RESEARCH ARTICLE

Gamification enhances student intrinsic motivation, perceptions of autonomy and relatedness, but minimal impact on competency: a meta-analysis and systematic review

Liuyufeng Li[1] · Khe Foon Hew[2] · Jiahui Du[1]

Abstract

Although many studies in recent years have examined the use of gamification as a motivational strategy in education, evidence regarding its effects on intrinsic motivation is inconsistent. To make the case for or against the adoption of gamification in education, this study examines its effects on students' intrinsic motivation and the underlying motivational factors: perceived competence, autonomy, and relatedness. In this review, we analyzed the results of studies comparing gamified learning with non-gamified learning published between 2011 and 2022. The results of our meta-analysis of 35 independent interventions (involving 2500 participants) indicated an overall significant but small effect size favoring gamified learning over learning without gamification (Hedges' $g = 0.257$, 95% CI [0.043, 0.471], $p = .019$) with no evidence of publication bias. Gamification also exerted a positive and significant effect on the students' perceptions of autonomy (Hedges' $g = 0.638$, 95% CI [0.139, 1.136], $p = .012$) and relatedness (Hedges' $g = 1.776$, 95% CI [0.737, 2.814], $p = .001$), but minimal impact on competence (Hedges' $g = 0.277$, 95% CI [0.001, 0.553], $p = .049$). To further investigate the possible reasons for the small impact on intrinsic motivation, a systematic review of 31 studies was conducted. The findings revealed two major challenges encountered in the adoption of gamification to increase students' intrinsic motivation: students' lack of perceived competence and lack of perceived autonomy in gamified classes.

Keywords Intrinsic motivation · Basic psychological needs · Gamification · Self-determination

✉ Khe Foon Hew
kfhew@hku.hk

[1] The University of Hong Kong, Hong Kong, China

[2] Faculty of Education, The University of Hong Kong, Room 111A, Runme Shaw Building, Hong Kong, China

Introduction

Intrinsic motivation is one of the most important psychological concepts in education research (Vallerand et al., 1992, p. 1004) and a key determinant of academic achievement (Ryan & Deci, 2020). When individuals are driven by intrinsic motivation, they actively engage in activities that interest them (Ryan & Deci, 2000a). It is therefore unsurprising that a substantial amount of research has been devoted to fostering students' intrinsic motivation (Cerasoli et al., 2014; Deci et al., 2001; Xu et al., 2021).

Studies have shown that students' intrinsic motivation in school-related activities tends to decline over the school year (Gnambs & Hanfstingl, 2016; Lepper et al., 2005; Ryan & Deci, 2020; Scherrer & Preckel, 2019). Researchers have argued that this decline in intrinsic motivation occurs because students' basic psychological needs are not being sufficiently satisfied in their education (Gnambs & Hanfstingl, 2016; Ryan & Deci, 2020). According to the self-determination theory, intrinsic motivation flourishes when an activity satisfies an individual's basic psychological needs (i.e., competence, autonomy, and relatedness) rather than when it is performed due to some separable external consequence such as pressure (Ryan & Deci, 2000c, 2002). Competence refers to the feeling of mastering a challenge, autonomy refers to one's volition in performing a task, while relatedness refers to a sense of connection with other people. However, traditional school settings do not always create need-supportive learning environments that foster students' fulfillment of basic psychological needs (Gnambs & Hanfstingl, 2016; La Guardia & Ryan, 2002; Raufelder & Kulakow, 2021; Ryan & Deci, 2020).

To reverse the abovementioned decline in intrinsic motivation, educational researchers have examined approaches to satisfy students' basic psychological needs and thus foster their intrinsic motivation. Gamification, which refers to providing students with exciting game-like experiences "in a non-game context" through the use of game elements such as points and badges (Deterding et al., 2011, p. 10), is a potential solution to this problem (e.g., Ryan & Deci, 2020; Xu et al., 2021). For example, game elements such as leaderboards may fulfill students' need for competence by visually showing their achievement with respect to other people; group competition may fulfill students' need for relatedness to a team; while providing students with various badge choices can help fulfill students' need for autonomy (Sailer et al., 2017). Gamification, therefore, has the potential to support the growth of students' intrinsic motivation (Xu et al., 2021).

Since its conceptualization around 2010, the body of empirical research on gamification in educational contexts has been steadily growing (Koivisto & Hamari, 2019). Unfortunately, findings on the effects of gamification on intrinsic motivation have been inconsistent (Hanus & Fox, 2015; Mekler et al., 2017; Sailer & Sailer, 2021; Sailer et al., 2017), and there is currently a lack of conclusive evidence regarding the effects of gamification on students' intrinsic motivation. The present study addressed this inconsistency in the literature by conducting a meta-analysis to offer insights into the effects of gamification on intrinsic motivation and how it can be used to foster students' intrinsic motivation. A meta-analysis that integrates the results of multiple intervention studies can provide more accurate estimates of an intervention's effects than a single study (Higgins et al., 2019). Furthermore, meta-analysis is superior to narrative synthesis, which in itself is not sufficient to synthesize conflicting results when numerous studies are involved (Hunter & Schmidt, 2004).

Several previous meta-analyses have reported that gamification positively affects students' motivation (Mula-Falcón et al., 2022; Ritzhaupt et al., 2021; Sailer & Homner, 2020; Zhang et al., 2021). However, a closer scrutiny of these published meta-analyses reveals

that they did not explicitly explain whether they focused on intrinsic or extrinsic motivation, or their motivation outcomes also included other constructs. For example, Sailer and Homner (2020) referred to motivational outcomes that included a host of constructs such as (intrinsic) motivation, dispositions, preferences, attitudes, engagement, confidence, and self-efficacy. Similarly, although Ritzhaupt et al. (2021) reported a positive and significant effect size for the influence of gamification on students' affective outcomes, the affective outcomes that they examined included not only motivation and interest but also learner self-efficacy, perceived learning, perceived ease of use, and attitude. This lack of clarity on the types of motivational outcomes investigated in prior studies has caused some confusion about the effect of gamification on intrinsic motivation and made it difficult to decide whether to use gamification in education.

This study contributes to the literature in two ways. First, we present a meta-analysis of quantitative intervention studies that focuses specifically on the effects of gamification on students' intrinsic motivation and the fulfillment of basic psychological needs (competence, autonomy, relatedness) in various educational settings. Second, this study explored the challenges encountered in using gamification to enhance intrinsic motivation. Understanding these challenges can help us gain deeper insight into how to use gamification to effectively improve students' intrinsic motivation.

Research questions

In this review, we first defined the concepts of intrinsic motivation and gamification based on the literature. The following are the research questions that guided the review:

RQ1. What instruments have been used to measure students' intrinsic motivation?
RQ2. What is the effect of gamification on students' intrinsic motivation?
RQ3. What is the effect of gamification on students' basic psychological needs (i.e., competence, autonomy, relatedness)?
RQ4. What are the current challenges of using gamification to enhance intrinsic motivation?

Conceptual and theoretical background

Motivation

Motivation has been a key focus of educational researchers, as it drives behavior and determines students' choices and their engagement, effort, and persistence in the learning process (Dörnyei & Ushioda, 2013). Motivation can be categorized into three types: (1) intrinsic motivation, which refers to the motivation to undertake activities "for their own sake" or for their intrinsic interest and enjoyment; (2) extrinsic motivation, which refers to the motivation to undertake activities for some separable outcome rather than intrinsic enjoyment; and (3) amotivation, which refers to a lack of intentionality that is common in the classroom and can be attributed to "either lack of felt competence to perform, or lack of value or interest" (see also Deci & Ryan, 2004; Ryan & Deci, 2020, p. 3).

Although both extrinsic and intrinsic motivation can contribute to student performance (Cerasoli et al., 2014), extrinsically motivated students—that is, students who do

a task for an external consequence (e.g., obtaining a reward or avoiding punishment)—are more likely to engage in surface and unsustainable learning (Lee et al., 2010), and extrinsic motivation in students may be associated with certain negative outcomes (Clanton Harpine, 2015). For example, once the external reward stops, such students may stop demonstrating the specific behavior. As Zichermann and Cunningham (2011, p. 27) stated, "once you start giving someone a reward, you have to keep her in that reward loop forever."

In contrast, intrinsic motivation is the more productive force behind any behavior (Deci & Ryan, 2000; Ryan & Deci, 2000d) because intrinsic motivation triggers an individual's inner drive to engage in activities based on their personal interests (Deci & Ryan, 2008). Intrinsically motivated students' engagement in learning activities is accompanied by positive effects such as improved mental health, enhanced creativity, and long-term learning outcomes (Ryan & Deci, 2000c), and intrinsic motivation increases the level of effort and quality of student input into a particular task (Cerasoli et al., 2014). In short, intrinsically motivated students—that is, students who find a task interesting—are more likely to persist in their learning and are more willing to voluntarily attempt different challenges (Deci & Ryan, 2004; Lee et al., 2010; Vansteenkiste et al., 2006).

To foster students' intrinsic motivation, many practitioners implement reward and incentive systems in educational settings (Cameron et al., 2001). However, there has long been a heated debate about the relationship between rewards and intrinsic motivation (e.g., Cameron et al., 2001; Deci et al., 1999). On the one hand, some researchers have suggested that providing extrinsic rewards for an initially enjoyable task can reduce individuals' subsequent intrinsic motivation for that task, because extrinsic rewards are designed to externally control a person (Greene, 2018). On the other hand, other researchers have argued that extrinsic rewards have the potential to maintain or enhance participants' intrinsic motivation depending on the types of rewards offered (tangible or verbal) and the types of reward contingencies adopted (e.g., Cameron & Pierce, 1994; Cameron et al., 2001; Cerasoli et al., 2014; Deci et al., 1999).

Verbal rewards refer to expressions of recognition or praise (Cameron & Pierce, 1994; Deci et al., 2001) that are delivered in either verbal or written form (Hewett & Conway, 2016). Meta-analytic evidence has suggested that verbal rewards have a more positive effect on free-choice motivation and self-reported intrinsic motivation than tangible rewards particularly for high-interest tasks (e.g., Cameron & Pierce, 1994; Cameron et al., 2001; Deci et al., 1999). Nevertheless, not all tangible rewards (such as gift cards or gold stars) have a negative effect on intrinsic motivation, and their effect depends on the type of reward contingency implemented (e.g., Cameron & Pierce, 1994; Cameron et al., 2001). Cameron et al. (2001) classified reward contingencies into seven types (see Table 1) and conducted a meta-analysis to investigate the effect of various types of reward contingencies on intrinsic motivation. Their overall conclusion was that offering extrinsic rewards for low-interest tasks can enhance free-choice intrinsic motivation and leave task interest unaffected. Furthermore, extrinsic rewards either enhance or do not harm free-choice and self-reported intrinsic motivations when the rewards are explicitly linked to performance (the rewards may be linked to performance based on an absolute standard, such as exceeding a specified score, or a relative standard, such as surpassing others' scores). Rewards that are tied to performance can enhance an individual's perceived competence (see next section for an in-depth discussion of competence), and a greater sense of competence can lead to higher interest in a task (Cameron et al., 2001).

Table 1 Types of reward contingencies (Cameron et al., 2001, p. 12)

Reward contingency	Description
Task noncontingent	Reward is offered for agreeing to participate, for coming to the study, or for waiting for the experimenter
Rewards offered for doing well	Reward is offered for doing well on the task or for doing a good job No specification is given as to what it means to do a good job or to do well
Rewards offered for doing a task	Reward is offered to engage in the experimental activity No instructions are given about how well participants must perform or whether they must complete the task
Rewards offered for finishing or completing a task	Reward is offered to finish an activity, to complete a task, or to get to a certain point on the task The reward is not related to quality of performance
Rewards offered for each unit solved	Reward is offered for each unit, puzzle, problem, etc., that is solved
Rewards offered for surpassing a score	Reward is offered for surpassing a particular specified score (absolute standard) In some cases, the better the score, the higher the reward
Rewards offered for exceeding a norm	Reward is offered to meet or exceed the performance of others on the task (relative standard)

Self-determination theory as a framework

Self-determination theory (SDT), the dominant theory of intrinsic motivation, explains how the environment promotes intrinsic motivation. It posits that individuals' intrinsic motivation is enhanced in environments in which they are able or have the opportunity to perceive autonomy, competence, and relatedness (Deci & Ryan, 1985, 2004). Several studies have also demonstrated that the satisfaction of one's competence needs (Fransen et al., 2018), autonomy needs (Karimi & Sotoodeh, 2020), and relatedness needs (Xiang et al., 2017) increases their intrinsic motivation. These needs and their relationship to intrinsic motivation can be summarized as follows:

(1) Competence refers to the feeling of mastering a challenge and flourishes when direct and positive (informative) feedback is received (Deci & Ryan, 2004). The positive effects of perceived competence on intrinsic motivation typically occur when it is accompanied by a sense of autonomy (Deci & Ryan, 2004).

(2) Autonomy refers to psychological freedom and the volition to perform tasks (Deci & Ryan, 2000, p. 231; Van den Broeck et al., 2010; Vansteenkiste et al., 2010). The sense of making decisions based on one's interests is the expression of psychological freedom (Deci & Ryan, 2012; Ryan & Deci, 2002), whereas volition is the sense of acting with no external pressure or coercion (Vansteenkiste et al., 2010). When a person perceives a sense of autonomy, they show more interest in an activity and greater confidence in engaging in it, which enhances performance and increases persistence (Ryan & Deci, 2000d).

(3) Relatedness refers to a sense of belonging and connection (Ryan & Deci, 2020). It represents an individual's underlying desire for integration into the social environment (Baumeister & Leary, 1995; Deci & Ryan, 2000, 2004). When individuals form

intimate relationships and feel a sense of communion with others, they perceive greater levels of relatedness (Deci & Ryan, 2000). In environments characterized by a sense of relatedness, intrinsic motivation is more likely to thrive (Ryan & Deci, 2000d; Ryan & La Guardia, 2000).

Gamification

Gamification is often depicted as being different from entertainment games and serious games (Bai et al., 2020). Games are usually developed for entertainment purposes (e.g., World of Warcraft), while serious games, also known as game-based learning (Boyle et al., 2016) and are developed to train certain skills or learn academic content (Annetta, 2010). In both entertainment games and serious games, the development of the game product typically requires a significant amount of money and effort. In contrast, gamification does not involve the development of a game product but rather the application of game elements to motivate participants' behaviors in non-gaming contexts (Educause, 2011).

Although game elements are the fundamental building blocks of gamification, there is no commonly acknowledged classification of game elements (Bai et al., 2020). Various authors have proposed their own classification schemes (e.g., Deterding et al., 2011; Dicheva et al., 2015; Zichermann & Cunningham, 2011). Although these schemes are distinct, several common game elements can be identified across them, including levels, narratives or storytelling, competition, badges, leaderboards, and points (see Ritzhaupt et al., 2021 for details).

Gamification and intrinsic motivation

The use of gamification in a learning context is referred to as gamified learning (Armstrong & Landers, 2017; Landers, 2014). The integration of game elements and learning activities into gamified learning can potentially increase students' intrinsic motivation by making learning activities enjoyable and satisfying (Koivisto & Hamari, 2019).

Viewed from the perspective of SDT, gamified learning has the potential to help students satisfy their basic psychological needs of competence, autonomy, and relatedness (Deterding, 2012; Przybylski et al., 2010):

(1)   As competence refers to the feeling that one is succeeding when interacting with the environment (Rigby & Ryan, 2011; Vansteenkiste & Ryan, 2013), feedback mechanisms in gamified learning can help satisfy students' needs for competence. For example, feedback mechanisms such as points, medals, and leaderboards can visually communicate students' achievements and competence (Xi & Hamari, 2019). In addition, to motivate students effectively, tasks in gamified learning should be designed so that they are not easy but just outside the comfort zone of the students at a level of difficulty they find achievable (Roy & Zaman, 2017). When tasks are at such a level of difficulty, students persist in improving themselves to accomplish them (Deci & Ryan, 1985; Peng et al., 2012).

(2)   As autonomy refers to a person's sense of freedom in their actions (Ryan & Deci, 2020), providing students with choice can help satisfy students' needs for autonomy. For example, Jones et al. (2022) addressed students' need for autonomy by providing multiple options for assignments that the students could engage in, which allowed them to choose their own path to achieve their desired outcomes (grades). The results

demonstrated that the students who participated in gamified learning had higher perceived autonomy and intrinsic motivation than those who participated in non-gamified learning.

(3) As relatedness refers to a person's sense of belonging to a group (Ryan & Deci, 2017), frequent communication and idea-sharing via group work in gamified learning can help learners perceive relatedness (Fernandez-Rio et al., 2021). Furthermore, group competition can create a sense of belonging to a team by reinforcing the sense of community (van Roy & Zaman, 2019).

Nevertheless, depending on how it is implemented, gamified learning may be ineffective at enhancing students' intrinsic motivation and even lead to negative consequences such as negative emotions and poor learning outcomes (Hanus & Fox, 2015; Mekler et al., 2017; Mitchell et al., 2017). Whereas rewards such as points and badges can reinforce extrinsically motivated behavior, they may shift the focus of students to the rewards rather than the learning process (Gladun, 2016). Leaderboards may also have unintended adverse effects as they may increase the sense of embarrassment for students in low positions of the leaderboards (Bai et al., 2020).

In short, although gamified learning, through the use of game elements such as badges, social interactions, points, and leaderboards in online learning environments, has the potential to increase intrinsic motivation (Xu et al., 2021), one of the most pressing challenges in this field is that there is little consensus on whether gamification actually improves intrinsic motivation. Empirical studies have reported mixed results, with some studies reporting positive effects (Fernandez-Rio et al., 2021; Sailer & Sailer, 2021; Segura-Robles et al., 2020) and others finding no effects or even negative effects (Hanus & Fox, 2015; Jones et al., 2022; Mekler et al., 2017; Tasadduq et al., 2021).

Methods

Search strategy

The Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) statement guided the procedure for choosing relevant studies (Moher et al., 2009). We searched 10 electronic databases that were likely to contain relevant and high-quality papers: ACM Digital Library, Emerald Insight, EBSCO host research databases (including Academic Search Complete, British Education Index, ERIC, and Education Full Text), IEEE, ProQuest, Scopus, Springer, Sage Journals, Science Direct, and Web of Science. We did not restrict the language of instruction or the location of the studies in our search, but the studies had to be reported in English. In addition, to widen the range of studies considered, we did not restrict our sources of papers to peer-reviewed journals and instead also included conference papers and dissertations.

Search string

The Boolean operators AND and OR were used to gather as many relevant papers as possible. Asterisks were used to capture a variety of common terms used in gamified learning. Three sets of search terms were used for this review. The first set consisted of possible gamification terms; we used the expression "gamif*" to cover all

morphological variations of "gamification," "gamified," and "gamify." The second set of search terms contained terms related to intrinsic motivation. We used the expression "intrinsic motiva*" to cover all morphological variations of "intrinsic motivation," "intrinsic motivated," and "intrinsically motivated." The third set contained terms that were related to course, classroom, education, or learning. The following is the search string that was used: gamif* AND intrinsic motiva* AND (course OR class* OR educat* OR learn*).

Selection criteria

Empirical studies published between January 2011 and October 2022 (11 years) were considered for this review as the concept of gamification was defined in 2011 (Deterding et al., 2011). To examine the possible effect of gamification on students' intrinsic motivation and to detail the challenges in current gamification research, a meta-analysis and a systematic review of relevant articles were respectively conducted. The criteria that were used to select articles for the meta-analysis and the systematic review are shown in Table 2.

For the purpose of meta-analysis, we excluded studies that contained mere descriptions of gamification without presenting any empirical data. For the purpose of systematic review, we excluded studies that did not explicitly present the qualitative findings regarding students' perceptions of whether and how gamified learning affected intrinsic motivation.

Table 2 Criteria for article selection

Meta-analysis
(a) The study reported an empirical examination of gamified practices using at least one clearly described game element
(b) The study was conducted in a K-12 or higher education setting and written in English
(c) The study contained at least one comparison of motivational outcomes between a gamified class and a non-gamified class
(d) The study measured students' intrinsic motivation in both gamified and non-gamified classes for the same course topic using a survey such as the Intrinsic Motivation Inventory or a scale measuring free-choice behavior. Such surveys have been frequently used by researchers to assess participants' motivation during an intervention
(e) The study reported sufficient data such as means, standard deviations, sample sizes, t values of t-tests, F values, and z scores of Mann–Whitney U tests
(f) The study was published in a peer-reviewed journal or was a conference paper or thesis/dissertation. The study had to be accessible via a library subscription or freely available online

Systematic review
(a) The study reported on empirical gamified learning practices using at least one explicitly described game element
(b) The study was conducted in a K-12 or higher education setting
(c) The study involved data collected through student interviews and/or open-ended survey questions
(d) The study was published in a peer-reviewed journal or was a conference paper or thesis/dissertation. The study had to be accessible via a library subscription or freely available online

Study selection

As of October 2022 (the time of writing), the use of the search string to obtain relevant articles from the databases resulted in the retrieval of 3125 articles. Among these, 195 articles were removed as they were duplicate entries. 32 Articles were removed by the automated tool provided in the academic database because they failed to fulfill one or more of the following criteria: (a) all studies were published between 2011 and 2022; (b) all studies focused on K-12 or higher education; (c) all selected studies were peer-reviewed articles; and (d) all selected studies were written in English. Two additional articles were identified using the snowball method, which was performed by scanning the references of the relevant articles. Although our search string enabled us to capture a variety of terms used to refer to gamified classes, it also returned many unrelated articles (e.g., research about game-based learning). Therefore, after scanning their titles and abstracts, numerous articles that were irrelevant to this review were eliminated. During this scanning process, we were aware that some studies may have evaluated certain aspects of students' intrinsic motivation without explicitly mentioning them in the title or abstract. Therefore, we also scanned the articles' sections and subsections. We carefully examined each comparison item and determined through mutual discussion whether the item referred to intrinsic motivation or to factors that trigger intrinsic motivation. For example, we categorized the "interest/enjoyment" subscale of the Intrinsic Motivation Inventory (Deci & Ryan, 2022) as describing intrinsic motivation, because this subscale mainly emphasizes interest and enjoyment (e.g., "I enjoyed doing this activity very much").

After completing this process, 132 full-text articles were determined to be eligible. These articles reported comparisons of non-gamified classes with gamified classes. However, about three quarters of them were then eliminated because they did not compare any aspect of students' intrinsic motivation between the two instructional environments, yielding 31 articles of which 24 provided sufficient data for the meta-analysis. Facey-Shaw et al. (2020), Brom et al. (2019), and Leitão et al. (2022) reported more than one gamified classroom intervention. Consequently, we obtained 35 unique gamified classroom interventions, involving 2500 participants in total. Therefore, we effectively covered 35 intervention studies (in each of which one gamified intervention group was compared with a control group) in this review. Figure 1 outlines the article selection process.

Data analysis

Data extraction

The following key information was extracted from each article: (a) sample size; (b) geography; (c) school level; (d) intervention duration; (e) control factors used in the study (instructor equivalence values and student equivalence values); (f) discipline; (g) measurement instrument; (h) game element; (i) type of reward and reward contingencies; and (j) statistical results (e.g., mean and standard deviation). One coder independently extracted the information from all 24 articles, and another coder randomly selected 90% of the articles for independent coding to test the reliability of the extracted information. Inter-coder reliability was 98%. Coding differences were resolved by the coders through discussion.
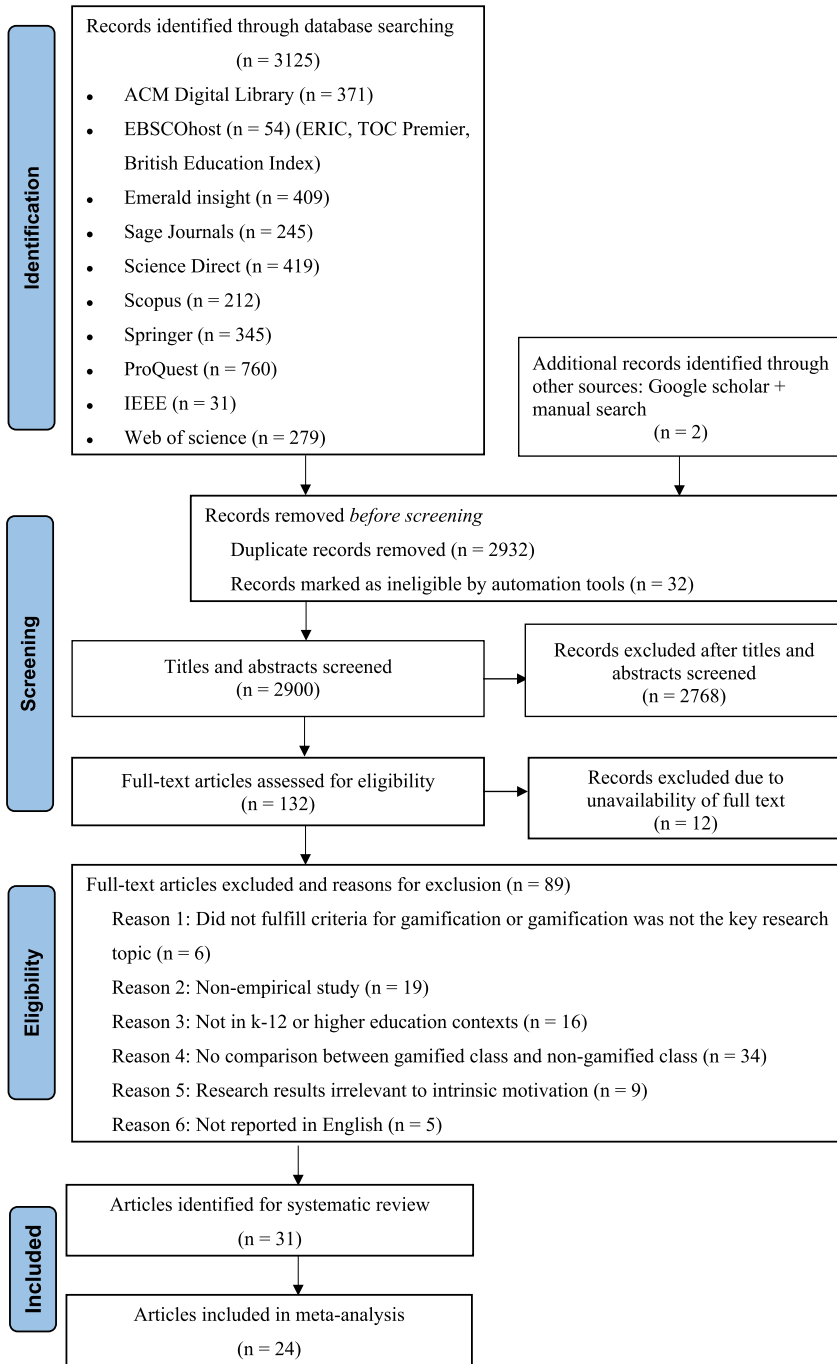
Records identified through database searching
(n = 3125)

- ACM Digital Library (n = 371)
- EBSCOhost (n = 54) (ERIC, TOC Premier, British Education Index)
- Emerald insight (n = 409)
- Sage Journals (n = 245)
- Science Direct (n = 419)
- Scopus (n = 212)
- Springer (n = 345)
- ProQuest (n = 760)
- IEEE (n = 31)
- Web of science (n = 279)

Additional records identified through other sources: Google scholar + manual search
(n = 2)

Identification

Records removed *before screening*
Duplicate records removed (n = 2932)
Records marked as ineligible by automation tools (n = 32)

Screening

Titles and abstracts screened
(n = 2900)

Records excluded after titles and abstracts screened
(n = 2768)

Full-text articles assessed for eligibility
(n = 132)

Records excluded due to unavailability of full text
(n = 12)

Full-text articles excluded and reasons for exclusion (n = 89)

Reason 1: Did not fulfill criteria for gamification or gamification was not the key research topic (n = 6)

Reason 2: Non-empirical study (n = 19)

Reason 3: Not in k-12 or higher education contexts (n = 16)

Reason 4: No comparison between gamified class and non-gamified class (n = 34)

Reason 5: Research results irrelevant to intrinsic motivation (n = 9)

Reason 6: Not reported in English (n = 5)

Eligibility

Articles identified for systematic review
(n = 31)

Included

Articles included in meta-analysis
(n = 24)

Fig. 1 PRISMA flow diagram for article selection

Computing effect sizes

The Comprehensive Meta-Analysis software version 4 (Biostat, Inc., Englewood Cliffs, NJ) was used to calculate effect sizes. Unless otherwise stated, a p-value of 0.05 or less is considered statistically significant (Baker, 2016), and all p-values reported here are two-tailed. Effect sizes were calculated using a random effects model (Gurevitch & Hedges, 1999) to shed light on inter-study variation (Raudenbush, 2009).

Hedges' *g* is useful in meta-analyses of studies with varying sample sizes because it is the corrected standardized mean difference between two groups based on the pooled standard deviation (Korpershoek et al., 2016). For studies in which means and standard deviations were not provided, standardized mean differences were computed using other sources of information such as F-values, t-tests, and p-values (Borenstein et al., 2021; Rosenthal & DiMatteo, 2001). If standard errors were used instead of standard deviations in a study, we calculated the standard deviation using the following formula (Altman & Bland, 2005):

$$SE = \frac{SD}{\sqrt{sample\ size}}$$

To satisfy the hypothesis of effect size independence for independent student samples, the effect size was computed for each study (Scammacca et al., 2014). However, if a study reported effect sizes for multiple student groups (e.g., Group a, Group b, and Group c) and the groups did not overlap, then the effect size for each group was included in the meta-analysis because each group represented an independent sample (Lipsey & Wilson, 2001). When this was done for a study, we verified that both coders concurred that the multiple student groups in the study were completely independent.

Furthermore, it is not always possible to code all items from the studies included in a meta-analysis, as some studies may not report the results for all items they administered (Lipsey & Wilson, 2001). Although some researchers may discard such studies with missing data, doing so is not ideal because the findings arrived at based only on studies that reported all items under consideration may be misleading (Lipsey & Wilson, 2001). To reduce the potential for misleading results, we included a "not reported" option for items in the coding protocol as proposed by Lipsey and Wilson (2001). We also conducted a moderator analysis to determine whether there were differences between the findings of studies with and without missing data.

Analysis of heterogeneity

The I^2 test was used to detect the presence of heterogeneity across the samples. According to Shamseer et al. (2015), q-values of 0–40% indicate that heterogeneity is likely to be non-significant, 30–60% indicate moderate heterogeneity, 50% or more indicate large heterogeneity, and 75–100% indicate substantial heterogeneity.

To identify the causes of potential differences in effect sizes across the samples, we conducted moderator analyses. The moderating variables were classified into six main categories based on prior studies (Bai et al., 2020; Cameron & Pierce, 2002; Cameron et al., 2001; Chen et al., 2018; Landers et al., 2014; Zheng et al., 2016), namely, participant characteristics (school level and geography), course characteristics (sample size, study design, and intervention duration), control level (student equivalence and instructor equivalence), number of game elements, and the type of reward and reward contingency implemented:

(1) Participant characteristics. Aimed at analyzing whether there were any differences between participants across school levels and geographic regions.

(2) Course characteristics. Aimed at analyzing whether there were any differences across course subjects and whether the durations and sample sizes of the experimental interventions influenced the final effect size (Chen et al., 2018; Zheng et al., 2016). To ensure a precise analysis, we used a sample size coding scheme adapted from Chen et al. (2018) and a time coding scheme from Bai et al. (2020) (see Table 3).

(3) Control level. Aimed at determining whether the various control levels implemented in the interventions may have affected the final effect size (Freeman et al., 2014). According to Bai et al. (2020), two control levels can be considered: student equivalence and instructor equivalence. A study can be categorized into one of three types based on student equivalence: (1) no-significant-difference group, i.e., the study conducted and reported an initial statistical assessment of the control and experimental groups, and showed that the students in the two groups were initially at the same level in a statistically significant manner; (2) significant-difference group, i.e., the results of the initial statistical assessment showed that the initial levels of the students were different between the groups; and (3) no data reported, i.e., the study did not provide statistical data on whether the initial levels of the students were equivalent. Similarly, a study can be categorized into one of three types based on instructor equivalence: (1) identical instructor, i.e., the same instructor oversaw the treatment and control groups; (2) different instructor, i.e., two or more instructors oversaw the treatment and control groups respectively; and (3) no data reported, i.e., the authors did not provide this information about the instructors.

(4) Number of game elements. Aimed at investigating whether the number of game elements used moderated the effect size.

(5) Type of reward and reward contingency. Aimed at determining whether the type of reward and reward contingency moderated the effect sizes. Rewards were categorized into verbal rewards (i.e., praise or positive feedback) and tangible rewards (e.g., sweets, toys, and badges). We coded the reward contingencies using the reward contingency framework developed by Cameron et al. (2001) (see Table 1).

Following Cameron and Pierce (2002) studies that did not provide enough information to code specific characteristics of reward were omitted from the reward contingency analyses. In addition, subgroups with only one intervention were excluded from the moderator analyses, as suggested by Tondello et al. (2017). This is because the number of interventions was considered too small to yield meaningful results (Bai et al., 2020; Fu et al., 2011).

Table 3 Coding scheme (Sample size and Intervention duration)

	1	2	3	4	5	6
Sample size	< 50	50–100	100–150	≥ 150	No data provided	
Intervention duration	< 1 week	1 week–1 month	1 month–3 months	3 months–1 semester	≥ 1 semester	No data provided

Analysis of publication bias

One cause of publication bias is that researchers tend to only report favorable results (i.e., significant results), which may lead to overestimation of the effects of interventions (Borenstein et al., 2021). We conducted four analyses to assess publication bias: the funnel plot, the classic fail-safe N test, and the calculation of Egger's regression and Begg and Mazumdar's rank correlation.

Qualitative analysis

To further analyze the possible challenges of using gamification to facilitate students' intrinsic motivation, the self-determination theory of motivation (SDT) was used to construct our coding framework. Preliminary coding of several studies indicated that SDT elements (i.e., competence, relatedness, and autonomy) can address the challenges of employing gamification to foster intrinsic motivation. More specifically, the first author developed the coding scheme based on previous gamification literature (e.g., Bai et al., 2020; Yaşar et al., 2020) and preliminary coding of several empirical studies identified in the present review (Table 4). The first author coded all the articles using the constant-comparative approach (Lincoln & Guba, 1985) based on the coding scheme. The second author independently coded 20% of the articles using the same coding scheme. Although we used SDT as a priori, we remained open to the identification of new categories (if any) during the coding process. We also did not force any data into a particular category. Any discrepancies were resolved through mutual discussion.

The following example illustrates how the data were analyzed and coded: "Some students also reported being unaware of the badges. This may lead to the statistically significant drop in scores for perceived competence" (Facey-Shaw et al., 2020, p. 46). This example was coded as an "unfamiliarity with game elements" subcategory since the most salient element appeared to be students being unacquainted with the game elements. The "unfamiliarity with game elements" subcategory is subsumed within the main category of "lack of perceived competence". This is because the game elements (e.g., badges) were linked to student's individual competence in solving a task. Therefore, students who were unacquainted with the game elements were less likely to feel a sense of competence.

Analysis of the data corpus continued until each coding category was saturated, which means that new data began to confirm rather than shed new light on the types of challenge categories.

Results

Characteristics of the studies

Thirteen (42%) of the 31 studies were conducted in Europe (e.g., Brom et al., 2019; Ferriz-Valero et al., 2020; Garcia-Cabot et al., 2020; Jurgelaitis et al., 2019; Kyewski & Krämer, 2018) and 9 (29%) were conducted in America (e.g., Challco et al., 2019; Hanus & Fox, 2015; Hazan et al., 2018). Three (10%) were conducted in the Asia–Pacific region, specifically, China (Sun & Hsieh, 2018), Malaysia (Hong & Masood, 2014), and Pakistan (Tasadduq et al., 2021). The remaining six studies (e.g., De Schutter & Abeele, 2014; Facey-Shaw

Table 4 Self-determination theory-based coding scheme to code the challenges of using gamification

Challenges related to competence	Challenges pertaining to relatedness	Challenges pertaining to autonomy
Failure to achieve a sense of mastery, a feeling of success	Failure to achieve a sense of connection with other people	Failure to attain a sense of initiative and ownership of individual actions
• Discomfort among the lowest ranked participants in a public absolute leaderboard • Unsuitability of the difficulty level of the gamified tasks • Unfamiliarity with gamification elements	• Feelings of jealousy among peers (Bai et al., 2020) • Grudge towards peers (Yaşar et al., 2020)	• Lack of autonomy in choosing learning content and activities

et al., 2020; Stansbury & Earnest, 2017) did not specify where the interventions were conducted. This review included studies conducted at both the K-12 and higher education levels. Twenty-one of the studies were conducted at the higher education level (undergraduate: n = 20, graduate: n = 1) and 10 studies were conducted at the K-12 level (primary school: n = 2, secondary school: n = 6, junior high school: n = 2). The subjects covered varied between the studies, including physical education (e.g., Fernandez-Rio et al., 2021; Segura-Robles et al., 2020), algorithms (e.g., Rodrigues et al., 2021), and mathematics (e.g., Stoyanova et al., 2017). In most of the studies (n = 12), the interventions were between 2 and 12 weeks long, and in seven studies the interventions were longer than 16 weeks (1 semester). The duration of seven interventions was less than 1 week. Five studies did not provide explicit information about the duration of their interventions.

RQ1. What instruments have been used to measure students' intrinsic motivation in the gamified classroom approach?

Self-report questionnaires

We found that 29 of the studies used self-report survey measures (e.g., the interest/enjoyment scale of the IMI and the intrinsic motivation subscales of the Basic Psychological Needs in Exercise Scale and the Sport Motivation Scale) to assess the students' intrinsic motivation. The self-report measures used were generally not discipline-specific with the exception of a few used to assess intrinsic motivation in specific domains such as mathematics (Stoyanova et al., 2017) or physical education (Ferriz-Valero et al., 2020).

Overall, self-report methods were the most frequently used methods to assess intrinsic motivation. Students' intrinsic motivation can also be inferred from behavior such as voluntary re-engagement with a task, which indicates the resumption of an activity without instruction or compulsion over a free-choice period (Ryan & Deci, 1987; Ryan et al., 1991). However, to the best of our knowledge, none of the studies included in this review assessed intrinsic motivation by monitoring the students' free-choice behavior.

Interview

A few of the studies (n = 5) used interviews to assess the students' intrinsic motivation. Most of the interviews used a semi-structured approach in which the participants were interviewed in an open, unstructured setting based on pre-designed questions with a focus on listening to the students' stories. One of the advantages of this approach is that it enables the interviewer to gain insight into the causes of variations in intrinsic motivation across students in a relaxed atmosphere and to understand why some students' intrinsic motivation may have increased while other students' intrinsic motivation may have shown no change or even decreased.

RQ2. What is the effect of gamification on students' intrinsic motivation?

Overall effect size

In this meta-analysis, 35 independent interventions involving 2500 participants were examined. The overall effect of gamification on the students' intrinsic motivation was
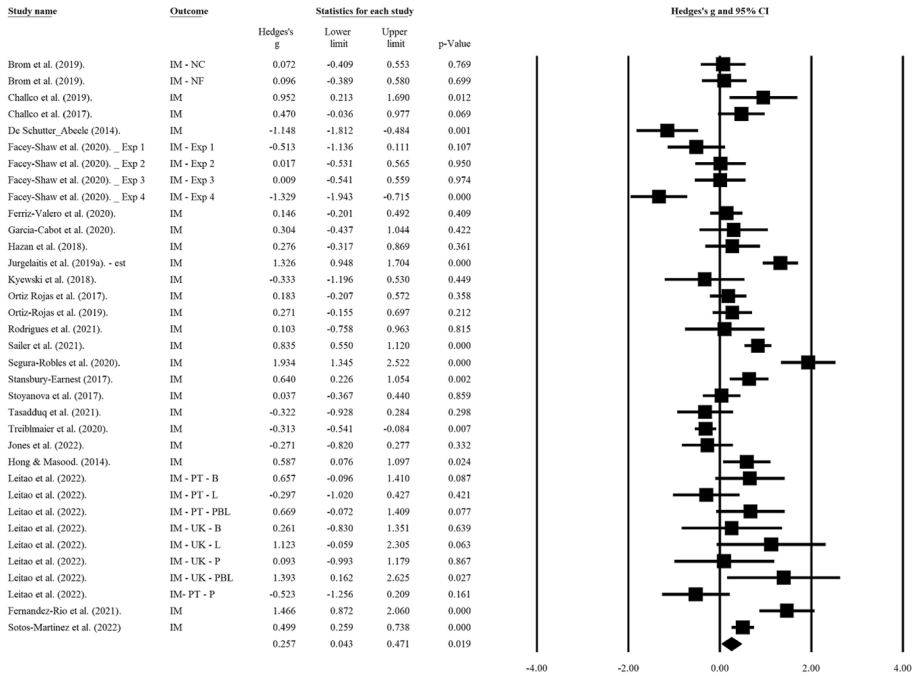
Fig. 2 Intrinsic motivation—Forest plot of effect sizes (Hedges' g) using the random effects model (n = 35)

statistically significant (Hedges' $g = 0.257$, 95% CI [0.043, 0.471], $p = 0.019$) (see Fig. 2). This result indicated that the gamified settings had a significant but small effect on the students' intrinsic motivation. A significant Q statistic ($Q = 206.403$, I = 83.527%, $p < 0.001$) indicated the presence of heterogeneity. We conducted several moderator analyses and assumed that the variables were unequal across subgroups to explore the possible reasons for this heterogeneity (see Tables 5, 6).

Table 5 Results of the Q-test for heterogeneity for the first four categories of moderators

Moderator	Q_B	Df	P-value
Participant characteristics			
School level	2.820	1	.093
Geography	5.490	3	.139
Curriculum characteristics			
Sample size	1.231	3	.746
Research design	3.119	3	.374
Intervention duration	9.509	4	.050*
Control level			
Student equivalence	0.463	2	.794
Instructor equivalence	12.596	2	.002*
Number of game elements			
Number of game elements	0.975	2	.614

Table 6 Results of the Q-test for heterogeneity for initial task interest and reward contingency

Moderator	n	p	g	95% CI		Q_B (p)
				LL	UL	
Type of Reward Contingency						99.486 (.000*)
Each unit solved	10	.041	.401	.016	.785	
Each unit solved + Surpassing a score	2	.631	.084	− .258	.425	
Completing a task + Each unit solved + Exceeding a norm	2	.768	− .085	− .647	.478	
Each unit solved + Surpassing a score + Exceeding a norm	2	.005	.633	.187	1.079	
Completing a task + Each unit solved + Surpassing a score + Exceeding a norm	2	.000	− 1.246	− 1.697	− .795	
Doing a task + Completing a task + Each unit solved + Surpassing a score	3	.575	.325	-.809	1.459	

n is the number of independent interventions, p is the p-value, g is Hedges' g, LL is the lower limit, UL is the upper limit, Q_B is between-group variance in effect sizes, and IM is Intrinsic Motivation

Table 7 Comparing effect sizes across values of instructor equivalence

Moderator	n	Hedges' g	SE	95% CI		Q$_B$(p)
				LL	UL	
Instructor equivalence						12.596 (.002*)
Different instructors	8	− .143	0.176	− 0.487	0.202	
Identical instructor	9	.613	0.124	0.371	0.855	
No data reported	18	.264	0.187	− 0.102	0.630	
Intervention duration						9.509 (.050*)
< 1 week	12	.371	.153	.071	.672	
1 week–1 month	5	.243	.208	− .165	.651	
1 month–3 months	7	.610	.253	.115	1.105	
> = 1 semester	8	− .068	.337	− .729	.592	
Not reported	3	− .127	.146	− .414	.159	

*$p \leq .05$

The heterogeneity analysis (Table 7) showed no significant variation in the effects of gamification across the samples attributable to participant characteristics, curriculum characteristics, and the number of game elements, i.e., (a) participants at different grade levels ($Q = 2.820$, $df = 1$, $p = 0.093$); (b) studies conducted in different countries ($Q = 5.490$, $df = 3$, $p = 0.139$); (c) studies with different sample sizes ($Q = 1.231$, $df = 3$, $p = 0.746$); (d) studies adopting different research designs ($Q = 3.119$, $df = 3$, $p = 0.374$); (e) student equivalence ($Q = 0.463$, $df = 2$, $p = 0.794$); and (f) studies using different numbers of game elements ($Q = 0.975$, $df = 2$, $p = 0.614$). However, there was significant variation across reward contingencies ($Q = 99.486$, $p = 0.000$). The effect sizes were greater when "each unit solved" or a combination of rewards (e.g., "each unit solved + surpassing a score + exceeding a norm") was used as the contingency for offering rewards.

Furthermore, significant differences were found across various intervention durations ($Q = 9.509$, $df = 4$, $p = 0.050$), and interventions conducted over a period of 1 to 3 months had the highest effect sizes. There were also significant differences between studies that had the same instructor for all groups and those that had different instructors ($Q = 12.596$, $df = 2$, $p = 0.002$). Effect sizes were larger in studies in which both the control and experimental groups were taught by the same instructor (Hedges' $g = 0.613$) than in those in which the groups were taught by different instructors (Hedges' $g = − 0.143$) or the study provided no data on whether the groups were taught by the same instructor or by different instructors (Hedges' $g = 0.264$) (see Table 6).

Publication bias

The following tests were performed to examine the possibility of publication bias in the studies examined in this review: Begg and Mazumdar's rank correlation analysis, the classic fail-safe N test, Egger's regression, and a funnel plot. The funnel plot is shown in Fig. 3. A visual inspection of the figure does not indicate any publication bias. Two statistical indicators supported this finding: Kendall's tau correlation coefficient with continuity correction ($\tau = − 0.02689$, two-tailed $p = 0.82025$) and Egger's regression intercept ($\alpha = − 0.14994$, two-tailed $p = 0.88721$).
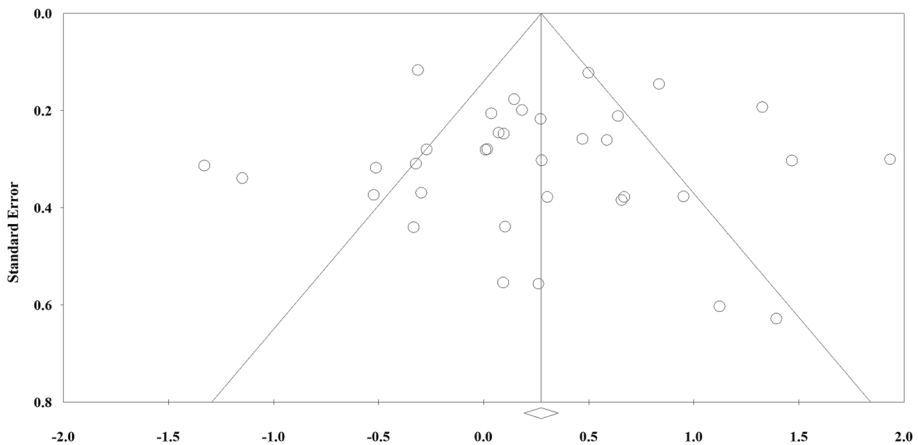
Fig. 3 Funnel plot of standard error by Hedges' g

Furthermore, we estimated the number of null studies (i.e., studies that were not published or did not report the effect of gamification on intrinsic motivation, because the findings were not significant) required to raise the p-value related with the mean effect to any alpha level ($a = 0.05$) by using the classic fail-safe N test. For the overall effect to be statistically nonsignificant, 281 missing studies with zero mean effect size were required. We thus concluded that the overall average effect size observed in our study was not exaggerated by publication bias, as such an exaggeration would have required a disproportionately large number of unreported studies with zero effects.

RQ3: What is the effect of gamification on the basic psychological needs that contribute to intrinsic motivation?

To better understand how gamification affects students' intrinsic motivation, we also calculated the effect sizes of the influence of gamification on the fulfillment of the three basic psychological needs (competence, autonomy, and relatedness) that contribute to intrinsic motivation.

Twelve of the 35 independent studies reported statistical data on competence. These data showed that the effect size of the influence of gamification on perceived competence was marginally significant (Hedges' $g = 0.277$, 95% CI [0.001, 0.553], $p = 0.049$) (see Fig. 4). We can therefore infer that these gamification interventions yielded minimal effect in enhancing students' perceived competence.

Eleven independent studies reported statistical data on autonomy, and the results showed that gamification contributed to the students' perceived autonomy (Hedges' $g = 0.638$, 95% CI [0.139, 1.136], $p = 0.012$) (see Fig. 5). Only four independent interventions reported statistical data on relatedness, and these data showed that gamification significantly facilitated the students' perceptions of relatedness (Hedges' $g = 1.776$, 95% CI [0.737, 2.814], $p = 0.001$) (see Fig. 6).
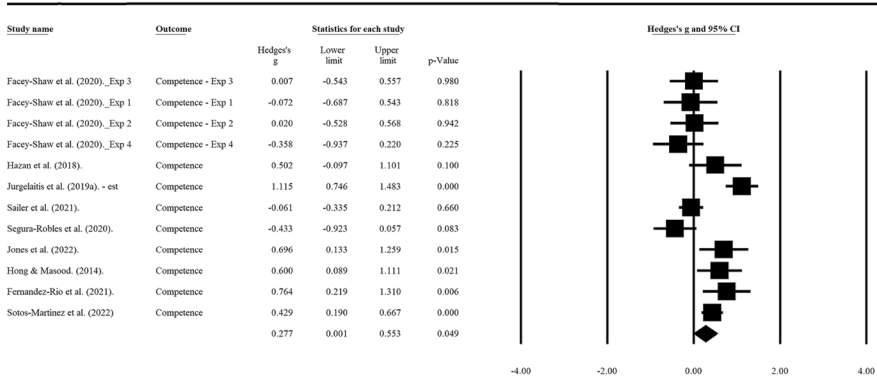
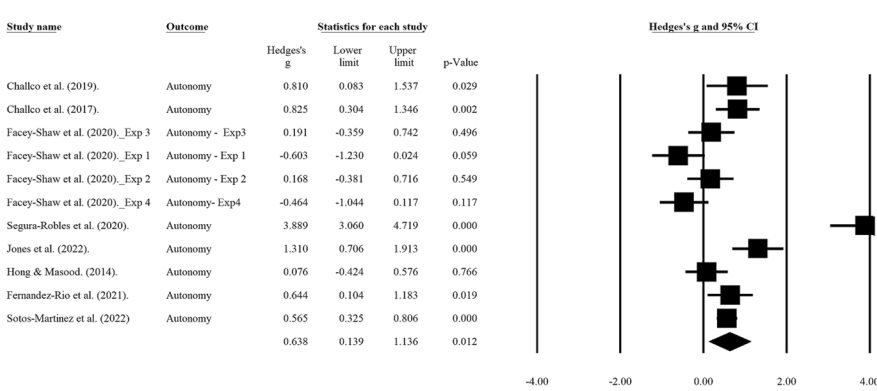**Fig. 4** Competence—Forest plot of effect sizes (Hedges' g) using the random effects model (n=12)

Fig. 5 Autonomy—Forest plot of effect sizes (Hedges' g) using the random effects model (n=11)
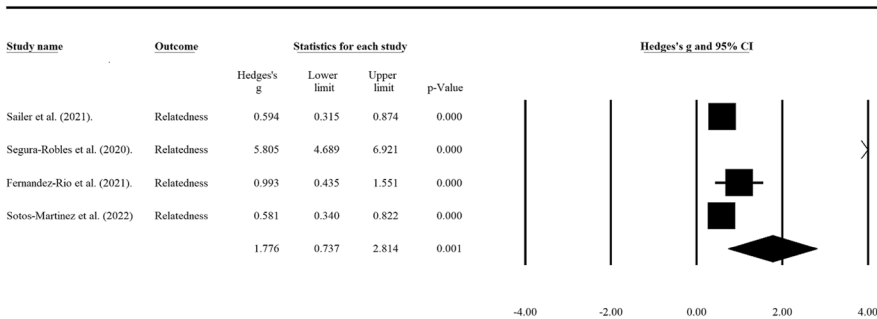

Fig. 6 Relatedness—Forest plot of effect sizes (Hedges' g) using the random effects model (n=4)

RQ4. What are the current challenges that gamification research must address?

Although the results of our meta-analysis suggested that gamification contributed to the learners' intrinsic motivation, its overall effect size was small. To further understand the challenges that gamification research must address, we conducted a systematic review of 31 gamification studies and identified two broad challenges currently faced when using gamification to facilitate students' intrinsic motivation: insufficient perceived competence and insufficient perceived autonomy of students in gamified classes. Details and examples of these challenges are summarized in Table 8.

In terms of perceived competence, the most frequently reported challenge was the discomfort experienced by learners (n = 7) who ranked low in the rankings displayed publicly on absolute leaderboards. Absolute leaderboards (also known as infinite leaderboards) display the positions of all players and are often used in educational settings (e.g., Bai et al., 2020; Tsay et al., 2018). On such a leaderboard, each participant can view the position of every other participant, and those at the top of the leaderboard may have a greater sense of achievement than those at the bottom (Ortiz-Rojas et al., 2019). This can contribute to

Table 8 Challenges in using gamification to support intrinsic motivation

Themes and sub-themes (Percentage)	Representative citations
Self-determination theory	
Lack of perceived competence	
Discomfort among the lowest ranked participants in a public absolute leaderboard (23%)	"The way in which the rewards were presented to students could have also played a role in their perception of the software. For example, whereas all students received points and a position on the leaderboard, the fact that only the top three students were awarded badges could have limited the effect of this gamification element. Furthermore, being high on the leaderboards or receiving a badge could have been particularly encouraging only for those students concerned, while the rest of the class could have potentially been negatively affected" (Andrade et al., 2020, p. 17)
Unsuitability of the difficulty level of the gamified tasks (10%)	"Students were not interested in certain badges, such as those given for merely attending classes…Some badges, such as attendance badges, were not considered valuable as students wanted badges that required achievement effort and not those that 'everyone can earn'" (Facey-Shaw et al., 2020, p. 42)
Unfamiliarity with gamification elements (10%)	"Some students also reported being unaware of the badges. This may have led to the statistically significant drop in scores for perceived competence" (Facey-Shaw et al., 2020, p. 46)
Lack of perceived autonomy	
Lack of autonomy in choosing learning content and activities (19%)	"Based on these findings and the way the students responded to the course in class, the following recommendations should be made: … Provide students with freedom of choice in how they want to show their mastery of the materials" (De Schutter & Abeele, 2014, p. 7)

social pressure and frustration for students who are ranked low (e.g., Andrade et al., 2020; Ferriz-Valero et al., 2020). Another challenge in terms of perceived competence, reported in three of the studies, was the unsuitable difficulty level of the gamified tasks. For example, some of the students perceived valueless and easy tasks as unchallenging (e.g., Facey-Shaw et al., 2020). Questions that exceeded the students' expected difficulty, combined with negative feedback, also undermined their competence needs (e.g., Sailer & Sailer, 2021). In addition, three studies reported concerns about the clarity and purpose of the gamification rules or elements. For example, Facey-Shaw et al. (2020) reported that some of the students were unaware of the existence of badges in the gamified classes, leading to surprise when they received them and confusion over how they were earned.

Six studies reported challenges in addressing the learners' needs for autonomy. The most frequently reported challenge was the learners' perceived lack of autonomy in choosing what to learn and which activities to engage in (e.g., De Schutter & Abeele, 2014; Hong & Masood, 2014; Ortiz-Rojas et al., 2019).

Discussion

Gamification has increasingly attracted the attention of educational researchers due to its potential to motivate students in their learning, and has been shown to positively influence students' behavior and learning outcomes (e.g., Bai et al., 2020; Huang & Hew, 2021; Ritzhaupt et al., 2021). However, despite its popularity, there is little consensus on whether gamification enhances students' intrinsic motivation. This review provides an overview of quantitative research on the use of gamification in educational settings to influence students' intrinsic motivation. Specifically, we conducted a meta-analysis of studies on gamification and its effects on intrinsic motivation, measured primarily through self-reported data obtained from students.

Based on a comprehensive and careful selection process, we identified 31 relevant articles among which 24 reported sufficient data for the meta-analysis. In total, 35 separate interventions (reported in these 24 articles) were examined in the meta-analysis. In addition to the meta-analysis, we conducted a systematic review and found that there were two main challenges that gamification research must address in relation to intrinsic motivation, namely, students' lack of perceived competence and lack of perceived autonomy in gamified classes.

Measurement of intrinsic motivation

We found that most of the gamification studies examined in the meta-analysis used self-reports (questionnaires or interviews) to assess the students' intrinsic motivation. One reason for the widespread use of self-reports is that this method enables researchers to measure large and diverse samples at a relatively low cost (Fredricks & McColskey, 2012). In addition, self-reports are relatively easy to administer in the classroom (Fredricks & McColskey, 2012). However, as mentioned earlier, there are challenges to the objectivity of the responses obtained via self-reports because participants are prone to self-favoring bias, exaggeration, and falsification when reporting about themselves (Paulhus & Vazire, 2007).

To address this drawback of self-reports, future studies should use behavioral assessment methods that are more objective such as free-choice behavior (Deci et al., 1999; Mekler et al., 2017). Although previous reviews of studies using self-reports to assess

intrinsic motivation have found similar results to those of studies using free-choice behavior (Deci et al., 1999), the adoption of behavioral free-choice measures of intrinsic motivation, such as by allowing participants the choice to continue engaging in a task without any reward (Deci & Ryan, 2004) and recording the outcome, may yield additional insights.

Effect of gamification on intrinsic motivation

The meta-analysis showed that gamified learning was more effective in increasing students' intrinsic motivation than non-gamified learning, although the effect size was small. One possible explanation for this greater effectiveness of gamified learning in enhancing students' intrinsic motivation is that gamified learning, using various game elements, meets students' basic psychological needs. However, the gamification features adopted in most gamification studies did not cater to all basic psychological needs (Xi & Hamari, 2019). This may be the reason for the small effect size found in our meta-analysis. Since the relationships between the three basic psychological needs are complementary (Ryan & Deci, 2000b), the fulfillment of all three needs has an additive impact on intrinsic motivation (Deci & Ryan, 2004; Rigby & Ryan, 2011). Conversely, if individuals perceive that the fulfillment of one of their basic psychological needs is diminished or hindered, a loss of motivation is likely to occur (Ryan & Deci, 2000b, 2020; van Roy & Zaman, 2019). This is an issue that future research on gamification must consider.

We also examined whether different moderators attenuated or exacerbated the differences in effect sizes between the interventions. We found evidence of differences in effect sizes that were attributable to differences in instructor equivalence between the interventions. Our comparison of the empirical studies based on the instructor equivalence adopted suggested that the average effect sizes of the influence of gamification was higher when both the gamified and non-gamified classrooms were taught by the same instructor than when different instructors taught them. However, considering that many of the studies did not report whether they ensured instructor equivalence, this finding should be interpreted with caution, and the causes of this variation cannot be conclusively established.

Furthermore, the results showed that gamification interventions that were short had a larger mean effect size than those that were long. The interventions that were 1 to 3 months long had the largest mean effect size ($g = 0.610$), whereas those that were more than 1 semester long had an almost negligible and negative mean effect size. One possible reason for this finding is that gamification is a non-traditional teaching method and the interest shown by the participants in the short-term interventions may be due to the fact that the approach is a new and exciting game-like learning approach. Over time, however, this novelty wears off and students become less engaged or even negative (e.g., bored) (Bai et al., 2020). Future research should pay more attention to psychological mechanisms in gamification design to meet the psychological needs of students and thus promote meaningful and lasting outcomes. Our results also indicated that the offering of rewards (either a single reward or a combination of rewards) based on the level of performance achieved was associated with a greater average effect size than the offering of combined rewards based simply on the completion of or participation in an activity. Specifically, reward combinations that consisted of rewards offered for solving each unit or problem, exceeding an absolute performance standard, and exceeding a relative performance standard had the largest average effect size ($g = 0.633$). A possible reason for this finding is that the effect of rewards on individuals' intrinsic motivation depends on how the reward events affect their perceptions

of competence, and when performance criteria are graded and attainable, rewards have a positive effect (Cameron et al., 2001).

Effect of gamification on basic psychological needs

The small overall effect size found in the meta-analysis necessitates reflection on whether gamification interventions, in their current implementations, fulfill basic psychological needs (competence, autonomy, and relatedness). Although some studies have expressed optimism regarding the use of gamification to meet basic psychological needs from a theoretical perspective, there is a lack of empirical evidence on whether gamification actually meets these needs. Therefore, one of the objectives of our meta-analysis was to determine whether prior gamification interventions fulfilled these basic psychological needs.

The results showed that the gamification interventions had a positive and significant effect on the students' perceptions of autonomy and relatedness. According to SDT, when a person can freely pursue an outcome or engage in an activity, they perceive a high sense of autonomy, which in turn promotes intrinsic motivation (Peng et al., 2012; Xi & Hamari, 2019). That is, when students perceive that they have freedom of choice in their actions, they perceive higher levels of autonomy, which in turn enhances their intrinsic motivation (e.g., Jones et al., 2022); conversely, when students feel forced to participate, their intrinsic motivation decreases (e.g., Hanus & Fox, 2015). Only a few of the studies examined in the meta-analysis reported statistical data on the need for relatedness. Nevertheless, the results based on those studies suggested that gamified learning was more conducive to enhancing the students' perceived relatedness than non-gamified learning. One explanation for this finding is that the high frequency of communication and idea-sharing between the students during group work in gamified classes may have contributed to fulfilling their need for relatedness (Fernandez-Rio et al., 2021). In addition, gamification may have stimulated competition between teams, thereby increasing the participants' sense of belonging to their team by strengthening their sense of community (van Roy & Zaman, 2019).

While the effect of gamification on the students' perceived competence, which refers to a sense of achievement and the recognition of one's competence by others, was statistically significant yet minimal. In the gamification studies examined in the meta-analysis, competence needs were the most frequently tested motivational factor (compared with autonomy and relatedness). Although it might seem that competence needs can be easily met in a gamified environment through performance indicators (e.g., leaderboards) or symbolic achievement icons (e.g., badges), the results showed that the gamification interventions did not satisfy the students' competence needs significantly better than non-gamified learning. One explanation for this result is that some of the game elements used may have had a negative effect on the students' perceived competence. Leaderboards, for example, publicly display students' ranks calculated based on some success criterion (Costa et al., 2013; Sailer et al., 2017). They are therefore a game element that directly communicates students' success relative to the entire class and elicits social comparison (Sailer et al., 2017). However, for students performing poorly, leaderboards convey negative feedback (e.g., an unpleasant sense of competition) and generate social pressure (Ferriz-Valero et al., 2020; Ortiz-Rojas et al., 2019), leading to a sense of incompetence and thus lower intrinsic motivation. Furthermore, although rewards such as badges and points create a sense of competence, they may undermine students' autonomy if they are perceived as controlling (Ryan & Deci, 2020), resulting not in a sense of accomplishment but rather in a loss of intrinsic motivation.

Challenges of using gamification to facilitate intrinsic motivation and potential solutions

Regarding the fourth research question, the results revealed two broad challenges to current implementations of gamification aimed at promoting intrinsic motivation: the lack of perceived competence and the lack of perceived autonomy among students in gamified classes. Three main factors were reported to contribute to the lack of perceived competence, namely, unsuitable difficulty levels of gamified tasks, unfamiliarity with gamification elements, and the discomfort caused to underperforming students by public absolute leaderboards.

To deal with the unsuitability of the difficulty levels of gamified tasks, gamification designers may consider providing tasks of varying difficulty for learners to choose from. As Deci and Ryan (1985) argued, activities that are trivial or simple and therefore provide no challenge are not intrinsically interesting, even for somebody who perceives themselves to be extremely competent. Facey-Shaw et al. (2020) also reported that students are not interested in trivial rewards that do not require any effort to achieve. In contrast, Jones et al. (2022) successfully satisfied students' perceived need for competence by providing a variety of assignments with different requirements for students to choose from.

To help reinforce students' understanding of gamification rules or elements, designers may consider combining storylines with tasks (Zarraonandia et al., 2015) and providing students with clear goals, thus promoting transparency regarding whether and how they might succeed in their attempts (Sailer et al., 2014; Xi & Hamari, 2019). To avoid the negative effects caused by public absolute leaderboards, other types of leaderboards should be considered. Relative leaderboards, which help to reduce lower ranked students' frustration and discouragement (Ortiz-Rojas et al., 2019), may reduce the likelihood that such students will lose their intrinsic motivation. To address the issue of students' perceived lack of autonomy in gamified classes, they should be provided with more choices for learning and opportunities to express themselves. For example, Fernandez-Rio et al. (2021) found that allowing students to choose their preferred path to explore a subject enhanced their perceived autonomy, as it gave them a sense of being in charge of their own actions. The use of avatars may also help enhance students' perceived autonomy (Sailer et al., 2017; Xi & Hamari, 2019) by, for example, allowing them to choose their preferred avatar.

Limitations

This study has certain limitations that must be acknowledged when interpreting its results. First, although the search string used in the study was broad in scope, enabling us to capture as many empirical studies as possible, important information was missing in many of the studies (e.g., insufficient statistical data). Missing information and unclear reporting of findings are common problems encountered in meta-analyses and research synthesis studies (Karabulut-Ilgu et al., 2018), and our study also had to contend with these challenges.

Second, the fact that self-reports were the most commonly used measurement approach in the studies may have affected the objectivity of the data gathered in this meta-analysis to examine the effect of gamification on intrinsic motivation (Paulhus & Vazire, 2007). Future studies could adopt more objective approaches to collecting data such as the observation of actual behavior. For example, free-choice behavior can be recorded in a manner that is "typically unobtrusive" (Deci et al., 1999, p. 656), whereby participants assume that the

experimenter is not aware of whether they persist in performing an activity during the free-choice period and thus decide whether to persist based on their own motivation.

Third, the coding and analysis of the moderating variables were based on information that was explicitly reported in the original articles. Thus, when information regarding a variable was not reported explicitly in the original article, we coded that information as not available. This may have led to minor differences between our recording of the variables reported in a study and the variables actually observed in the study.

Conclusion

Although gamification has garnered substantial interest in education research over the past decade, evidence of its ability to enhance students' intrinsic motivation remains unclear. To bridge this gap, we conducted a meta-analysis of 35 independent interventions in which we estimated the overall mean effect of the influence of the gamification interventions on students' intrinsic motivation and the fulfillment of the three basic psychological needs of autonomy, relatedness, and competence. We also identified the challenges faced when using gamification to enhance intrinsic motivation. The findings suggested that it is possible to foster students' intrinsic motivation by using gamified learning. This review contributes to the literature in three ways. First, it clarifies the effectiveness of existing gamification interventions in fostering intrinsic motivation. Second, the review enables educators to better understand whether gamification supports the basic psychological needs of students from a statistical perspective. Finally, the review identifies several challenges associated with the adoption of gamification to foster intrinsic motivation and offers possible solutions to these challenges.

We conclude by highlighting three directions for future gamification research. First, instead of self-report questionnaires, other approaches to measuring students' intrinsic motivation should be considered. For example, free-choice behavior could be recorded to examine whether and for how long students are intrinsically motivated to engage in learning activities in a free-choice period (Cameron & Pierce, 2002; Ryan & Deci, 1987).

Second, future research should examine more closely the effects of various types of rewards and reward contingencies used in gamification interventions on students' intrinsic motivation. In this meta-analysis, we examined reward types and reward contingencies as moderating variables. The results indicated that rewards tied to performance had a positive influence on intrinsic motivation, and this finding echoes that of Cameron and Pierce (2002). However, the number of papers that reported this information was small and we therefore call for more gamification research on the effects of various reward types and reward conditions on intrinsic motivation.

Third, whereas the results indicated that gamification enhanced intrinsic motivation, a substantial part of the heterogeneity in the effects of gamification across the studies could not be explained by the moderators investigated in the analysis. That is, the question of what factors facilitate gamification's influence on intrinsic motivation remains to be addressed. One reason why a substantial part of the heterogeneity could not be explained may be that our coding and analyses were based on what was explicitly reported in the articles reviewed. Furthermore, there could be discrepancies between what was documented in the articles and the actual research. Future research should explore how gamification should be designed to better foster intrinsic motivation rather than simply attempt to determine whether gamification is effective. To address the challenge of heterogeneity in

findings across studies, first, the design of instructional gamification interventions should be based on a comprehensive theoretical framework, and studies should provide a clear description of the instructional arrangements and the types of instructional activities used. Second, all aspects of a study's design, such as study characteristics and control group arrangements, should be reported transparently to facilitate a comprehensive meta-analytic investigation of the factors influencing the effectiveness of gamification.

Data availability Not applicable.

References

Altman, D. G., & Bland, J. M. (2005). Standard deviations and standard errors. *Bmj, 331*(7521), 903.

Andrade, P., Law, E. L.-C., Farah, J. C., & Gillet, D. (2020). Evaluating the effects of introducing three gamification elements in STEM educational software for secondary schools. 32nd Australian Conference on Human-Computer Interaction

Annetta, L. A. (2010). The "I's" have it: A framework for serious educational game design. *Review of General Psychology, 14*(2), 105–113.

Armstrong, M. B., & Landers, R. N. (2017). An evaluation of gamified training: Using narrative to improve reactions and learning. *Simulation & Gaming, 48*(4), 513–538.

Bai, S., Hew, K. F., & Huang, B. (2020). Does gamification improve student learning outcome? Evidence from a meta-analysis and synthesis of qualitative data in educational contexts. *Educational Research Review, 30*, 100322.

Baker, M. (2016). Statisticians issue warning on P values. *Nature, 531*(7593), 151–151.

Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin, 117*(3), 497.

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2021). *Introduction to meta-analysis*. John Wiley & Sons.

Boyle, E. A., Hainey, T., Connolly, T. M., Gray, G., Earp, J., Ott, M., Lim, T., Ninaus, M., Ribeiro, C., & Pereira, J. (2016). An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games. *Computers & Education, 94*, 178–192.

Brom, C., Stárková, T., Bromová, E., & Děchtěrenko, F. (2019). Gamifying a simulation: Do a game goal, choice, points, and praise enhance learning? *Journal of Educational Computing Research, 57*(6), 1575–1613.

Cameron, J., Banko, K. M., & Pierce, W. D. (2001). Pervasive negative effects of rewards on intrinsic motivation: The myth continues. *The Behavior Analyst, 24*(1), 1–44.

Cameron, J., & Pierce, W. D. (1994). Reinforcement, reward, and intrinsic motivation: A meta-analysis. *Review of Educational Research, 64*(3), 363–423.

Cameron, J., & Pierce, W. D. (2002). *Rewards and intrinsic motivation: Resolving the controversy*. Bergin & Garvey.

Cerasoli, C. P., Nicklin, J. M., & Ford, M. T. (2014). Intrinsic motivation and extrinsic incentives jointly predict performance: A 40-year meta-analysis. *Psychological Bulletin, 140*(4), 980.

Challco, G. C., Isotani, S., & Bittencourt, I. I. (2019). The effects of ontology-based gamification in scripted collaborative learning. 2019 IEEE 19th International Conference on Advanced Learning Technologies (ICALT),

Chen, J., Wang, M., Kirschner, P. A., & Tsai, C.-C. (2018). The role of collaboration, computer use, learning environments, and supporting strategies in CSCL: A meta-analysis. *Review of Educational Research, 88*(6), 799–843. https://doi.org/10.3102/0034654318791584

Clanton Harpine, E. (2015). Is Intrinsic Motivation Better Than Extrinsic Motivation? In E. Clanton Harpine (Ed.), *Group-centered prevention in mental health: Theory, training, and practice* (pp. 87–107). Springer International Publishing. https://doi.org/10.1007/978-3-319-19102-7_6

Costa, J. P., Wehbe, R. R., Robb, J., & Nacke, L. E. (2013). Time's up: studying leaderboards for engaging punctual behaviour. Proceedings of the First International Conference on Gameful Design, Research, and Applications,

Deci, E. L., & Ryan, R. M. (2022). *Intrinsic Motivation Inventory (IMI)*. https://selfdeterminationtheory.org/intrinsic-motivation-inventory/#toc-description

Deci, E. L., Koestner, R., & Ryan, R. M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin, 125*(6), 627.

Deci, E. L., Koestner, R., & Ryan, R. M. (2001). Extrinsic rewards and intrinsic motivation in education: Reconsidered once again. *Review of Educational Research, 71*(1), 1–27.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. Springer Science & Business Media.

Deci, E. L., & Ryan, R. M. (2000). The" what" and" why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry, 11*(4), 227–268.

Deci, E. L., & Ryan, R. M. (2004). *Handbook of self-determination research*. University Rochester Press.

Deci, E. L., & Ryan, R. M. (2008). Self-determination theory: A macrotheory of human motivation, development, and health. *Canadian Psychology/psychologie Canadienne, 49*(3), 182.

Deci, E. L., & Ryan, R. M. (2012). Motivation, personality, and development within embedded social contexts: An overview of self-determination theory. In R. M. E. Ryan (Ed.), *The Oxford Handbook of Human Motivation* (pp. 85–107). Oxford University Press.

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: defining "gamification". Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments,

Deterding, S. (2012). Gamification: designing for motivation. *Interactions, 19*(4), 14–17.

Dicheva, D., Dichev, C., Agre, G., & Angelova, G. (2015). Gamification in education: A systematic mapping study. *Journal of Educational Technology & Society, 18*(3), 75–88.

Dörnyei, Z., & Ushioda, E. (2013). *Teaching and researching: Motivation*. Routledge.

Educause. (2011). 7 things you should know about gamification. https://library.educause.edu/resources/2011/8/7-things-you-should-know-about-gamification

Facey-Shaw, L., Specht, M., van Rosmalen, P., & Bartley-Bryan, J. (2020). Do Badges Affect Intrinsic Motivation in Introductory Programming Students? *Simulation & Gaming, 51*(1), 33–54. https://doi.org/10.1177/1046878119884996

Fernandez-Rio, J., Zumajo-Flores, M., & Flores-Aguilar, G. (2021). Motivation, basic psychological needs and intention to be physically active after a gamified intervention programme. *European Physical Education Review*. https://doi.org/10.1177/1356336X21105288

Ferriz-Valero, A., Østerlie, O., García Martínez, S., & García-Jaén, M. (2020). Gamification in physical education: Evaluation of impact on motivation and academic performance within higher education. *International Journal of Environmental Research and Public Health, 17*(12), 4465.

Fransen, K., Boen, F., Vansteenkiste, M., Mertens, N., & Vande Broek, G. (2018). The power of competence support: The impact of coaches and athlete leaders on intrinsic motivation and performance. *Scandinavian Journal of Medicine & Science in Sports, 28*(2), 725–745.

Fredricks, J. A., & McColskey, W. (2012). The measurement of student engagement: A comparative analysis of various methods and student self-report instruments. *Handbook of research on student engagement* (pp. 763–782). Springer.

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences, 111*(23), 8410–8415.

Fu, R., Gartlehner, G., Grant, M., Shamliyan, T., Sedrakyan, A., Wilt, T. J., Griffith, L., Oremus, M., Raina, P., & Ismaila, A. (2011). Conducting quantitative synthesis when comparing medical interventions: AHRQ and the Effective Health Care Program. *Journal of Clinical Epidemiology, 64*(11), 1187–1197.

Garcia-Cabot, A., Garcia-Lopez, E., Caro-Alvaro, S., Gutierrez-Martinez, J. M., & de Marcos, L. (2020). Measuring the effects on learning performance and engagement with a gamified social platform in an MSc program. *Computer Applications in Engineering Education, 28*(1), 207–223.

AECT

Gladun, M. (2016). Using gamification as a way of increasing students motivation. *Відкрите Освітнє е-Середовище Сучасного Університету, 2016*(2), 14–19.

Gnambs, T., & Hanfstingl, B. (2016). The decline of academic motivation during adolescence: An accelerated longitudinal cohort analysis on the effect of psychological need satisfaction. *Educational Psychology, 36*(9), 1691–1705.

Greene, R. J. (2018). *Rewarding performance: Guiding principles; custom strategies*. Routledge.

Gurevitch, J., & Hedges, L. V. (1999). Statistical issues in ecological meta-analyses. *Ecology, 80*(4), 1142–1149.

Hanus, M. D., & Fox, J. (2015). Assessing the effects of gamification in the classroom: A longitudinal study on intrinsic motivation, social comparison, satisfaction, effort, and academic performance. *Computers & Education, 80*, 152–161.

Hazan, B., Zhang, W., Olcum, E., Bergdoll, R., Grandoit, E., Mandelbaum, F., Wilson-Doenges, G., & Rabin, L. A. (2018). Gamification of an undergraduate psychology statistics lab: Benefits to perceived competence. *Statistics Education Research Journal, 17*(2), 255–265.

Hewett, R., & Conway, N. (2016). The undermining effect revisited: The salience of everyday verbal rewards and self-determined motivation. *Journal of Organizational Behavior, 37*(3), 436–455.

Higgins, J. P., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (2019). *Cochrane handbook for systematic reviews of interventions*. John Wiley & Sons.

Hong, G. Y., & Masood, M. (2014). Effects of gamification on lower secondary school students' motivation and engagement. *International Journal of Educational and Pedagogical Sciences, 8*(12), 3765–3772.

Huang, B., & Hew, K. F. (2021). Using gamification to design courses. *Educational Technology & Society, 24*(1), 44–63.

Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting error and bias in research findings*. Sage.

Jones, M., Blanton, J. E., & Williams, R. E. (2022). Science to practice: Does gamification enhance intrinsic motivation? *Active Learning in Higher Education*. https://doi.org/10.1177/14697874211066882

Jurgelaitis, M., Čeponienė, L., Čeponis, J., & Drungilas, V. (2019). Implementing gamification in a university-level UML modeling course: A case study. *Computer Applications in Engineering Education, 27*(2), 332–343.

Karabulut-Ilgu, A., Jaramillo Cherrez, N., & Jahren, C. T. (2018). A systematic review of research on the flipped learning method in engineering education. *British Journal of Educational Technology, 49*(3), 398–411.

Karimi, S., & Sotoodeh, B. (2020). The mediating role of intrinsic motivation in the relationship between basic psychological needs satisfaction and academic engagement in agriculture students. *Teaching in Higher Education, 25*(8), 959–975. https://doi.org/10.1080/13562517.2019.1623775

Koivisto, J., & Hamari, J. (2019). The rise of motivational information systems: A review of gamification research. *International Journal of Information Management, 45*, 191–210. https://doi.org/10.1016/j.ijinfomgt.2018.10.013

Korpershoek, H., Harms, T., de Boer, H., van Kuijk, M., & Doolaard, S. (2016). A meta-analysis of the effects of classroom management strategies and classroom management programs on students' academic, behavioral, emotional, and motivational outcomes. *Review of Educational Research, 86*(3), 643–680.

Kyewski, E., & Krämer, N. C. (2018). To gamify or not to gamify? An experimental field study of the influence of badges on motivation, activity, and performance in an online learning course. *Computers & Education, 118*, 25–37.

La Guardia, J., & Ryan, R. (2002). What adolescents need. *Academic motivation of adolescents* (Vol. 2, pp. 193–219). IAP Information Age Publishing.

Landers, R. N. (2014). Developing a theory of gamified learning: Linking serious games and gamification of learning. *Simulation & Gaming, 45*(6), 752–768.

Landers, R. N., Bauer, K. N., Callan, R. C., & Armstrong, M. B. (2014). *Psychological theory and the gamification of learning* (pp. 165–186). Springer International Publishing. https://doi.org/10.1007/978-3-319-10208-5_9

Lee, J. Q., McInerney, D. M., Liem, G. A. D., & Ortiga, Y. P. (2010). The relationship between future goals and achievement goal orientations: An intrinsic–extrinsic motivation perspective. *Contemporary educational psychology, 35*(4), 264–279. https://doi.org/10.1016/j.cedpsych.2010.04.004

Leitão, R., Maguire, M., Turner, S., & Guimarães, L. (2022). A systematic evaluation of game elements effects on students' motivation. *Education and Information Technologies, 27*(1), 1081–1103.

Lepper, M. R., Corpus, J. H., & Iyengar, S. S. (2005). Intrinsic and extrinsic motivational orientations in the classroom: Age differences and academic correlates. *Journal of Educational Psychology, 97*(2), 184.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Sage.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. SAGE Publications Inc.

Mekler, E. D., Brühlmann, F., Tuch, A. N., & Opwis, K. (2017). Towards understanding the effects of individual gamification elements on intrinsic motivation and performance. *Computers in Human Behavior, 71*, 525–534.

Mitchell, R., Schuster, L., & Drennan, J. (2017). Understanding how gamification influences behaviour in social marketing. *Australasian Marketing Journal (AMJ), 25*(1), 12–19.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Group, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS medicine, 6*(7), e1000097.

Mula-Falcón, J., Moya-Roselló, I., & Ruiz-Ariza, A. (2022). The active methodology of gamification to improve motivation and academic performance in educational context: A meta-analysis. *Review of European Studies*. https://doi.org/10.5539/res.v14n2p32

Ortiz-Rojas, M., Chiluiza, K., & Valcke, M. (2019). Gamification through leaderboards: An empirical study in engineering education. *Computer Applications in Engineering Education, 27*(4), 777–788.

Paulhus, D. L., & Vazire, S. (2007). The self-report method. *Handbook of research methods in personality psychology* (pp. 224–239). The Guilford Press.

Peng, W., Lin, J.-H., Pfeiffer, K. A., & Winn, B. (2012). Need satisfaction supportive game features as motivational determinants: An experimental study of a self-determination theory guided exergame. *Media Psychology, 15*(2), 175–196.

Przybylski, A. K., Rigby, C. S., & Ryan, R. M. (2010). A motivational model of video game engagement. *Review of General Psychology, 14*(2), 154–166.

Raudenbush, S. W. (2009). Analyzing effect sizes: Random-effects models. *The handbook of research synthesis and meta-analysis* (Vol. 2, pp. 295–316). Russell Sage Foundation.

Raufelder, D., & Kulakow, S. (2021). The role of the learning environment in adolescents' motivational development. *Motivation and Emotion, 45*(3), 299–311. https://doi.org/10.1007/s11031-021-09879-1

Rigby, S., & Ryan, R. M. (2011). *Glued to games: How video games draw us in and hold us spellbound: How video games draw us in and hold us spellbound*. AbC-CLIo.

Ritzhaupt, A. D., Huang, R., Sommer, M., Zhu, J., Stephen, A., Valle, N., Hampton, J., & Li, J. (2021). A meta-analysis on the influence of gamification in formal educational settings on affective and behavioral outcomes. *Educational Technology Research and Development, 69*(5), 2493–2522.

Rodrigues, L., Toda, A. M., Oliveira, W., Palomino, P. T., Avila-Santos, A. P., & Isotani, S. (2021). Gamification Works, but How and to Whom? An Experimental Study in the Context of Programming Lessons. Proceedings of the 52nd ACM Technical Symposium on Computer Science Education,

Rosenthal, R., & DiMatteo, M. R. (2001). Meta-analysis: Recent developments in quantitative methods for literature reviews. *Annual Review of Psychology, 52*(1), 59–82.

Roy, R. V., & Zaman, B. (2017). Why gamification fails in education and how to make it successful: Introducing nine gamification heuristics based on self-determination theory. *Serious Games and edutainment applications* (pp. 485–509). Springer.

Ryan, R., & Deci, E. (1987). When free-choice behavior is not intrinsically motivated: Experiments on internally controlling regulation. *Unpublished manuscript, University of Rochester*.

Ryan, R. M., & Deci, E. L. (2000). Chapter 2 - When rewards compete with nature: The undermining of intrinsic motivation and Self-Regulation. In C. Sansone & J. M. Harackiewicz (Eds.), *Intrinsic and Extrinsic Motivation* (pp. 13–54). Academic Press. https://doi.org/10.1016/B978-012619070-0/50024-6

Ryan, R. M., & Deci, E. L. (2000b). The darker and brighter sides of human existence: Basic psychological needs as a unifying concept. *Psychological Inquiry, 11*(4), 319–338. https://doi.org/10.1207/S15327965PLI1104_03

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: classic definitions and new directions. *Contemporary Educational Psychology, 25*(1), 54–67. https://doi.org/10.1006/ceps.1999.1020

Ryan, R. M., & Deci, E. L. (2000d). 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., & Deci, E. L. (2002). Overview of self-determination theory: An organismic dialectical perspective. *Handbook of self-determination research* (Vol. 2, pp. 3–33). University of Rochester Press.

Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. Guilford Publications.

Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary educational psychology, 61*, 101860. https://doi.org/10.1016/j.cedpsych.2020.101860

Ryan, R. M., Koestner, R., & Deci, E. L. (1991). Ego-involved persistence: When free-choice behavior is not intrinsically motivated. *Motivation and Emotion, 15*(3), 185–205.

AECT

Ryan, R. M., & La Guardia, J. G. (2000). What is being optimized?: Self-determination theory and basic psychological needs. In S. H. Qualls & N. Abeles (Eds.), *Psychology and the aging revolution* (pp. 145–172). American Psychological Association. https://doi.org/10.1037/10363-008

Sailer, M., Hense, J., Mandl, J., & Klevers, M. (2014). Psychological perspectives on motivation through gamification. *Interaction Design and Architecture Journal, 19*, 28–37.

Sailer, M., Hense, J. U., Mayr, S. K., & Mandl, H. (2017). How gamification motivates: An experimental study of the effects of specific game design elements on psychological need satisfaction. *Computers in Human Behavior, 69*, 371–380. https://doi.org/10.1016/j.chb.2016.12.033

Sailer, M., & Homner, L. (2020). The gamification of learning: A meta-analysis. *Educational Psychology Review, 32*(1), 77–112.

Sailer, M., & Sailer, M. (2021). Gamification of in-class activities in flipped classroom lectures. *British Journal of Educational Technology, 52*(1), 75–90.

Scammacca, N., Roberts, G., & Stuebing, K. K. (2014). Meta-analysis with complex research designs: Dealing with dependence from multiple measures and multiple group comparisons. *Review of Educational Research, 84*(3), 328–364.

Scherrer, V., & Preckel, F. (2019). Development of motivational variables and self-esteem during the school career: A meta-analysis of longitudinal studies. *Review of Educational Research, 89*(2), 211–258.

De Schutter, B., & Abeele, V. V. (2014). Gradequest-Evaluating the impact of using game design techniques in an undergraduate course. In 9th international conference on the foundations of digital games., USA.

Segura-Robles, A., Fuentes-Cabrera, A., Parra-González, M. E., & López-Belmonte, J. (2020). Effects on personal factors through flipped learning and gamification as combined methodologies in secondary education. *Frontiers in Psychology, 11*, 1103.

Shamseer, L., Moher, D., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., & Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ, 349*, g7647–g7647. https://doi.org/10.1136/bmj.g7647

Stansbury, J. A., & Earnest, D. R. (2017). Meaningful gamification in an industrial/organizational psychology course. *Teaching of Psychology, 44*(1), 38–45.

Stoyanova, M., Tuparova, D., & Samardzhiev, K. (2017). Impact of motivation, gamification and learning style on students' interest in maths classes–a study in 11 high school grade. International Conference on Interactive Collaborative Learning,

Sun, J.C.-Y., & Hsieh, P.-H. (2018). Application of a gamified interactive response system to enhance the intrinsic and extrinsic motivation, student engagement, and attention of English learners. *Journal of Educational Technology & Society, 21*(3), 104–116.

Tasadduq, M., Khan, M. S., Nawab, R. M., Jamal, M. H., & Chaudhry, M. T. (2021). Exploring the effects of gamification on students with rote learning background while learning computer programming. *Computer Applications in Engineering Education, 29*(6), 1871–1891.

Tondello, G. F., Mora, A., & Nacke, L. E. (2017). Elements of gameful design emerging from user preferences. Proceedings of The Annual Symposium on Computer-Human Interaction in Play

Tsay, C.H.-H., Kofinas, A., & Luo, J. (2018). Enhancing student learning experience with technology-mediated gamification: An empirical study. *Computers & Education, 121*, 1–17.

Vallerand, R. J., Pelletier, L. G., Blais, M. R., Briere, N. M., Senecal, C., & Vallieres, E. F. (1992). The Academic Motivation Scale: A measure of intrinsic, extrinsic, and amotivation in education. *Educational and Psychological Measurement, 52*(4), 1003–1017.

Van den Broeck, A., Vansteenkiste, M., De Witte, H., Soenens, B., & Lens, W. (2010). Capturing autonomy, competence, and relatedness at work: Construction and initial validation of the work-related basic need satisfaction scale. *Journal of Occupational and Organizational Psychology, 83*(4), 981–1002.

van Roy, R., & Zaman, B. (2019). Unravelling the ambivalent motivational power of gamification: A basic psychological needs perspective. *International Journal of human-computer studies, 127*, 38–50. https://doi.org/10.1016/j.ijhcs.2018.04.009

Vansteenkiste, M., Lens, W., & Deci, E. L. (2006). Intrinsic versus extrinsic goal contents in self-determination theory: Another look at the quality of academic motivation. *Educational Psychologist, 41*(1), 19–31.

Vansteenkiste, M., Niemiec, C. P., & Soenens, B. (2010). The development of the five mini-theories of self-determination theory: An historical overview, emerging trends, and future directions. *The decade ahead: Theoretical perspectives on motivation and achievement.* Emerald Group Publishing Limited.

Vansteenkiste, M., & Ryan, R. M. (2013). On psychological growth and vulnerability: Basic psychological need satisfaction and need frustration as a unifying principle. *Journal of Psychotherapy Integration, 23*(3), 263.

Xi, N., & Hamari, J. (2019). Does gamification satisfy needs? A study on the relationship between gamification features and intrinsic need satisfaction. *International Journal of Information Management, 46*, 210–221. https://doi.org/10.1016/j.ijinfomgt.2018.12.002

Xiang, P., Agbuga, B., Liu, J., & McBride, R. E. (2017). Relatedness need satisfaction, intrinsic motivation, and engagement in secondary school physical education. *Journal of Teaching in Physical Education, 36*(3), 340–352.

Xu, J., Lio, A., Dhaliwal, H., Andrei, S., Balakrishnan, S., Nagani, U., & Samadder, S. (2021). Psychological interventions of virtual gamification within academic intrinsic motivation: A systematic review. *Journal of Affective Disorders, 293*, 444–465. https://doi.org/10.1016/j.jad.2021.06.070

Yaşar, H., Kiyici, M., & Karatas, A. (2020). The views and adoption levels of primary school teachers on gamification, problems and possible solutions. *Participatory Educational Research*, *7*(3), 265–279.

Zarraonandia, T., Diaz, P., Aedo, I., & Ruiz, M. R. (2015). Designing educational games through a conceptual model based on rules and scenarios. *Multimedia Tools and Applications, 74*(13), 4535–4559.

Zhang, Q., Yu, L., & Yu, Z. (2021). A content analysis and meta-analysis on the effects of classcraft on gamification learning experiences in terms of learning achievement and motivation. *Education Research International*. https://doi.org/10.1155/2021/9429112

Zheng, B., Warschauer, M., Lin, C.-H., & Chang, C. (2016). Learning in one-to-one laptop environments: A meta-analysis and research synthesis. *Review of Educational Research, 86*(4), 1052–1084. https://doi.org/10.3102/0034654316628645

Zichermann, G., & Cunningham, C. (2011). *Gamification by design: Implementing game mechanics in web and mobile apps*. O'Reilly Media Inc.

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

Liuyufeng Li is a PhD candidate at the Faculty of Education, University of Hong Kong. Her research focuses on the use of gamification in education contexts.

Khe Foon Hew is Professor of Instructional Design and Technology at the Faculty of Education, University of Hong Kong. His research focuses on improving student engagement in blended and online learning environments.

Jiahui Du is a PhD candidate at the Faculty of Education, University of Hong Kong. Her research focuses on student self-regulated learning.