (e.g., low- or high achievers, at-risk students) or participants with diseases were also not eligible (e.g., Fien et al., 2016).

Study Design

To avoid the 'garbage in – garbage out problem' (Döring & Bortz, 2016), only studies of good quality should contribute to the meta-analysis. Therefore, only studies with at least a quasi-experimental pre-post-control-design were eligible, studies that deviated from this requirement (e.g., no control group) were excluded (e.g., Tangsripairoj et al., 2019). The control group had to participate in a traditional learning setting or receive no intervention to satisfy the inclusion criteria. Studies that included control groups with digital pseudo-treatments (e.g., reading a website) were not eligible (e.g., Nussbaum et al., 2015), nor were studies whose experimental conditions developed a game on their own instead of participating in a DGBL intervention (e.g., Pellas & Vosinakis, 2018). Because the target population consisted of pupils, only studies conducted in a school setting were included, studies in a clinical setting or at the workplace were excluded (e.g., All et al., 2017).

Learning Outcomes

Selected studies had to measure at least one appropriate dependent variable representing students' cognitive, metacognitive, or affective-motivational learning outcomes. Studies were not eligible if they measured outcomes that did not address the analysis's objectives, e.g., studies that measured teachers' attitude towards digitalization or studies that reported children's frequency of media usage (e.g., Coombes et al., 2016).

Publication Type

Eligible studies had to be articles published in peer-reviewed journals or conference proceedings. Reviews, meta-analyses, reports, as well as all theses were excluded.

Language

Only studies written in German or English were selected. Studies written in other languages could not be considered for the synthesis (e.g., Ada et al., 2016).

Measures

Eligible studies had to use quantitative methods and provide sufficient statistical information to calculate effect sizes (e.g., sample size, *t*-value or *F*-value). Qualitative research, e.g., observational studies, and studies that did not report required statistical parameters were excluded.

Literature Search

The literature search period lasted from February 2020 until December 2020. The goal was a broad literature sample to avoid missing important findings. Thus, several psychological, medical, and computer science databases were consulted, namely *PubMed*, *ProQuest*, *IEEE*, *ACM*, *Web of Science*, *EBSCOhost,* and *Google Scholar*. Six research assistants, as well as the first author, conducted the literature search following a literature search manual.

To identify as many relevant studies between 2015 and 2020 as possible, the following search terms were used: 'Game-based learning', 'Simulation game', 'Serious game', 'Educational game' and each search term was also combined with 'training', 'learning' and 'education', respectively. Every database had special search settings to consider during the literature search. If possible (e.g., in *ProQuest*), a filter for peer-reviewed journals as well as Boolean search with "AND" (e.g., *EBSCOhost*) was used. If there were many hits for one search term, additional filters, e.g., 'undergraduate', 'patient', 'employees' and 'disorder' were applied to make the search more efficient.

Selection Process

Following Clark et al. (2016), eligible studies were selected following a three-step procedure based on a manual that contained the eligibility and exclusion criteria. In the first step, studies were selected by screening the titles in different databases. Thereby, a liberal approach was used and only studies were rejected that hit an exclusion criterion or were clearly irrelevant to answering the objectives of the analysis. In total 7,013 studies were selected in the first step. Altogether, 859 studies were selected from *IEE*, 632 from *PubMed*, 866 from *Google Scholar*, 59 from *ACM*, 2,625 from *Web of Science*, 675 from *ProQuest* and 1,298 from *EBSCOhost*. All duplicates were removed afterwards, and the abstracts of all remaining studies (n = 3,120) were screened to identify further eligible and ineligible articles. The full texts were considered in a last step to determine the final study sample for the analyses. After the selection process, N = 36 studies were extracted for the analyses. All selected studies are marked with an asterisk in the reference section. During the whole selection process, the first author was consulted in case of uncertainty regarding a study's eligibility. Figure 2 shows the flow chart for the selection process and the reasons why certain studies were not selected for the analysis. The category 'Other reasons', for example, includes studies in which the control group watched the experimental group playing or studies in which the teacher used the simulation, and the pupils were not allowed to interact with it. Furthermore, approaches that used just one game element rather than a complete game also joined the 'Other reasons' category.

Coding Procedure

The coding occurred after the selection process, by two raters independently following a coding guideline manual. For each study, the year of publication and study's location was recorded. Furthermore, processing included extracting the number of conditions and descriptive sample information (e.g., number of participants, mean age) and assessing participants' grade level. The number of training sessions, any additional non-game instruction (1 = yes, 0 = no) and the duration of the intervention were coded. For each study, the raters determined whether the intervention consisted of a serious game (coded with 1), a simulation (coded with 2) or a mix (coded with 3), the mode in which the participants played (1 = single player, 2 = multiplayer, 3 = mixed) and whether the game was competitive (1) or not (0). Coding also included further learning-environment characteristics, like dimensionality (1 = mixed, 2 = 2D, 3 = 3D), visual realism (1 = schematic, 2 = cartoon, 3 = realistic, 4 = mixed), narration (1 = yes, 0 = no) and the corresponding content domain. Moreover, coding included whether the players had an avatar with which to play (1 = yes, 0 = no) and non-player-characters (digital agents) with which to interact (1 = yes, 0 = no). All reported statistical parameters (e.g., *F*-value, *p*-value, degrees of freedom) were extracted to enable the calculation of the corresponding effect sizes, if necessary. With two raters and all coded variables nominally scaled, calculating the inter-rater reliability (Cohen's Kappa) assessed rater agreement while considering random agreements (Cohen, 1960). Based on conventional standards, values for Cohen's Kappa larger than .75 represent very good agreement, values between .60 and .75 are good, values between .40 and 60 are acceptable and values lower than .40 are not acceptable (Landis & Koch, 1977). For all coded variables, Cohens Kappa was between κ = .35 and κ = .76, depicting a variation of the rater agreement from 'low' to 'very good'. Due to the low congruence for playing mode (κ = .43) and visual realism (κ = .37), a third rater was consulted to rate the variables in question again. Hence, they achieved at least moderate agreement in all variables with a mean agreement of κ = .57.

Table **1** represents the particular inter-rater reliabilities in detail. Discussion with the first author until consensus was reached solved all remaining disagreements.

To categorize the different learning outcomes (cognitive, metacognitive, affective-motivational) for each study result, seven experts from an educational science department with expertise in self-regulated learning and its cognitive, metacognitive, and motivational components, were requested to estimate the learning outcomes for each study result. In total *n* = 57 cognitive learning outcomes, *n* = 5 metacognitive learning outcomes and *n* = 27 affective-motivational learning outcomes were observed.

Effect Size Measures

If reported, the pretest-adjusted posttest effect sizes for the outcomes of interest were extracted directly from the studies. In a first step, all reported effect sizes were converted into Cohen's *d* by using the pooled standard deviation in the denominator to enable the calculation of Hedges' g with small sample size correction. If an effect size was missing, the parameter was calculated from the provided statistical information in the study. After transferring all effect sizes coherent to Cohen's *d* with pooled standard deviation, all outcomes were converted to Hedges' *g*, corrected for small-sample bias. According to Cohen (1988), values smaller than .50 are interpreted as small, values between .50 and .80 as medium and values greater than .80 as large effect sizes.

However, knowing that this is just a rule of thumb, effect sizes will also be interpreted in relation to previous meta-analytic results. Some studies provided numerous results that rely on the same sample, causing dependent effect sizes. To cope with this circumstance, the average effect size was calculated from the different partial results.

Synthesis Method

In the present meta-analysis, one study was excluded as an outlier due to an unrealistically high effect size. Thus, the synthesis was conducted with a final sample of n = 35 studies. The random-effects meta-analysis and meta-regression models to analyze the impact of the moderator variables were carried out using R (version 4.1.2; R Core Team, 2020) and the *metafor* package (version 3.1.43; Viechtbauer, 2010). Due to the fact that game interventions in the selected studies varied, a random-effects model was fitted to the data and the amount of heterogeneity was estimated using the restricted maximum likelihood estimator (REML). To assess heterogeneity, the Q-statistic (Cochran, 1954) and the parameter τ^2, with its corresponding prediction intervals, were used to indicate the amount of heterogeneity in the random-effects model. Additionally, the inconsistency was also specified by using I^2 (Higgins & Thompson, 2002) and the confidence intervals were calculated with Wald-type CI. If heterogeneity was detected (i.e., $\tau^2 > 0$), a prediction interval for the true outcomes was also provided, which considers the intervention effect on the individual study level. Therefore, it depicts a more accurate way to interpret the results (Riley et al., 2011). Studentized residuals and Cook's distances were used to analyze whether studies may be outliers and examine their influence on the results (Viechtbauer & Cheung, 2010). Studies with a studentized residual larger than the $100 * (1 - \frac{0.05}{2 * k})$ th percentile of a standard normal distribution were considered potential outliers. Studies with a Cook's distance larger than the median plus six times the interquartile range of the Cook's distances were assumed to have a huge influence in the context of the model.

Given that the analyses are based on published articles, it is important to also consider the influence of publication bias. To analyze whether the sample reflected publication bias, Fail safe N, which determines the number of nonsignificant studies that must be included in the synthesis to annihilate a significant effect (Rosenthal, 1979), was calculated, and funnel plots were generated for a visual examination. Their symmetry was tested with Egger's regression test (Egger et al., 1997), using the standard error of the observed outcomes as predictor. An asymmetric funnel plot could be first evidence for the influence of publication bias. But neither the funnel plot nor Egger's regression test is appropriate for detecting and correcting publication bias (Hedges & Vevea, 1996). In fact, they deal with the influence of small study bias. To cope with this problem, selection models were calculated, because they take the publication probability into account by weighing the studies according to their p-value.

Results

Study characteristics

The n = 35 studies represent N = 7,139 participants in total. Participants' average age was 10.89 years (SD = 2.41), and the data was collected on average in grade six (SD = 2.81). Most of the studies were published in 2019 (n = 10). Ten studies were carried out in Europe, 6 in America, 16 in Asia, 2 in Australia and only one study in Africa. Most of the studies came from Taiwan (n = 6).

Table S2 (online only) shows further study characteristics, and Table S3(online only) illustrates an overview of the coded learning-environment factors.

Table S2 (online only) shows further study characteristics, and Table S3(online only) illustrates an overview of the coded learning-environment factors.

Results of syntheses

For the synthesis of the overall learning outcomes, the analysis included a total of κ = 35 studies. The observed outcomes ranged from -0.56 to 2.03, with 91% of the estimates being positive. Based on the random-effects model, the estimated average outcome was $g = 0.54$ (95% CI: [0.37, 0.72]) and differed significantly from zero ($z = 6.04$, $p < .001$). A forest plot showing the observed outcomes and the estimate appears in Figure S3 (online only).

The true outcomes appear to be heterogeneous (Q (34) = 3626.72, $p < .001$, $\tau^2 = 0.27$, $I^2 =$ 98.04 %). For the true outcomes, a 95% prediction interval is given at -0.49 to 1.58. Hence, although the average outcome is estimated to be positive, in some studies, the true outcome may be negative.

To analyze cognitive learning outcomes, the analysis included a total of κ = 29 studies. The observed outcomes ranged from 0.06 to 2.29, and all estimates were positive. Based on the random-effects model, the estimated average outcome was $g = 0.67$ (95% CI: 0.48 to 0.86). Therefore, the average outcome differed significantly from zero ($z = 6.83$, $p < .001$). Figure S4 (online only) shows a forest plot of the results.

According to the Q-test, the true outcomes seem to be heterogeneous (Q (28) = 3247.18, $p <$ .001, $\tau^2 = 0.27$, $I^2 = 98.16$ %) and a 95% prediction interval is given at -0.36 to 1.70. Although the average outcome is estimated to be positive, in some studies the true outcome may be negative.

Metacognitive outcomes were analyzed with a total of κ = 5 studies. The observed outcomes ranged from -0.56 to 1.14, with most estimates being positive (60%). Based on the random-effects model, the estimated average outcome was $g = 0.32$ (95% CI: -0.26 to 0.89) and did not differ significantly from zero ($z = 1.09$, p = .276). A forest plot showing the results appears in Figure S5 (online only).

According to the Q-test, the true outcomes seem to be heterogeneous (Q (4) = 157.12, p < .001, τ^2 = 0.42, I^2 = 97.62 %). A 95% prediction interval for the true outcomes is given at -1.07 to 1.71. Although the average outcome is estimated to be positive, in some studies the true outcome may be negative.

Affective-motivational learning outcomes were analyzed with a total of κ = 12 studies. The observed outcomes ranged from -0.24 to 1.78, with most estimates being positive (75%). Based on the random-effects model, the estimated average outcome was g = 0.32 (95% CI: 0.03 to 0.61) and differed significantly from zero (z = 2.14, p = .032). A forest plot showing the observed outcomes and the estimate is shown in Figure S6 (online only).

According to the Q-test, the true outcomes seem to be heterogeneous (Q (11) = 240.30, p < .001, τ^2 = 0.25, I^2 = 97.47 %) and for the true outcomes, a 95% prediction interval is given at -0.70 to 1.34. Although the average outcome is estimated to be positive, in some studies the true outcome may be negative.

Results of Moderator Analyses

To analyze the potential effects of different moderator variables, meta-regression models were conducted with the R package *metafor* (Viechtbauer, 2010). The results follow.

Results for Hypothesized Moderators

Age. The participants' mean age from each study was considered a continuous variable in the meta-regression model. No evidence could be found that participants' mean age influenced the impact of DGBL interventions (Q_M (1) = 0.25, p = .617). H2.1 was confirmed.

Gender. We considered the studies' gender distribution by calculating the percentage of female participants in each study. Information about the participants' gender was given in n = 23 studies, and n = 12 studies did not provide that information. The percentage of female participants ranged from 37% to 62%, with a mean percentage of M = 49.50% (SD = 5.60). The almost even distribution of males and females did not allow further examination because the differences between the studies were too small. As a result, H2.2 could not be tested.

Number of Sessions. The number of sessions ranged from 1 session to 50 sessions and, therefore, was split into quartiles. The first quartile contained studies with up to 2 sessions; the second quartile included studies with 3 to 7 sessions and the third quartile included studies with 8 to 19 sessions. Based on the quartiles, the categories became 0-2, 3-7, 8-19 and more than 19 sessions. As Table S4 (online only) shows, the estimated effect size was highest for games with more than 19 sessions, but there was no significant difference between the four categories (Q_M (3) = 1.82, p = .610). This did not conform with our hypothesis, and H2.3 was rejected.

Competition. As Table S4 (online only) shows, the findings in the current meta-analysis indicated that DGBL interventions without competitive elements are associated with higher learning gains than DGBL interventions with competition, but the meta-regression revealed no significant differences between DGBL interventions with and without competition (Q_M (1) = 0.21, p = .648). This did not conform with our hypothesis, and H2.4 was rejected.

Dimensionality. As Table S4 (online only) shows, DGBL interventions with mixed dimensions tended towards the strongest effect, followed by three-dimensional learning environments and two-dimensional DGBL interventions with the lowest effect. However, the meta-regression model showed no significant differences regarding dimensionality (Q_M (2) = 0.86, p = .651). This did not conform with our hypothesis, and H2.5 was rejected.

Visual Realism. The visual realism was coded using four different categories: schematic, cartoonlike, realistic, and mixed. The model results (see Table S4 (online only)) revealed a tendency towards the strongest effect for realistic games followed by cartoonlike and schematic games; DGBL interventions with mixed realism had the lowest effect on learning. But the meta-regression model found no significant differences across the mean effect sizes among the four categories (Q_M (3) = 3.23, p = .358). This did not conform with our hypothesis, and H2.6 was rejected.

Narration. As Table S4 (online only) shows, DGBL interventions with narration were associated with a greater increase in learning than DGBL interventions without narration, but the two categories do not differ significantly (Q_M (1) = 0.41, p = .523). This did not conform with our hypothesis, and H2.7 was rejected.

Avatar. Although the results in Table S4 (online only) suggested that DGBL interventions with avatars led to stronger effects on learning than DGBL interventions without avatars, the variable did not moderate the effect of DGBL interventions (Q_M (1) = 0.26, p = .609). This did not conform with our hypothesis, and H2.8 was rejected.

Digital Agent. Results suggested that DGBL interventions that use digital agents led to greater learning effects than DGBL interventions without digital agents (see Table S4 (online only)), but the meta-regression model indicated no significant differences between the two categories (Q_M (1) = 1.99, p = .159). This did not conform with our hypothesis, and H2.9 was rejected.

Results for Explorative Moderators

Additional Non-Game Instruction. The results indicated, as Table S4 (online only) shows, that DGBL interventions with additional non-game instruction seemed to be more effective than DGBL interventions without additional non-game instruction, but the two groups did not significantly differ (Q_M (1) = 0.44, p = .509).

Game Type. There was no evidence for a different impact regarding learning gains of serious games and simulations or mixed designs ($Q_M(2) = 4.41$, $p = .110$), but there was a trend, showing serious games as more effective than mixed DGBL interventions and simulations (see Table S4 (online only)).

Playing Mode. The tendency towards learning effects for mixed-mode DGBL interventions was the strongest, and games for single players closely followed (see Table S4 (online only)). The results indicated that DGBL interventions with multiple players have the least effect. Although the results hint at the superiority of mixed-mode games, the meta-regression model with a dummy indicator for playing mode revealed no significant differences across the variable ($Q_M(2) = 0.78$, $p = .677$) regarding DGBL intervention effectiveness.

Domain. The investigated studies covered a broad field of domains. The coded domains included mathematics, science, prevention, history, cybersecurity, literature, and cognitive skills. The results, that Table S4 (online only) presents, revealed a tendency towards the strongest effect for DGBL interventions with literature content and the weakest effect for DGBL interventions that promote cognitive skills. But the findings of the meta-regression model also promote no significant differences between the domains ($Q_M(6) = 10.55$, $p = .103$).

Publication Year. The publication year was considered as a continuous moderator variable and showed no significant influence ($Q_M(1) = 0.61$, $p = .434$).

Country of Data Collection. The countries of data collection were assigned to their continents, among which no significant differences appeared ($Q_M(4) = 0.20$, $p = .996$).

Reporting Biases

For the overall learning outcomes, an examination of the studentized residuals showed that no study had a value larger than $\pm$ 3.19, hence, there was no indication of outliers in the context of this model. According to Cook's distances, two studies (2; 28) could be strongly influential. Fail safe N was N_{fs} = 20,745. Because this number is higher than the rule of thumb ($N_{fs} > 5 * \kappa + 10$; Fragkos et al., 2014), the effect can be considered robust. A funnel plot of the estimates appears in Figure S7. The regression test did not indicate any funnel plot asymmetry ($z = 0.07$, $p = .947$).

The selection model showed no significant change in the estimate due to a nonsignificant likelihood ratio test ($\chi^2 (1) = 0.40$, $p = .528$), indicating no advantage of the adjusted model.

For cognitive learning outcomes, one study (23) may be a potential outlier. According to Cook's distances, two studies could have a very strong influence (2; 23). Fail safe N was N_{fs} = 22,796, which also indicates a robust effect. Figure S8 (online only) shows a funnel plot with the estimates. The regression test revealed no funnel plot asymmetry ($z = -0.28$, $p = .778$). The selection model indicated no advantage of an adjusted model because of a nonsignificant likelihood ratio test ($\chi^2 (1) = 3.70$, $p = .054$).

For metacognitive learning outcomes, no study had a residual larger than $\pm$ 2.58, so no study was regarded as an outlier. Furthermore, Cook's distances showed no highly influential studies. Fail safe N ($N_{fs} = 49$) suggested a robust effect. A funnel plot of the estimates appears in Figure S9 (online only), and there was no evidence of funnel plot asymmetry ($z = 1.44$, $p = .151$). The selection model showed no significant change in the estimate due to a nonsignificant likelihood ratio test ($\chi^2 (1) = 0.04$, $p = .845$), which indicates no advantage of the adjusted model.

For affective-motivational learning outcomes, Study 9 had residuals larger than $\pm$ 2.87 and might be an outlier. Cook's distances also showed that Study 9 could be overly influential. A robust effect could be assumed based on Fail safe N ($N_{fs} = 409$). A funnel plot of the estimates is shown in Figure S10 (online only). The regression test revealed a funnel plot asymmetry ($z = 2.15$, $p = .032$), that could be an indication of publication bias, but the selection model showed no significant change in the estimate due to a nonsignificant likelihood ratio test (χ^2 (1) = 0.07, p = .794).

In summary, no evidence of publication bias was revealed in the sample for all learning outcomes.

Discussion

Summary and interpretation of findings

The aim of the current meta-analysis was to examine the impacts DGBL interventions have in the school context, compared to traditional instruction methods. Furthermore, we analyzed whether different personal, environmental, or confounding factors influence the effectiveness of DGBL interventions. Based on the *Integrated Design Framework of Playful Learning* (Plass et al., 2015), cognitive and affective-motivational learning outcomes were included in the analyses and extended by considering metacognitive learning outcomes. With this approach, the current meta-analysis contributes to the replication of prior meta-analyses on the one hand and, on the other hand, extends the knowledge about DGBL interventions in the school context. The examination of metacognitive learning outcomes as well as avatars and digital agents as moderators were, to the knowledge of the authors, not considered in meta-analyses before. Unlike prior meta-analyses, the current meta-analysis focuses on pupils as a more specific target group which makes the results easier to generalize for the school context. The examination of DGBL interventions only in the last five years considers the rapid technologic development and enables a better comparability of the interventions used. Furthermore, the current meta-analysis contributes to theory-based research by deducing the examined learning outcomes on basis of the underlying theoretical model which increases the external validity of the current findings. Moreover, the aggregation of different study results leads to a higher validity which makes a meta-analytic approach superior to a single study.

For overall learning, a medium effect for DGBL interventions was revealed, confirming the assumption that DGBL interventions lead to improved learning compared to traditional instruction methods (H 1.1). This result is in accordance with previous meta-analytic evidence (e.g., Clark et al., 2016; Wouters et al., 2013), which suggests an advantage of DGBL interventions over traditional instruction methods. Hypothesis 1.2, the assumption that DGBL interventions improve cognitive learning outcomes more than traditional instruction methods, was also confirmed. A medium effect on learning was found for cognitive learning outcomes, accounting for the majority of studies. This result is in line with Lamb and colleagues (2018), who also found a medium effect of DGBL interventions on cognitive outcomes ($d = .67$). The assumption that DGBL interventions lead to higher metacognitive learning outcomes (H 1.3) could not be confirmed. It remains unclear whether DGBL interventions have an impact on metacognitive learning outcomes, because there was insufficient test power, due to a lack of studies ($n = 5$) that assessed this type of outcome. For affective-motivational learning outcomes (H 1.4), a small effect of DGBL interventions was detected, surprising because games are often described as motivating (Ryan & Rigby, 2020) and could lead to flow experiences (Chang et al., 2017). A reason for this small effect on affective-motivational learning outcomes could be the broad variability of DGBL interventions with different game designs. To be motivating, games must be designed thoughtfully (Lee & Hammer, 2011), and the game designers must ensure that the learning content is well integrated and does not overshadow the engaging game mechanics (De Freitas et al., 2018). In our sample, it could be possible that the balance between gameplay and learning was not achieved in every DGBL intervention, which could have impacted the effect on affective-motivational learning.

The results concerning these hypotheses regarding learning outcomes reinforce the effectiveness of DGBL interventions, based on the *Integrated Design Framework for Playful*

Learning (Plass et al., 2015), and highlight the cognitive and affective-motivational engagement the model assumes.

In addition to examining of the hypotheses regarding learning outcomes, the influence of personal, environmental, and confounding moderators was analyzed by using meta-regression models. Hypothesis 2.1 assumed no age differences for the impact of DGBL interventions and was confirmed. In the current meta-analysis, participants' age was not influential regarding the effectiveness of DGBL interventions. Hypothesis 2.2., the influence of gender in DGBL interventions, could not be tested because of the balanced distribution of male and female participants in the included studies.

The number of sessions showed an advantage of multiple sessions on a descriptive level in the current study, but unlike the studies of Wouters et al. (2013) and Clark et al. (2016), no significant difference between the different numbers of sessions was found. The current findings suggest that the number of sessions for DGBL interventions does not have an influence on their effectiveness. Thus, Hypothesis 2.3, the assumption that DGBL interventions with multiple sessions lead to greater learning gains, could not be confirmed. Moreover, using competitive or noncompetitive environments made no difference regarding the effectiveness of DGBL interventions. Therefore, the assumption that competitive DGBL interventions lead to greater learning gains (H 2.4) was rejected. Competitive games did not lead to greater learning gains than noncompetitive games in the current meta-analysis which is surprising, because Abdul Jabbar and Felicia (2015) identified competition in games in their systematic review 'as a gameplay element that could emotionally and cognitively engage players and could have a significant impact on learning' (p. 762). Lamb et al. (2018) found a significantly stronger effect for three-dimensional over two-dimensional DGBL interventions, which the current data could not replicate because no significant differences between the dimensions were found (H2.5). Based on the current findings, the dimensionality of a game which is used as DGBL intervention seems to make no difference in the effectiveness of DGBL interventions. The same applies to visual realism, narrative, avatars, and digital agents, for which no moderating effect could be detected (H2.6–H 2.9). A trend exists on the descriptive level that DGBL interventions with a narrative, avatars and digital agents outperform DGBL interventions without these features, but with no significant evidence.

For the exploratory research questions, no significant influence of additional non-game instruction, game type, playing mode, and learning domain appeared. Furthermore, the year of publication and the country of data collection had no confounding impact on the results.

The unexpected results for the hypotheses and research questions could have been caused by the sample's very high heterogeneity. The DGBL interventions may differ too much from each other, exacerbating the analyses and not leading to meaningful results. Furthermore, some analyzed subgroups were very small, which could have also prevented detecting significant differences.

Because the basis for the current analyses was published literature, we examined their robustness and the influence of publication bias. Fail Safe N strengthened the assumption that the results were robust, and we checked publication bias visually with funnel plots, testing their symmetry with Egger's regression test and finding no evidence for asymmetry, except for affective-motivational learning outcomes. Since these methods do not very accurately detect and correct for publication bias (Hedges & Vevea, 1996), selection models were used additionally. Hence, the selection models did not significantly better fit to the data, leading to the conclusion that the current data showed no publication bias.

Limitations

Although the present meta-analysis could partially extend the findings of prior meta-analyses, several limitations must be considered when interpreting the results. Regarding the literature research, *Google Scholar* only shows the first 1,000 hits for each search term, which did not allow an in-depth literature search. Furthermore, the different databases had varying options to refine the search process, which could have caused a failure to find all relevant studies, always a possibility when working with data based on a literature search. To minimize the probability of missing important studies, a broad literature search in different data bases should occur. Furthermore, since the authors can only read papers in English and German, a language bias in the data is possible. Inability to include studies in other languages in the synthesis could confound the results.

Concerning the coding procedure, determining visual realism was based on the graphics in the papers provided. Depending on the number of photos, deciding to which category of visual realism the game belongs was not easy for the raters. One cannot rule out the possibility that games were miscategorized due to a lack of pictures, which would also affect the meta-regression results.

The decision to use the mean of dependent effect sizes for each study could have led to loss of information in the dataset. However, a multivariate meta-analysis, sometimes suggested as an alternative procedure, has high costs (e.g., complicated procedure, high dataset requirements) and its use is controversial. In our opinion, its benefits (e.g., less loss of information) did not balance the costs, the reason we chose to use the dependent effect sizes' mean.

Another limitation is the small number of studies assessing metacognitive learning outcomes—only five studies in the current meta-analysis to analyze the effect of DGBL interventions on those outcomes. Therefore, the results should be interpreted with caution, and the impact of DGBL interventions on metacognitive learning outcomes remains a desideratum on which prospective studies should focus.

Regarding the moderator analyses, one limiting factor is the unequal distribution of studies in the analyzed categories. For example, only one study with a mixed playing mode was compared to 29 single-player games and only 4 multiplayer games. Furthermore, the primary studies did not provide some information on the moderators, leading to a smaller sample that made the analyses and the detection of moderator effects more unlikely. Moreover, the considered moderator 'number of sessions' has limited informative value because the number of sessions does not reveal the duration of the intervention (e.g., five sessions could have occurred in one week or in three weeks). Future studies should consider the exact duration as a more accurate variable than the number of sessions. Also, the year of publication is not a very precise variable. If the paper does not mention it, we do not know when the data was collected. A study published in 2015 could potentially use data from 2010, which could also influence the results.

Implications and future directions

Despite these limitations, the current study strengthens the assumption that DGBL interventions are also effective for pupils' learning, especially for cognitive and affective-motivational outcomes.

As a practical implication, the findings confirm that DGBL in school provides another tool for teachers to use in class, to offer variation in learning methods for their pupils. Based on the current findings, well-designed DGBL interventions could motivate pupils to engage with learning materials and, therefore, promote their learning process. This could be especially beneficial for pupils who are bored in traditional learning settings or pupils who have the need either to work more autonomous or to collaborate with other pupils. Due to the high heterogeneity in the classroom, teachers cannot always satisfy all their pupils' needs. DGBL interventions in the classroom can be designed to adapt to their users' needs by, for example, offering exercises with an adequate difficulty for the user. This could contribute to a learner-centered pedagogy, which focusses on learners' needs and enables them to learn at their own pace and to avoid excessive demands and resignation. In fact, DGBL interventions offer a safe learning environment, which allows 'graceful failure' (Plass et al., 2015, p. 261) that promotes experiences of success (mastery experiences) and could lead to an increased self-efficacy. Cognitive outcomes seem especially suitable for DGBL interventions, making them an adequate tool for teachers to convey knowledge to their pupils. Furthermore, DGBL could also occur during distance learning and, therefore, be a learning solution for pupils who cannot attend school in person, e.g., because of quarantine. Indeed, pupils need access to digital devices or computers, requiring well-equipped school and home environments. However, there are still schools which lack the equipment for DGBL interventions. To improve the possibilities to offer DGBL interventions in the classroom or at home, it should be the governance's responsibility to invest in education and, therefore, provide financial support for schools and families.

Another important factor for the successful implementation of DGBL interventions in school is teachers' technology literacy (Marklund & Alklind Taylor, 2016). Studies reveal that teachers are interested in using DGBL interventions but are often worried whether they have enough knowledge and gaming literacy to implement DBGL interventions in the classroom (Jong, 2016). This shows the need to foster the knowledge and self-efficacy to use digital games by integrating the use of different digital media, including digital games, into the teacher education curriculum.

An implication for research is the revealed lack of studies that analyze metacognitive learning outcomes, calling for more studies that target those outcomes. Another topic for prospective research is that to be motivating, future DGBL interventions need well-integrated learning content balanced with intrinsic or extrinsic game elements. Game designers must consider individually which game elements suit the intervention's learning goals. A general recommendation for fitting a DGBL intervention to all pupils is not possible because pupils have individual needs and preferences during learning. Therefore, adaptive games that align with pupils' needs during the learning process could be a solution. But the development of DGBL interventions in general, and especially adaptive games, is time-consuming and costly. If it fits the learning goal, researchers could make use of publicly available games whenever possible. The question of which game-design elements cause the positive effects of DGBL is still unclear, and future studies should examine it. Future DGBL intervention studies should focus on specific game-design elements separately and compare different game versions.

The impact of DGBL interventions on teacher education would also be interesting to study in future DGBL intervention research because teachers would gain experience with digital games and could act as models for their prospective students (Peeters et al., 2014). Also involving the teachers in acting as a multiplicator could promote the learning with digital games on two different levels. This means that pupils could acquire knowledge by playing digital games and learn by observing and adopting teachers' behavior.

Furthermore, the results of the current meta-analysis cannot be transferred to all DGBL interventions in general. They include only games without embodiment, augmented or virtual reality. Due to the different quality of immersion (Skarbez et al., 2021), they are not comparable to games that do not use these methods. Therefore, future studies considering DGBL interventions with embodiment, augmented or virtual reality would be interesting. Given that the current meta-analysis did not find any moderating effects, a deeper examination of mentioned and additional moderators (e.g., intervention duration, device type) could also provide new insights into DGBL interventions' effectiveness.

References

Studies which were used in the analysis are marked with *

Abdul Jabbar, A. I., & Felicia, P. (2015). Gameplay Engagement and Learning in Game-Based Learning: A Systematic Review. *Review of Educational Research*, 85(4), 740–779. https://doi.org/10.3102/00346543155772

Ada, M., Akdeniz, M. F., Gümgüm, Ö., Keskin, M., Yazıcı, A., & Yayan, U. (2016, May). Interactive Serious Games With Visual Programming for Mobile Robot Learning. In *2016 24th Signal Processing and Communication Application Conference (SIU)* (pp. 485–488). IEEE. https://doi.org/10.1109/SIU.2016.7495783

Admiraal, W., Huizenga, J., Akkerman, S., & Ten Dam, G. (2011). The Concept of Flow in Collaborative Game-Based Learning. *Computers in Human Behavior*, *27*(3), 1185–1194. https://doi.org/10.1016/j.chb.2010.12.013

Ak, O., & Kutlu, B. (2017). Effect of Game Environments on Students' Learning. *British Journal of Educational Technology*, 48, 129–144. https://doi.org/10.1111/bjet.12346

Al-Azawi, R., Al-Faliti, F., & Al-Blushi, M. (2016). Educational Gamification vs. Game Based Learning: Comparative Study. *International Journal of Innovation, Management and Technology*, *7*(4), 131–136. https://doi.org/10.18178/ijimt.2016.7.4.659

All, A., Plovie, B., Castellar, E. P. N., & Van Looy, J. (2017). Pre-Test Influences on the Effectiveness of Digital-Game Based Learning: A Case Study of a Fire Safety Game. *Computers & Education*, 114, 24–37. https://doi.org/10.1016/j.compedu.2017.05.018

Allen, K. N., & Friedman, B. D. (2010). Affective Learning: A Taxonomy for Teaching Social Work Values. *Journal of Social Work Values and Ethics*, *7*(2), 1–12.

Bandura, A. (1989). Human Agency in Social Cognitive Theory. *American Psychologist*, *44*(9), 1175–1184. https://doi.org/10.1037/0003-066X.44.9.1175

Bandura, A. (1997). *Self-Efficacy: The Exercise of control*. Freeman.

* Barros, C., Carvalho, A. A., & Salgueiro, A. (2020). The Effect of the Serious Game Tempoly on Learning Arithmetic Polynomial Operations. *Education and Information Technologies*, *25*(3), 1497–1509. https://doi.org/10.1007/s10639-019-09990-4

Bowers, C. A., Serge, S., Blair, L., Cannon-Bowers, J., Joyce, R., & Boshnack, J. (2013). The Effectiveness of Narrative Pre-Experiences for Creating Context in Military Training. *Simulation & Gaming*, *44*(4), 514–522. https://doi.org/10.1177/1046878113475341

* Brezovszky, B., McMullen, J., Veermans, K., Hannula-Sormunen, M. M., Rodríguez-Aflecht, G., Pongsakdi, N., Laakkonen, E., & Lehtinen, E. (2019). Effects of a Mathematics Game-Based Learning Environment on Primary School Students' Adaptive Number Knowledge. *Computers & Education*, *128*, 63–74. https://doi.org/10.1016/j.compedu.2018.09.011

Cagiltay, N. E., Ozcelik, E., & Ozcelik, N. S. (2015). The Effect of Competition on Learning in Games. *Computers & Education*, *87*, 35–41. https://doi.org/10.1016/j.compedu.2015.04.001

* Cayvaz, A., Akcay, H., & Kapici, H. O. (2020). Comparison of Simulation-Based and Textbook-Based Instructions on Middle School Students' Achievement, Inquiry Skills and Attitudes. *International Journal of Education in Mathematics, Science and Technology*, *8*(1), 34–43. https://doi.org/10.46328/ijemst.v8i1.758

* Cejudo, J., López-Delgado, M. L., & Losada, L. (2019). Effectiveness of the Videogame "Spock" for the Improvement of the Emotional Intelligence on Psychosocial Adjustment in Adolescents. *Computers in Human Behavior*, *101*, 380–386. https://doi.org/10.1016/j.chb.2018.09.028

Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed Practice in Verbal Recall Tasks: A Review and Quantitative Synthesis. *Psychological Bulletin*, *132*(3), 354–380. https://doi.org/10.1037/0033-2909.132.3.354

Chang, C.-C., Liang, C., Chou, P.-N., & Lin, G.-Y. (2017). Is Game-Based Learning Better in Flow Experience and Various Types of Cognitive Load Than Non-Game-Based Learning? Perspective From Multimedia and Media Richness. *Computers in Human Behavior*, *71*, 218–227. https://doi.org/10.1016/j.chb.2017.01.031

* Chen, C.-C., & Huang, P.-H. (2020). The Effects of STEAM-Based Mobile Learning on Learning Achievement and Cognitive Load. *Interactive Learning Environments*, 1–17. https://doi.org/10.1080/10494820.2020.1761838

Chen, C.-H. (2019). The Impacts of Peer Competition-Based Science Gameplay on Conceptual Knowledge, Intrinsic Motivation, and Learning Behavioral Patterns. *Educational Technology Research and Development*, *67*(1), 179–198. https://doi.org/10.1007/s11423-018-9635-5

Chen, C.-H., & Chiu, C.-H. (2016). Employing Intergroup Competition in Multitouch Design-Based Learning to Foster Student Engagement, Learning Achievement, and Creativity. *Computers & Education*, *103*, 99–113. https://doi.org/10.1016/j.compedu.2016.09.007

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

* 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, G.-D., Lee, J.-H., Wang, C.-Y., Chao, P.-Y., Li, L.-Y., & Lee, Y. (2012). An Empathic Avatar in a Computer-Aided Learning Program to Encourage and Persuade Learners. *Journal of Educational Technology & Society, 15*(2), 62–72.

* Chen, H.-R., Liao, K.-C., & Chang, J.-J. (2015). Design of Digital Game-Based Learning System for Elementary Mathematics Problem Solving. *2015 8th International Conference on Ubi-Media Computing (UMEDIA)*, 303–307. https://doi.org/10.1109/UMEDIA.2015.7297475

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

Cochran, W. G. (1954). The Combination of Estimates From Different Experiments. *Biometrics*, *10*(1), 101–129. https://doi.org/10.2307/3001666

Cohen, E. L. (2016). Enjoyment of a Counter-Hedonic Serious Digital Game: Determinants and Effects on Learning and Self-Efficacy. *Psychology of Popular Media Culture*, *5*(2), 157–170. http://dx.doi.org/10.1037/ppm0000052

Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement*, *20*(1), 37–46. https://doi.org/10.1177/001316446002000104

Cohen, J. (1988). *Statistical Power Analysis for the Behavioural Sciences* (2. Ed.). Erlbaum Associates.

Cohen, J. (2001). Defining Identification: A Theoretical Look at the Identification of Audiences With Media Characters. *Mass Communication & Society, 4*(3), 245–264. https://doi.org/10.1207/S15327825MCS0403_01

Cojocariu, V.-M., & Boghian, I. (2014). Teaching the Relevance of Game-Based Learning to Preschool and Primary Teachers. *Procedia - Social and Behavioral Sciences*, *142*, 640–646. https://doi.org/10.1016/j.sbspro.2014.07.679

Connolly, T. M., & Stansfield, M. (2009). From e-learning to games-based e-learning. In *Encyclopedia of Information Communication Technology* (pp. 268–275). IGI Global. https://doi.org/10.4018/978-1-59904-845-1.ch035

Coombes, L., Chan, G., Allen, D., & Foxcroft, D. R. (2016). Mixed-Methods Evaluation of the Good Behaviour Game in English Primary Schools. *Journal of community & applied social psychology*, *26*(5), 369-387. https://doi.org/10.1002/casp.2268

De Freitas, V., Mohan, P., & Kinshuk, (2018). A Game Designers' Guide for Creating Learning Games for Mathematics. *2018 IEEE 18th International Conference on Advanced Learning Technologies (ICALT)*, 146–148. https://doi.org/10.1109/ICALT.2018.00144

Dickey, M. D. (2020). Narrative in Game-Based Learning. Plass, J. L., Mayer, R. E. & Homer, B. D. (Eds..), *Handbook of Game-Based Learning,* (pp. 283–306). MIT Press.

Domagk, S., Schwartz, R. N., & Plass, J. L. (2010). Interactivity in Multimedia Learning: An Integrated Model. *Computers in Human Behavior*, *26*(5), 1024–1033. https://doi.org/10.1016/j.chb.2010.03.003

Döring, N., & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-41089-5

* Dorji, U., Panjaburee, P., & Srisawasdi, N. (2015). A Learning Cycle Approach to Developing Educational Computer Game for Improving Students' Learning and Awareness in Electric Energy Consumption and Conservation. *Educational Technology & Society, 18*(1), 91–105.

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in Meta-Analysis Detected by a Simple, Graphical Test. *BMJ*, *315*(7109), 629–634. https://doi.org/10.1136/bmj.315.7109.629

Erhel, S., & Jamet, E. (2013). Digital Game-Based Learning: Impact of Instructions and Feedback on Motivation and Learning Effectiveness. *Computers & Education*, *67*, 156–167. https://doi.org/10.1016/j.compedu.2013.02.019

* Fendt, M. W., & Ames, E. (2019). Using Learning Games to Teach Texas Civil War History to Public Middle School Students. *IEEE Conference on Games (CoG)*, 1–4. https://doi.org/10.1109/CIG.2019.8847968

Fien, H., Doabler, C. T., Nelson, N. J., Kosty, D. B., Clarke, B., & Baker, S. K. (2016). An Examination of the Promise of the NumberShire Level 1 Gaming Intervention for Improving Student Mathematics Outcomes. *Journal of Research on Educational Effectiveness*, *9*(4), 635-661. https://doi.org/10.1080/19345747.2015.1119229

* Fiorella, L., Kuhlmann, S., & Vogel-Walcutt, J. J. (2019). 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

Fragkos, K. C., Tsagris, M., & Frangos, C. C. (2014). Publication Bias in Meta-Analysis: Confidence Intervals for Rosenthal's Fail-Safe Number. *International Scholarly Research Notices*, 1–17. https://doi.org/10.1155/2014/825383

Franciosi, S. J. (2017). The Effect of Computer Game-Based Learning on FL Vocabulary Transferability. *Educational Technology & Society*, 20 (1), 123–133.

* Freina, L., Bottino, R., & Ferlino, L. (2018). Visuospatial Abilities Training With Digital Games in a Primary School. *International Journal of Serious Games*, *5*(3), 23–35. https://doi.org/10.17083/ijsg.v5i3.240

* Giannakas, F., Papasalouros, A., Kambourakis, G., & Gritzalis, S. (2019). A Comprehensive Cybersecurity Learning Platform for Elementary Education. *Information Security Journal: A Global Perspective*, *28*(3), 81–106. https://doi.org/10.1080/19393555.2019.1657527

* Hannel, S. L., & Cuevas, J. (2018). A Study on Science Achievement and Motivation Using Computer-Based Simulations Compared to Traditional Hands-on Manipulation. *Georgia Educational Researcher*, *15*(1), 38–55. https://doi.org/10.20429/ger.2018.15103

* Hawkins, R. D., Ferreira, G. A. M., & Williams, J. M., (2019). The Development and Evaluation of 'Farm Animal Welfare': An Educational Computer Game for Children. *Animals*, *9*(3), 1–17. https://doi.org/10.3390/ani9030091

Hedges, L. V., & Vevea, J. L. (1996). Estimating Effect Size Under Publication Bias: Small Sample Properties and Robustness of a Random Effects Selection Model. *Journal of Educational and Behavioral Statistics, 21*(4), 299–332. https://doi.org/10.3102/10769986021004299

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying Heterogeneity in a Meta-Analysis. *Statistics in Medicine*, *21*(11), 1539–1558. https://doi.org/10.1002/sim.1186

* Hussein, M. H., Ow, S. H., Cheong, L. S., & Thong, M.-K. (2019). A Digital Game-Based Learning Method to Improve Students' Critical Thinking Skills in Elementary Science. *IEEE Access*, *7*, 96309–96318. https://doi.org/10.1109/ACCESS.2019.2929089

* Hwang, G.-J., Hsu, T.-C., Lai, C.-L., & Hsueh, C.-J. (2017). Interaction of Problem-Based Gaming and Learning Anxiety in Language Students' English Listening Performance and Progressive Behavioral Patterns. *Computers & Education*, *106*, 26–42. https://doi.org/10.1016/j.compedu.2016.11.010

Jackson, L. C., O'Mara, J., Moss, J., & Jackson, A. C. (2018). A Critical Review of the Effectiveness of Narrative-Driven Digital Educational Games. *International Journal of Game-Based Learning*, *8*(4), 32–49. https://doi.org/10.4018/IJGBL.2018100103

Johnson, W. L., & Lester, J. C. (2016). Face-to-Face Interaction with Pedagogical Agents, Twenty Years Later. *International Journal of Artificial Intelligence in Education*, *26*(1), 25–36. https://doi.org/10.1007/s40593-015-0065-9

Jong, M. S. Y. (2016). Teachers' concerns about adopting constructivist GBL. *British Journal of Educational Technology*, 47: 601–617. https://doi.org/10.1111/bjet.12247

* Kapici, H. O., Akcay, H., & de Jong, T. (2020). How do Different Laboratory Environments Influence Students' Attitudes Toward Science Courses and Laboratories? *Journal of Research on Technology in Education*, *52*(4), 534–549. https://doi.org/10.1080/15391523.2020.1750075

Karakoç, B., Eryılmaz, K., Turan Özpolat, E., & Yıldırım, İ. (2020). The Effect of Game-Based Learning on Student Achievement: A Meta-Analysis Study. *Technology, Knowledge and Learning*, *27*(1), 207–222. https://doi.org/10.1007/s10758-020-09471-5

* Ke, F. (2019). Mathematical Problem Solving and Learning in an Architecture-Themed Epistemic Game. *Educational Technology Research and Development*, *67*(5), 1085–1104. https://doi.org/10.1007/s11423-018-09643-2

* Khan, A., Ahmad, F. H., & Malik, M. M. (2017). Use of Digital Game Based Learning and Gamification in Secondary School Science: The Effect on Student Engagement, Learning and Gender Difference. *Education and Information Technologies*, *22*(6), 2767–2804. https://doi.org/10.1007/s10639-017-9622-1

* Kiili, K., Moeller, K., & Ninaus, M. (2018). Evaluating the Effectiveness of a Game-Based Rational Number Training - In-Game Metrics as Learning Indicators. *Computers & Education*, *120*, 13–28. https://doi.org/10.1016/j.compedu.2018.01.012

Kinzer, C. K., Hoffman, D. L., Turkay, S., Gunbas, N., Chantes, P., Dvorkin, T., & Chaiwinij, A. (2012). The Impact of Choice and Feedback on Learning, Motivation, and Performance in an Educational Video Game. *Proceedings of the Games, Learning, and Society Conference, 2*, 175–181.

Klimmt, C., Hefner, D., Vorderer, P., Roth, C., & Blake, C. (2010). Identification With Video Game Characters as Automatic Shift of Self-Perceptions. *Media Psychology*, *13*(4), 323–338. https://doi.org/10.1080/15213269.2010.524911

Kuhn, D., & Pease, M. (2006). Do Children and Adults Learn Differently? *Journal of Cognition and Development*, *7*(3), 279–293. https://doi.org/10.1207/s15327647jcd0703_1

Lai, A. F., Chen, C. H., & Lee, G. Y. (2019). An Augmented Reality-Based Learning Approach to Enhancing Students' Science Reading Performances From The Perspective of the Cognitive Load Theory. *British Journal of Educational Technology*, *50*(1), 232–247. https://doi.org/10.1111/bjet.12716

* Lamb, R. L. (2016). Examination of the Effects of Dimensionality on Cognitive Processing in Science: A Computational Modeling Experiment Comparing Online Laboratory Simulations and Serious Educational Games. *Journal of Science Education and Technology*, *25*(1), 1–15. https://doi.org/10.1007/s10956-015-9587-z

Lamb, R. L., Annetta, L., Firestone, J., & Etopio, E. (2018). A Meta-Analysis With Examination of Moderators of Student Cognition, Affect, and Learning Outcomes While Using Serious Educational Games, Serious Games, and Simulations. *Computers in Human Behavior*, *80*, 158–167. https://doi.org/10.1016/j.chb.2017.10.040

Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, *33*(1), 159–174. https://doi.org/10.2307/2529310

* Lee, H. K., & Choi, A. (2020). Enhancing Early Numeracy Skills With a Tablet-Based Math Game Intervention: A Study in Tanzania. *Educational Technology Research and Development*, *68*(6), 3567–3585. https://doi.org/10.1007/s11423-020-09808-y

Lee, J. J., & Hammer, J. (2011). Gamification in Education: What, How, Why Bother? *Academic Exchange Quarterly, 15*(2), 146.

Lester, J. C., Spain, R. D., Rowe, J. P., & Mott, B. W. (2020). Instructional Support, Feedback, and Coaching in Game-based Learning. In J. L. Plass, R. E. Mayer & B. D. Homer (Eds.), *Handbook of Game-based Learning* (pp. 209–238). MIT Press.

Lester, J. C., Spires, H. A., Nietfeld, J. L., Minogue, J., Mott, B. W., & Lobene, E. V. (2014). Designing Game-Based Learning Environments for Elementary Science Education: A Narrative-Centered Learning Perspective. *Information Sciences*, *264*, 4–18. https://doi.org/10.1016/j.ins.2013.09.005

Liew, T. W., & Tan, S.-M. (2016). Virtual Agents With Personality: Adaptation of Learner-Agent Personality in a Virtual Learning Environment. *Eleventh International Conference on Digital Information Management (ICDIM)*, 157–162. https://doi.org/10.1109/ICDIM.2016.7829758

Loderer, K., Pekrun, R., & Plass, J. L. (2020). Emotional Foundations of Game-Based Learning. In J. L. Plass, R. E. Mayer & B. D. Homer (Eds.), *Handbook of Game-based learning* (pp. 111–152). MIT Press.

López-Fernández, D., Gordillo, A., Alarcón, P. P., & Tovar, E. (2021). Comparing Traditional Teaching and Game-Based Learning Using Teacher-Authored Games on Computer Science

Education. *IEEE Transactions on Education*, 64(4), 367–373. https://doi.org/10.1109/TE.2021.3057849

Malliarakis, C., Tomos, F., Shabalina, O., & Mozelius, P. (2018). Andragogy and E.M.O.T.I.O.N: 7 Key Factors of Successful Serious Games. In M. Ciussi (Ed.), *Proceedings of the 12th European Conference on Games Based Learning* (pp. 371–378). ACI.

Marklund, B. B., & Alklind Taylor, A. S. (2016). Educational Games in Practice: The Challenges Involved in Conducting a Game-Based Curriculum. *The Electronic Journal of E-Learning*, 14(2), 122–135.

Martha, A. S. D., & Santoso, H. B. (2019). The Design and Impact of the Pedagogical Agent: A Systematic Literature Review. *Journal of Educators Online*, *16*(1). https://doi.org/10.9743/jeo.2019.16.1.8

Mayer, R. E. (2014). Computer Games for Learning: An Evidence-Based Approach. *Teachers College Record* (pp. 1–3). MIT Press.

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). MIT Press.

McLaren, B. M., Adams, D. M., Mayer, R. E., & Forlizzi, J. (2017). A Computer-Based Game That Promotes Mathematics Learning More Than a Conventional Approach. *International Journal of Game-Based Learning*, *7*(1), 36–56. https://doi.org/10.4018/IJGBL.2017010103

Nelson, B., & Kim, Y. (2020). Multimedia Design Principles in Game-Based Learning. In J. L. Plass, R. E. Mayer & B. D. Homer (Eds.), *Handbook of Game-Based Learning* (pp. 307–328). MIT Press.

Novak, E. (2015). A Critical Review of Digital Storyline-Enhanced Learning. *Educational Technology Research and Development*, *63*(3), 431–453. https://doi.org/10.1007/s11423-015-9372-y

Nussbaum, E. M., Owens, M. C., Sinatra, G. M., Rehmat, A. P., Cordova, J. R., Ahmad, S., Harris Jr., F. C., & Dascalu, S. M. (2015). Losing the Lake: Simulations to Promote Gains in Student Knowledge and Interest About Climate Change. *International Journal of Environmental and Science Education*, *10*(6), 789–811.

* O'Rourke, J., Main, S., & Hill, S. M., (2017). Commercially Available Digital Game Technology in the Classroom: Improving Automaticity in Mental-Maths in Primary-Aged Students. *Australian Journal of Teacher Education*, *42*(10), 50–70. https://doi.org/10.14221/ajte.2017v42n10.4

Partovi, T., & Razavi, M. R. (2019). The Effect of Game-Based Learning on Academic Achievement Motivation of Elementary School Students. *Learning and Motivation*, 68: 101592. https://doi.org/10.1016/j.lmot.2019.101592

Pawar, S., Tam, F., & Plass, J. L. (2020). Emerging Design Factors of Game-Based Learning: Emotional Design, Musical Score, and Game Mechanics Design. In J. L. Plass, R. E. Mayer & B. D. Homer (Eds.), *Handbook of Game-Based Learning* (pp. 347–367). MIT Press.

Peeters, J., De Backer, F., Reina, V. R., Kindekens, A., Buffel, T., & Lombaerts, K. (2014). The Role of Teachers' Self-Regulatory Capacities in the Implementation of Self-Regulated Learning Practices. *Procedia - Social and Behavioral Sciences*, *116*, 1963–1970. https://doi.org/10.1016/j.sbspro.2014.01.504

Pekrun, R. (2006). The Control-Value Theory of Achievement Emotions: Assumptions, Corollaries, and Implications for Educational Research and Practice. *Educational Psychology Review*, *18*(4), 315–341. https://doi.org/10.1007/s10648-006-9029-9

Pellas, N., & Vosinakis, S. (2018). The Effect of Simulation Games on Learning Computer Programming: A Comparative Study on High School Students' Learning Performance by Assessing Computational Problem-Solving Strategies. *Education and Information Technologies*, *23*(6), 2423–2452. https://doi.org/10.1007/s10639-018-9724-4

Plass, J. L., Homer, B. D., Kinzer, C., Frye, J., & Perlin, K. (2011). Learning Mechanics and Assessment Mechanics for Games for Learning. *White Paper, 1*, 1–19. https://doi.org/10.13140/2.1.3127.1201

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., Home r, B. D., Mayer, R. E., & Kinzer, C. K. (2020). Theoretical Foundations of Game-Based and Playful Learning. In J. L. Plass, R. E. Mayer & B. D. Homer (Eds.), *Handbook of Game-Based Learning* (pp. 3–24). MIT Press.

Plass, J. L., & Pawar, S. (2020). Adaptivity and Personalization in Game-Based Learning. In J. L. Plass, R. E. Mayer & B. D. Homer (Eds.), *Handbook of Game-Based Learning* (pp. 263–282). MIT Press.

Post, L. S., Guo, P., Saab, N., & Admiraal, W. (2019). Effects of Remote Labs on Cognitive, Behavioral, and Affective Learning Outcomes in Higher Education. *Computers & Education*, *140*, 1–9. https://doi.org/10.1016/j.compedu.2019.103596

Prensky, M., & Prensky, M. (2007). Digital Game-Based Learning. *Paragon House, 1.*

R Core Team. (2020). A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing.* https://www.R-project.org/

* Ramos, D. K., & Melo, H. M. (2019). Can Digital Games in School Improve Attention? A Study of Brazilian Elementary School Students. *Journal of Computers in Education*, *6*(1), 5–19. https://doi.org/10.1007/s40692-018-0111-3

* Ratamun, M. M., & Osman, K. (2018). The Effectiveness of Virtual lab Compared to Physical lab in the Mastery of Science Process Skills for Chemistry Experiment. *Problems of Education in the 21st Century*, *76*(4), 544–560. https://doi.org/10.33225/pec/18.76.544

Rigby, S., & Ryan, R. M. (2011). Glued to Games: How Video Games Draw us in and Hold us Spellbound. *Praeger*, 510-511.

Riley, R. D., Higgins, J. P. T., & Deeks, J. J. (2011). Interpretation of Random Effects Meta-Analyses. *BMJ*, *342*, 964–967. https://doi.org/10.1136/bmj.d549

Rosenthal, R. (1979). The „File Drawer Problem“ and Tolerance for Null Results. *Psychological Bulletin, 86*(3), 638–641. https://doi.org/10.1037/0033-2909.86.3.638

Rowe, J. P., Shores, L. R., Mott, B. W., & Lester, J. C. (2011). Integrating Learning, Problem Solving, and Engagement in Narrative-Centered Learning Environments. *International Journal of Artificial Intelligence in Education, 21*(1-2), 115–133.

Ryan, R. M., & Deci, E. L. (2000). Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. *American Psychologist, 55*(1), 68–78. https://doi.org/10.1037/0003-066X.55.1.68

Ryan, R. M., & Rigby, C. S. (2020). Motivational Foundations of Game-Based Learning. In J. L. Plass, R. E. Mayer & B. D. Homer (Eds.), *Handbook of Game-Based Learning* (pp. 153–176). MIT Press.

* 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

Sala, G., & Gobet, F. (2017). Does Chess Instruction Improve Mathematical Problem-Solving Ability? Two Experimental Studies With an Active Control Group. *Learning & behavior*, *45*(4), 414–421. https://doi.org/10.3758/s13420-017-0280-3

* Shum, A. K., Lai, E. S., Leung, W. G., Cheng, M. N., Wong, H. K., So, S. W., Law, Y. W., & Yip, P. S. (2019). A Digital Game and School-Based Intervention for Students in Hong Kong: Quasi-Experimental Design. *Journal of Medical Internet Research*, *21*(4), 1–13.

* Siew, N. M., Geofrey, J., & Lee, B. N. (2016). Students' Algebraic Thinking and Attitudes Towards Algebra: The Effects of Game-Based Learning Using Dragonbox 12+ App. *The Research Journal of Mathematics and Technology*, *5*(1), 66–79.

Sitzmann, T. (2011). A Meta-Analytic Examination of the Instructional Effectiveness of Computer-Based Simulation Games. *Personnel Psychology*, *64*(2), 489–528. https://doi.org/10.1111/j.1744-6570.2011.01190.x

Skarbez, R., Smith, M., & Whitton, M. C. (2021). Revisiting Milgram and Kishino's Reality-Virtuality Continuum. *Frontiers in Virtual Reality*, *2*, 1–8. https://doi.org/10.3389/frvir.2021.647997

* Srisawasdi, N., & Panjaburee, P. (2019). Implementation of Game-Transformed Inquiry-Based Learning to Promote the Understanding of and Motivation to Learn Chemistry. *Journal of Science Education and Technology*, *28*(2), 152–164. https://doi.org/10.1007/s10956-018-9754-0

* Stapinski, L. A., Reda, B., Newton, N. C., Lawler, S., Rodriguez, D., Chapman, C., & Teesson, M. (2018). Development and Evaluation of 'Pure Rush': An Online Serious Game for Drug Education. *Drug and Alcohol Review*, *37*(1), 420–428. https://doi.org/10.1111/dar.12611

Steinkuehler, C., & Tsaasan, A. M. (2020). Sociocultural Foundations of Game-Based Learning. In J. L. Plass, R. E. Mayer, & B. D. Homer (Hrsg.), *Handbook of Game-Based Learning* (pp. 177–208). MIT Press.

Sweller, J. (2011). Cognitive Load Theory. In J. P. Mestre & B. H. Ross (Eds.), *Psychology of Learning and Motivation* (Vol. 55, pp. 37–76). Academic Press. https://doi.org/10.1016/B978-0-12-387691-1.00002-8

Tam, F., & Pawar, S. (2020). Emerging Design Factors in Game-Based Learning: Incentives, Social Presence, and Identity Design. In J. L. Plass, R. E. Mayer & B. D. Homer (Eds.), *Handbook of Game-Based Learning* (pp. 367–386). MIT Press.

* Tan, J. L., Goh, D. H.-L., Ang, R. P., & Huan, V. S. (2016). Learning Efficacy and User Acceptance of a Game-Based Social Skills Learning Environment. *International Journal of Child-Computer Interaction*, *9*, 1–19. https://doi.org/10.1016/j.ijcci.2016.09.001

Tangsripairoj, S., Sukkhet, M., Sumanotham, J., & Yusuk, B. (2019, July). Kiddy Manner: A Game-Based Mobile Application for Children Learning Thai Social Etiquette. In *2019 16th*

International Joint Conference on Computer Science and Software Engineering (JCSSE) (pp. 109–114). IEEE. https://doi.org/10.1109/JCSSE.2019.8864177

* Tao, S.-Y., Huang, Y.-H., & Tsai, M.-J. (2016). Applying the Flipped Classroom With Game-Based Learning in Elementary School Students' English Learning. *International Conference on Educational Innovation through Technology (EITT)*, 59–63. https://doi.org/10.1109/EITT.2016.19

Theng, Y.-L., & Aung, P. (2012). Investigating Effects of Avatars on Primary School Children's Affective Responses to Learning. *Journal on Multimodal User Interfaces*, *5*(1), 45–52. https://doi.org/10.1007/s12193-011-0078-0

* Tsai, C.-Y., Lin, H.-S., & Liu, S.-C. (2020). The Effect of Pedagogical GAME Model on Students' PISA Scientific Competencies. *Journal of Computer Assisted Learning*, *36*(3), 359–369. https://doi.org/10.1111/jcal.12406

van der Meij, H., van der Meij, J., & Harmsen, R. (2015). Animated Pedagogical Agents Effects on Enhancing Student Motivation and Learning in a Science Inquiry Learning Environment. *Educational Technology Research and Development*, *63*(3), 381–403. https://doi.org/10.1007/s11423-015-9378-5

* van der Ven, F., Segers, E., Takashima, A., & Verhoeven, L. (2017). Effects of a Tablet Game Intervention on Simple Addition and Subtraction Fluency in First Graders. *Computers in Human Behavior*, *72*, 200–207. https://doi.org/10.1016/j.chb.2017.02.031

van Reijmersdal, E. A., Jansz, J., Peters, O., & van Noort, G. (2013). Why Girls go Pink: Game Character Identification and Game-Players' Motivations. *Computers in Human Behavior*, *29*(6), 2640–2649. https://doi.org/10.1016/j.chb.2013.06.046

* 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*, 1–47. https://doi.org/10.1016/j.compedu.2019.103680

Veenman, M. V. J., Van Hout-Wolters, B. H. A. M., & Afflerbach, P. (2006). Metacognition and Learning: Conceptual and Methodological Considerations. *Metacognition and Learning*, *1*(1), 3–14. https://doi.org/10.1007/s11409-006-6893-0

Veltri, N. F., Krasnova, H., Baumann, A., & Kalayamthanam, N. (2014). Gender Differences in Online Gaming: A Literature Review. *Social-Technical Issues and Social Inclusion.*

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.-L. (2010). Outlier and Influence Diagnostics for Meta-Analysis. *Research Synthesis Methods*, *1*(2), 112–125. https://doi.org/10.1002/jrsm.11

Vogel, J. J., Vogel, D. S., Cannon-Bowers, J., Bowers, C. A., Muse, K., & Wright, M. (2006). Computer Gaming and Interactive Simulations for Learning: A Meta-Analysis. *Journal of Educational Computing Research*, *34*(3), 229–243. https://doi.org/10.2190/FLHV-K4WA-WPVQ-H0YM

Wigfield, A., & Eccles, J. S. (2000). Expectancy–Value Theory of Achievement Motivation. *Contemporary Educational Psychology*, *25*(1), 68–81. https://doi.org/10.1006/ceps.1999.1015

Wouters, P., van Nimwegen, C., van Oostendorp, H., & van der Spek, E. D. (2013). A Meta-Analysis of the Cognitive and Motivational Effects of Serious Games. *Journal of Educational Psychology*, *105*(2), 249–265. https://doi.org/10.1037/a0031311

Wrzesien, M., Rodríguez, A., Rey, B., Alcañiz, M., Baños, R. M., & Vara, M. D. (2015). How the Physical Similarity of Avatars can Influence the Learning of Emotion Regulation Strategies in Teenagers. *Computers in Human Behavior*, *43*, 101–111. https://doi.org/10.1016/j.chb.2014.09.024

Table 1

Inter-Rater Reliabilities for Coded Moderator Variables

Variable	κ
Game type	0.46
Mode	0.43 (0.62)
Additional non-game instruction	0.48
Dimensionality	0.47
Visual realism	0.37 (0.53)
Competition	0.79
Narration	0.71
Avatar	0.58

Digital agent	0.52
Mean	0.53 (0.57)

Note. κ = Cohen's Kappa. Inter-rater reliability after consulting a third rater in brackets.

Figure 1

The Integrated Design Framework of Playful Learning (adapted from Plass et al., 2015)

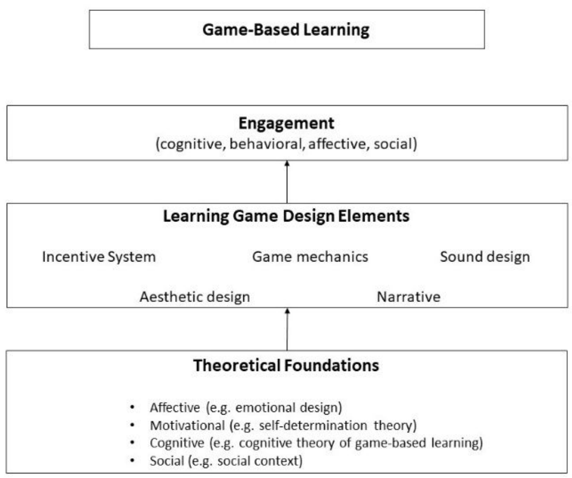

Figure 2

Flow Chart of the Selection Process