REVIEW ARTICLE

Open Access



Gamifying education: what is known, what is believed and what remains uncertain: a critical review

Christo Dichev and Darina Dicheva*

Abstract

Gamification of education is a developing approach for increasing learners' motivation and engagement by incorporating game design elements in educational environments. With the growing popularity of gamification and yet mixed success of its application in educational contexts, the current review is aiming to shed a more realistic light on the research in this field by focusing on empirical evidence rather than on potentialities, beliefs or preferences. Accordingly, it critically examines the advancement in gamifying education. The discussion is structured around the used gamification mechanisms, the gamified subjects, the type of gamified learning activities, and the study goals, with an emphasis on the reliability and validity of the reported outcomes. To improve our understanding and offer a more realistic picture of the progress of gamification in education, consistent with the presented evidence, we examine both the outcomes reported in the papers and how they have been obtained. While the gamification in education is still a growing phenomenon, the review reveals that (i) insufficient evidence exists to support the long-term benefits of gamification in educational contexts; (ii) the practice of gamifying learning has outpaced researchers' understanding of its mechanisms and methods; (iii) the knowledge of how to gamify an activity in accordance with the specifics of the educational context is still limited. The review highlights the need for systematically designed studies and rigorously tested approaches confirming the educational benefits of gamification, if gamified learning is to become a recognized instructional approach.

Keywords: Gamification in education, Gamifying learning, Critical literature review, Empirical studies

The idea of incentivizing people is not new but the term "gamification" didn't enter the mainstream vocabulary until 2010. Only a year later it became a viable trend. The growing popularity of gamification is stemming from the belief in its potential to foster motivation, behavioral changes, friendly competition and collaboration in different contexts, such as customer engagement, employee performance and social loyalty. As with any new and promising technology it has been applied in a diversity of domains, including marketing, healthcare, human resources, training, environmental protection and wellbeing. Gamification is a multidisciplinary concept spanning a range of theoretical and empirical knowledge, technological domains and platforms and is driven by an



^{*} Correspondence: dichevad@wssu.edu Winston-Salem State University, 601 S. Martin Luther King Jr. Drive, Winston Salem, NC 27110, USA

array of practical motivations (Seaborn & Fels, 2015). In an attempt to best capture the essence of the underlying concepts and practices, the term gamification has been defined in several ways, such as "the use of game design elements in nongame contexts" (Deterding, Dixon, Khaled, & Nacke, 2011), "the phenomenon of creating gameful experiences" (Hamari, Koivisto, & Sarsa, 2014), or "the process of making activities more game-like" (Werbach, 2014). Empirical work across disciplines has begun to explore how gamification can be used in certain contexts and what behavioral and experiential effects gamification has on people in the short and long terms.

Ever since its advent gamification has sparked controversy between game designers, user experience designers, game theorists and researchers in human-computer interaction (Mahnič, 2014). This controversy is reflected also in some scientific studies of gamification, which show that its effect on motivation or participation is lower than the expectations created by the hype (Broer, 2014). Even so, substantial efforts have sought to take advantage of the alleged motivational benefits of gamification approaches.

One key sector where gamification is being actively explored (mainly for its potential to motivate) is education. Motivation is among the important predictors of student academic achievements, which influences the effort and time a student spends engaged in learning (Linehan, Kirman, Lawson, & Chan, 2011). Given that games, known to engender motivation and engagement, are notably popular, the proposal to incorporate game mechanics and principles to motivate the learner is appealing. Gamification in education refers to the introduction of game design elements and gameful experiences in the design of learning processes. It has been adopted to support learning in a variety of contexts and subject areas and to address related attitudes, activities, and behaviors, such as participatory approaches, collaboration, self-guided study, completion of assignments, making assessments easier and more effective, integration of exploratory approaches to learning, and strengthening student creativity and retention (Caponetto et al. 2014). The rationality at the basis of gamifying learning is that adding elements, such as those found in games to learning activities will create immersion in a way similar to what happens in games (Codish & Ravid, 2015). This leads to the belief that by incorporating game mechanics in the design of a learning process, we can engage learners in a productive learning experience, and more generally, change their behavior in a desirable way (Holman et al. 2013). Yet, the design of successful gamification applications in education that can sustain the intended behavior changes is still more of a guessing practice than science. This fact is in line with the Gartner Hype Cycle (Gartner, 2013), a research methodology that outlines a technology's viability for commercial success, which points out that an emerging technology first climbs the 'peak of inflated expectations' followed by a subsequent strong fall down into the 'trough of disillusionment,' before reaching the 'slop of enlightenment', which marks the stage where its benefits and limitations are understood and demonstrated.

The Gartner model is intended for representing the level of maturity and adoption of certain emerging technologies. We maintain the view that gamification is not just a technology but also a methodology which some organizations adopt as way to increase motivation. In this aspect, gamification is not a purely marketing trend but a behavioral/affective design trend that can be applied to different areas, including education. As such, gamification is also a growing area of research. However, research efforts and

trends should be driven and evaluated based on distinct factors. Thus Gartner's model is used here metaphorically and as a comparison model. We borrow it to illustrate observed trends in emerging research areas, demonstrating some sorts of 'peaks of inflated expectations' and 'enlightenments'.

In 2014 we conducted a systematic mapping study of the empirical research published between January 2010 and June 2014 intended to recognize the emerging trends within the area of applications of gamification to education and to identify patterns, educational contexts and configurations of used game elements (Dicheva et al. 2015). For classifying the research results, the study used a categorical structure (based on the topics discussed in the reviewed papers) including game elements, context of the application of gamification, gamification implementation and evaluation. Although most of the reviewed 34 papers have been reporting promising results, the review concluded that more substantial empirical research is needed to determine whether both extrinsic and intrinsic motivation of the learners can be actually influenced by gamification. Given the exponential growth of publications on gamification, a year later we conducted a follow-up study covering the period July 2014-December 2015. Our goal was twofold: from one side, to complement the previous study and compare it with the findings derived from the papers published within the last year, and from another, to identify any shifts and new trends in this evolving field. The results from that review were published in (Dicheva and Dichev 2015).

In terms of the Gartner's hype cycle, our first review (Dicheva et al. 2015) covered works from the rise-in-expectations period of gamification, where the reported outcomes of the early empirical work were often influenced by the hype prompting desire to demonstrate that gamification is an effective tool for motivating and engaging learners in educational contexts. We believe that the progress in the research, including educational research, unlike technological evolutions should differ from the Gartner's hype cycle and evolve independently of media attention using instead scientific indicators for recognizing promising trends and thus minimizing inflated expectations. More importantly, the research efforts should be directed at understanding the phenomenon triggering the new interest and at generating evidence for or against the trend causing that interest. This suggests that the research should progress following a pattern different from the Gartner's hype cycle and marked by stages, such as early studies, emerging research area, research topics formation, etc. In this sense, our second review was intended to take another snapshot in an attempt to verify this view. Despite the growing body of studies, we found the level of understanding of how to promote engagement and learning by incorporating game design elements to be questionable. In parallel, a significant part of the empirical research was nonetheless reporting success stories and possibly contributing to the 'inflated expectations'. Because the empirical studies (on gamification) explore the unknown, uncertainty is an unavoidable part of the investigations. While the publication of valid and reliable studies reduces the uncertainty and adds to the knowledge on gamifying education, thus helping to shape future research in the field, invalid or unreliable findings obscure our understanding of the studied phenomenon. In this context and unlike the systematic mapping studies, the goal of this critical review is to see how the new studies are shaping the evolving research in educational gamification. In particular, compared to the previous reviews the focus here is shifted to analyzing and critically appraising the collected evidence

from the latest empirical research with the aim of distinguishing facts from hypotheses or opinions. From this perspective, the present review adds to the first two by trying to subject educational gamification research to similar standards as in social or health sciences.

Accordingly, in this article the focus is on analyzing the understanding of the motivational mechanisms provided by gamification in educational settings and its impact on learning. The guiding questions in this context were:

- What empirical evidence exists for the impact of gamification on motivational processes and effectiveness of learning?
- What is the level of progress towards a systematic understanding of how to use gamification in educational contexts?

With the growing popularity of gamification and yet mixed opinions about its successful application in educational contexts, the current review is aiming to shed a more realistic light on the research in this field focusing on empirical evidence rather than on potentialities, beliefs and preferences.

On the technical side, the article includes several tables that summarize and add to the information provided in the text. The article also includes two appendices that summarize the relevant features of the reviewed studies.

The study

Search strategy and sources

In search for empirical research papers, that is, papers based on actual observations or experiments on educational gamification, we searched the following databases: Google Scholar, ACM Digital Library, IEEE Explore and ScienceDirect using the following search terms: (gamification OR gamify OR gameful) AND (education OR learning OR training) AND (since 2014). In the cases when the OR option was not available in the provided Boolean search functionality, an equivalent search strategy was carried out through multiple searches with alternative terms. This search yielded a total of 4998 results depicted in Table 1. We have chosen the definition of (Deterding et al., 2011) for gamification ("the use of game design elements in non-gaming contexts") to measure each found publication for relevance. Accordingly, publications discussing full-fledged games were filtered out. Peer-reviewed empirical research papers where no findings were reported were also excluded. For example, purely descriptive papers such as (Morrison & DiSalvo, 2014), which describes the implementation of gamification within Khan

Table 1 Distribution of retrieved papers among sources

	1 1 9	
Database	Total Amount	No of Selected
ACM Digital Library	285	12 + 1 ^a
Google Scholar	4021	$21 + 8^{a}$
IEEE Explore	574	6
ScienceDirect	158	12 + 3 ^a

Note: ^aTheoretical Papers

Academy, were not included. At the end of this step, all papers that appeared in the review presented in (Dicheva et al., 2015) were also filtered out. The review was restricted to papers appearing in the searched databases between June 30, 2014 and December 31, 2015. The result was a list of 51 empirical research papers. In sum, in the past one and a half years, several hundred articles pertaining to gamification in education have been published however only 51 studies met our criteria and are reviewed in this article.

For completeness of the review of the research in the field, we decided this time to include also theoretical papers dealing with gamification in education. Following (Seaborn & Fels, 2015), the "theoretical papers" category includes papers that propose an explanation of the underlying nature of gamification in education and such that propose relevant pedagogies or test already existing explanatory models from other domains with respect to gamification. We also added the published literature reviews to the group of theoretical papers. The end result was a list of 11 theoretical papers appearing in the searched databases between June 30, 2014 and December 31, 2015. Thus the final number of selected papers (empirical and theoretical) amounted to 63 in total. The last column of Table 1 shows the results after filtering out irrelevant papers and removing duplicates. For comparison, the total number of papers included in the previous review covering the period January 2010–June 2014 was 34.

Following the division empirical studies vs. theoretical papers, the first part of this review covers the published empirical research on the topic, while the second part surveys briefly publications targeting theoretical aspects of educational gamification.

Data extraction

A literature survey typically employs a framework for structuring the evaluation of the works in the targeted area. This framework captures the potential properties of interest and enables a comparison of the surveyed works and drawing meaningful conclusions. The use of gamification in learning involves a number of aspects, including game elements, educational context, learning outcomes, learner profile and the gamified environment. Gamification is receiving attention, particularly for its potential to motivate learners. Accordingly, our objective involving evaluation of the level of understanding of the motivational impacts of gamification in educational contexts has shaped our decision of what categories of information to be included in the framework for evaluating the surveyed works. More specifically, we looked for information that can facilitate the process of identifying and analyzing the empirical evidence demonstrating the motivational effects of gamification. Motivation as a psychological process that gives behavior purpose and direction is contextual. Not only are individuals motivated in multiple ways, but also their motivation varies according to the situation or context of the task. To provide support for analyzing the contextual aspect, the information collected from the studies include the educational level, academic subject, and type of the gamified learning activity. We also included the used game elements, mechanics and dynamics since they are inherently related to the success of a gamification application. A number of motivation measures have been used in attempts to establish the effect of gamification on student motivation. In addition to appropriate measures, the verification of the validity of reported results

requires availability of relevant statistical information about the studies. In order to provide support for our decision on how conclusive the reported results of a study are, we added the following categories: study sample, study duration, method of data collection, and outcome. Thus the final structure of information to be derived from the reviewed studies included the following categories: game elements, educational level, academic subject, learning activity, study sample, study duration, data collection, and outcome.

Appendix 1 presents a description of the reviewed papers structured according to this framework. Obviously, the task of representing high-dimensional data in a table format is challenging, which implies a tradeoff between completeness and clarity.

Review results for empirical studies

For a systematic presentation of the review results we classify and interpret them in accordance with the described above framework.

What educational level is targeted?

Considering the educational level, the bulk of gamification studies in the survey period were conducted at university level (44 papers), with less attention to K-12 education (7 papers). At university level, 1 study has reported results involving graduate students (Nevin et al., 2014), while at K-12 level, 3 studies have reported results involving elementary school students (Boticki, Baksa, Seow, & Looi, 2015; Simoes, Mateus, Redondo, & Vilas, 2015; Su & Cheng, 2015), 2 studies have reported results involving middle school students (Attali & Arieli-Attali, 2015; Long & Aleven, 2014) and 2 studies have reported results involving high school students (Davis & Klein, 2015; Paiva, Barbosa, Batista, Pimentel, & Bittencourt, 2015). A possible explanation of this disproportion is that perhaps it is easier for college instructors to experiment with using gamification in their own courses. This might be because they are better supported technically or have necessary computer-related skills, which allow them to implement some gamification features, e.g. an electronic leaderboard. Studies involving different demographic groups however are beneficial, as we cannot necessarily generalize the results of a study conducted with one demographic group to another demographic group.

What subjects are gamified?

The collection of papers covers a wide range of academic subjects (32) organized in six categories (see Table 2). The category "Others" includes studies with unspecified subjects, where the gamified activities are independent of a subject and the focus is on: the platform supporting gamification (Barrio et al., 2015; Chang & Wei, 2015; Davis & Klein, 2015; Lambruschini & Pizarro, 2015; Mekler et al., 2015), the game elements used (Boticki et al., 2015; Pedro et al., 2015a), a personal learning environment (Morschheuser et al., 2014), measurements (Simoes et al., 2015) or learners' personalities (Tu et al., 2015).

One emerging area which is not an academic subject in its own but rather referring to a set of tools offering new affordances for enhancing students' understanding of

Table 2 Distribution of papers among subjects of studies

Subjects	Nu (%)	Papers
CS/IT	20 (39%)	(Amriani, Aji, Utomo, Wahidah, & Junus, 2014 ^a ; Anderson, Nash, & McCauley, 2015; Auvinen, Hakulinen, & Malmi, 2015; Bernik, Bubaš, & Radošević, 2015; Codish & Ravid, 2014; Codish & Ravid, 2015; Hakulinen, Auvinen, & Korhonen, 2015; Herbert, Charles, Moore, & Charles, 2014; Ibanez, Di Serio, & Delgado-Kloos, 2014; Knutas, Ikonen, Maggiorini, Ripomonti, & Porras, 2014a; Knutas, Ikonen, Nikula, & Porras, 2014b; Krause, Mogalle, Pohl, & Williams, 2015; Laskowski & Badurowicz, 2014; Leach, Laur, Code, Bebbington, & Broome, 2014; Lehtonen, Aho, Isohanni, & Mikkonen, 2015; Poole, Kemp, Patterson, & Williams, 2014; Sillaots, 2014; Sillaots, 2015; Smith, Herbert, Kavanagh, & Reidsema, 2014; Tvarozek & Brza, 2014)
Math	5 (10%)	(Attali & Arieli-Attali, 2015 ^a ; Christy & Fox, 2014; Long & Aleven, 2014; Paiva et al., 2015 ^a ; Pedro, Lopes, Prates, Vassileva, & Isotani, 2015b) ^a
Multimedia/ Communication	6 (12%)	(Barata, Gama, Jorge, & Gonçalves, 2014; Hanus & Fox, 2015; Holman et al., 2015; Jang, Park, & Yi, 2015; Leach et al., 2014; Utomo & Santoso, 2015)
Medicine/Biology/ Psychology	5 (10%)	(Bonde et al., 2014; Landers & Landers, 2015; Nevin et al., 2014; Pettit, McCoy, Kinney, & Schwartz, 2015; Su & Cheng, 2015) ^a
Languages	4 (8%)	(Hasegawa, Koshino, & Ban, 2015; Hew, Huang, Chu, & Chiu, 2016; Perry, 2015; Smith et al., 2014)
Others	11 (21%)	(Barrio, Organero, & Soriano, 2015; Boticki et al., 2015 ^a ; Chang & Wei, 2015; Davis & Klein, 2015; Lambruschini & Pizarro, 2015; Mekler, Brühlmann, Tuch, & Opwis, 2015; Morschheuser, Rivera-Pelayo, Mazarakis, & Zacharias, 2014; Pedro, Santos, Aresta, & Almeida, 2015a; Shi, Cristea, Hadzidedic, & Dervishalidovic, 2014; Sillaots, 2015 ^a ; Tu, Yen, Sujo-Montes, & Roberts, 2015)

Note: aK-12 Schools

dynamic processes and systems is interactive simulations (dynamic computer-based models which can help students observe or interact with scientific phenomena). Although gamifying the use of such simulations can help overcome the problems with insufficient motivation and engagement, there is a lack of studies evaluating the effects of gamified simulation-based learning. In this context, the work of Bonde et al. (2014), who studied the effect of combining gamification elements with simulations for improving learning effectiveness and motivation of biotech students addresses a critical gap. The results show that a gamified laboratory simulation can increase both learning outcomes and motivation levels when compared with traditional teaching. Further research is needed to examine whether these results can be extrapolated to a general tendency of the effectiveness of gamified simulations.

As shown in Table 2, the vast majority of gamification studies are dealing with Computer Science (CS) and Information Technology (IT). This fact provokes the question: Are CS and IT more suitable to gamification than the other subjects? The present studies however do not provide conclusive answer to this question. In the lack of other evidences, speculative answers can be given similar to the ones for the observed disproportion in gamifying college vs. school level activities, namely that perhaps it is easier for CS and IT instructors to experiment in their own courses. In sharp contrast, gamification experiments targeting activities related to disciplines from humanity and social sciences are extremely limited, with only one example (Holman et al., 2015) touching this subject. Another interesting observation is the low proportion of studies on gamifying STEM disciplines, excluding CS/IT and mathematics, where reinforcement of motivation is particularly beneficial: only two out of thirty two (Bonde et al., 2014) and (Su & Cheng, 2015).

What kind of learning activities is targeted?

Formal learning typically involves a mix of instructional activities and supporting materials, such as lectures, tutorials, assignments, projects, labs, exercises, class discussions and team work. A sizable part of the papers (16) studied gamification of courses as a whole, which implies gamifying a range of learning activities. Half of these are studies of gamified online courses (Amriani et al., 2014; Bernik et al., 2015; Jang et al., 2015; Krause et al., 2015; Leach et al., 2014; Sillaots, 2014; Utomo & Santoso, 2015), while the remaining part are regular courses typically with web-based learning support. Online learning normally requires stronger motivation, which makes it a somewhat more promising field for applying gamification. Although this presumes a higher concentration of studies on gamified online learning our findings indicate the opposite.

As illustrated in Table 3, the majority of works (36) studied the effect of gamification on general class activities (16) or a particular learning activity, such as exercises (6), collaboration/discussion forums (4), projects/labs (6) or tests (4). Another part of the papers addresses activities with indirect effect on learning, such as engaging students in more regular interactions with the learning environment (11). The category "Others" includes perception studies (Davis & Klein, 2015), augmented game mechanics studies (Pedro et al., 2015a), a specific activity (Mekler et al., 2015) or platform dependent studies (Su & Cheng, 2015).

Although 6 studies are addressing "Exercises", still limited attention is given to gamifying activities where students can learn through experimenting and retrying without fear of negative consequences. One observation that can be drawn from this distribution is that learning activities which involve tasks that are decomposable into simpler subtasks or tasks where performance is measurable (according to an obvious rewarding scheme or skills) are better candidates for gamification.

What combinations of game elements are studied?

According to (Deterding et al., 2011) gamification is the use of game design elements in non-game contexts. In turn, game design elements which are used in the creation of

Table 3 Distribution of papers based on learning activities

		-
Learning Activity	Nu (%)	Papers
Course driven class/online learning activities	16 (31%)	(Amriani et al., 2014; Barata et al., 2014; Bernik et al., 2015; Codish & Ravid, 2014; Hanus & Fox, 2015; Holman et al., 2015; Ibanez et al., 2014; Jang et al., 2015; Krause et al., 2015; Lambruschini & Pizarro, 2015; Laskowski & Badurowicz, 2014; Latulipe, Long, & Seminario, 2015; Leach et al., 2014; Poole et al., 2014; Sillaots, 2014; Utomo & Santoso, 2015)
Interaction with learning environment	11 (21%)	(Barrio et al., 2015; Boticki et al., 2015; Chang & Wei, 2015; Codish & Ravid, 2015; Herbert et al., 2014; Morschheuser et al., 2014; Nevin et al., 2014; Paiva et al., 2015; Pedro et al., 2015b; Perry, 2015; Pettit et al., 2015)
Exercises	6 (12%)	(Auvinen et al., 2015; Hakulinen et al., 2015; Hasegawa et al., 2015; Lehtonen et al., 2015; Long & Aleven, 2014; Tvarozek & Brza, 2014)
Collaboration/discussions/ social interactions	4 (8%)	(Knutas et al., 2014; Knutas et al., 2014; Shi et al., 2014; Smith et al., 2014)
Projects/labs	6 (12%)	(Bonde et al., 2014; Boskic & Hu, 2015; Hew et al., 2016; Landers & Landers, 2015; Sillaots, 2015; Simoes et al., 2015)
Tests	4 (8%)	(Anderson et al., 2015; Attali & Arieli-Attali, 2015; Christy & Fox, 2014; Tu et al., 2015)
Others	4 (8%)	(Davis & Klein, 2015; Mekler et al., 2015; Pedro et al., 2015a; Su & Cheng, 2015)

gamification scenarios can be divided into three categories: dynamics, mechanics and components (Werbach & Hunter, 2012). Dynamics represents the highest conceptual level in a gamified system. It includes constraints, emotions, narrative, progression and relationships. Mechanics are a set of rules that dictate the outcome of interactions within the system, while dynamics are users' responses to collections of those mechanics. The game mechanics refer to the elements that move the action forward. They include challenges, chance, competition, cooperation, feedback, resource acquisition, rewards. Components are at the basic level of the gamification process and encompass the specific instances of mechanics and dynamics. They include: achievements, avatars, badges, collections, content unlocking, gifting, leaderboards, levels, points, virtual goods, etc. For instance, points (components) provide rewards (mechanics) and create a sense of progression (dynamics). However, we note that the gamification terminology is still unsettled and various variations of the introduced above terms exist. When there is no danger of confusion, we will use the terms mechanics and dynamics to refer also to their specific instances, that is, components. Also, for consistency with our previous studies (Dicheva et al. 2015), we will use the term game elements to refer to game components.

Most of the educational gamification studies and applications are driven by the presumption that gamification in education consists chiefly of incorporating a suitable combination of game elements within learning activities. However, our review shows that the empirical studies on understanding what kind of game elements under what circumstances can drive desired behavior are not quite systematic. In the reviewed collection, 11 papers report studies of the effect of a single game element, 8 papers study gamified systems using 2 game elements, 16 papers study gamified systems with 3 game elements, while the remaining 16 papers report results of gamifying systems by incorporating more than three elements (see Table 4).

In all reviewed works with the exception of (Tu et al., 2015), which investigates the relation between gamers' personality and their game dynamics preferences, the gamification studies focus on the use of game elements (i.e. game components in terms of (Werbach & Hunter, 2012)). Typically, no justification is given for the selection of particular game elements. There is a need of more studies that can improve our understanding of how individual game elements are linked to behavioral and motivational

Table 4 Number of game elements tested in the studies

Number game elements used	Nu (%)	Papers
1 element	11 (22%)	(Attali & Arieli-Attali, 2015; Barrio et al., 2015; Boticki et al., 2015; Christy & Fox, 2014; Davis & Klein, 2015; Hakulinen et al., 2015; Landers & Landers, 2015; Long & Aleven, 2014; Pedro et al., 2015a; Tu et al., 2015; Tvarozek & Brza, 2014)
2 elements	8 (16%)	(Auvinen et al., 2015; Bonde et al., 2014; Ibanez et al., 2014; Leach et al., 2014; Paiva et al., 2015; Perry, 2015; Poole et al., 2014; Utomo & Santoso, 2015)
3 elements	16 (31%)	(Anderson et al., 2015; Auvinen et al., 2015; Bernik et al., 2015; Boskic & Hu, 2015; Codish & Ravid, 2015; Hanus & Fox, 2015; Hew et al., 2016; Knutas et al., 2014; Lambruschini & Pizarro, 2015; Laskowski & Badurowicz, 2014; Latulipe et al., 2015; Lehtonen et al., 2015; Mekler et al., 2015; Morschheuser et al., 2014; Simoes et al., 2015; Su & Cheng, 2015)
More than 3 elements	16 (31%)	(Amriani et al., 2014; Barata et al., 2014; Chang & Wei, 2015; Codish & Ravid, 2014; Hasegawa et al., 2015; Herbert et al., 2014; Holman et al., 2015; Jang et al., 2015; Knutas et al., 2014; Krause et al., 2015; Nevin et al., 2014; Pedro et al., 2015b; Pettit et al., 2015; Shi et al., 2014; Sillaots, 2014; Sillaots, 2015)

outcomes and how they function in a given educational context. Without understanding the effect of individual game elements, it is difficult to identify their contribution in studies that mix several game elements together.

The majority of gamification studies feature a subset of the following game elements: points, badges, levels, leaderboards and progress bars. This is in line with the finding of other authors, e.g. (Nicholson, 2015) that the combination of points, badges and leaderboards (sometimes referred to as PBL) is the most used one (see Table 5).

In the absence of other justification for the overuse of points, badges and leaderboards, one possible explanation is that they somewhat parallel the traditional classroom assessment model and are also easiest to implement. This combination in its trivial form can be applied to almost any context, even if there isn't a good reason to do so. Gamification with "deeper game elements" (Enders & Kapp, 2013) incorporating game design principles involving game mechanics and dynamics such as challenges, choice, low risk failure, role-play or narrative are still scarce. Only one work (Tu et al., 2015) among the reviewed studies addresses game dynamics explicitly. Studies utilizing to some extent "deeper game elements" are demonstrated in (Bonde et al., 2014; Boskic & Hu, 2015; Holman et al., 2015; Krause et al., 2015; Pettit et al., 2015). We believe that in addition to reward and feedback mechanisms, gamified systems should provide safe places where learners can gain experience without being judged or punished for failure, drawing upon approaches similar to the online learning environments proposed by (Hakulinen et al., 2015) and (Lehtonen et al., 2015), where students can improve their algorithmic skills by practicing with interactive exercises (Dichev et al. 2014).

Three questions related to the use of combinations of game elements remain open: "Do more game elements produce better results than less?", "Is the task of identifying the right combination of game elements with respect to a given context and user group practically feasible?" and "How to balance points and rewards with play and intrinsic engagement?". For answering these questions and for advancing the understanding of how to build successful gamified educational systems, there is a need for testing systems that support examining the effect of game elements and experimentally

Table 5 Game elements tested in the studies

Game elements	Nu (%)	Papers
Points only	1 (2%)	(Barrio et al., 2015)
Badges only	9 (18%)	(Boticki et al., 2015; Davis & Klein, 2015; Hakulinen et al., 2015; Leach et al., 2014; Long & Aleven, 2014; Pedro et al., 2015a; Perry, 2015; Tu et al., 2015; Tvarozek & Brza, 2014)
Leaderboards only	3 (6%)	(Christy & Fox, 2014; Landers & Landers, 2015; Poole et al., 2014)
PBL ^a	14 (27%)	(Amriani et al., 2014; Anderson et al., 2015; Barata et al., 2014; Codish & Ravid, 2014; Hanus & Fox, 2015; Hew et al., 2016; Lambruschini & Pizarro, 2015; Laskowski & Badurowicz, 2014; Latulipe et al., 2015; Lehtonen et al., 2015; Nevin et al., 2014; Pedro et al., 2015b; Sillaots, 2015; Smith et al., 2014)
Others	24 (47%)	(Attali & Arieli-Attali, 2015; Auvinen et al., 2015, Bernik et al., 2015; Bonde et al., 2014; Boskic & Hu, 2015, Chang & Wei, 2015; Codish & Ravid, 2015; Hasegawa et al., 2015; Herbert et al., 2014; Holman et al., 2015; Ibanez et al., 2014; Jang et al., 2015; Knutas et al., 2014; Knutas et al., 2014; Krause et al., 2015; Mekler et al., 2015; Morschheuser et al., 2014; Paiva et al., 2015; Pettit et al., 2015; Shi et al., 2014; Sillaots, 2014; Smith et al., 2014; Su & Cheng, 2015; Utomo & Santoso, 2015)

^aNote: PBL – Points, Badges, Leaderboards

validating it. In particular, it implies the need of gamification platforms that support easy configuration of gamified learning prototypes with specific characteristics leveraging different game features and principles.

The available evidences indicate that in a learning context gamification is more than mapping game elements on to existing learning content. It should offer stronger ways to motivate students, rather than be simply a stream of extrinsic motivators.

What types of studies?

The reviewed papers expand the scope of the empirical research on educational gamification, as compared to (Dicheva et al. 2015). Although the majority of empirical works still examine the impact of the gamification on students' engagement, performance, participation or retention, they are widening and deepening the focus of their studies. A growing body of papers is exploring a range of learning and behavioral outcomes including:

- knowledge acquisition outcomes (Jang et al., 2015); Laskowski & Badurowicz, 2014;
 Paiva et al., 2015; Su & Cheng, 2015)
- perceptual outcomes (Christy & Fox, 2014; Codish & Ravid, 2014; Davis & Klein, 2015; Pedro et al., 2015b; Sillaots, 2014; Sillaots, 2015; Christy & Fox, 2014)
- behavioral outcomes (Barata et al., 2014; Codish & Ravid, 2015; Hakulinen et al., 2015; Hew et al., 2016; Pedro et al., 2015b)
- engagement outcomes (Boskic & Hu, 2015; Chang & Wei, 2015; Ibanez et al., 2014; Latulipe et al., 2015; Morschheuser et al., 2014; Poole et al., 2014)
- *motivational outcomes* (Hasegawa et al., 2015; Herbert et al., 2014; Mekler et al., 2015; Pedro et al., 2015a; Utomo & Santoso, 2015)
- social outcomes (Hanus & Fox, 2015; Christy & Fox, 2014; Shi et al., 2014).

Under the perceptual outcome category, we have included also some works that initiate a new line of studies - the impact of gamification on different demographic groups. For example, (Pedro et al., 2015b) reported that the game mechanics implemented in a virtual learning environment did not have any effect on motivation and performance of the female students. This findings are in line with the conclusions reported by (Koivisto & Hamari, 2014), who have shown in a more general context that women experience a greater effect when the gamification contains social aspects and men - when there is a sort of competition. (Christy & Fox, 2014), on the other hand, concluded that the use of leaderboards within educational settings may act to create stereotype threat (a belief that one may be evaluated based on a negative stereotype). The results of the study found that women in the female-dominated leaderboard condition demonstrated stronger academic identification than those in the control and male-dominated leaderboard conditions. These results suggest that the use of leaderboards in academic environments can, in some circumstances, affect academic performance of different demographics differently.

The motivational outcome category concerns concepts derived from motivational principles of games such as explicit goals, rules, a feedback system, and voluntary

participation (McGonigal, 2011). Motivation is demonstrated by an individual's choice to engage in an activity and the intensity of effort or persistence in that activity. Since video games are explicitly designed for entertainment, they can produce states of desirable experience and motivate users to remain engaged in an activity with unparalleled intensity and duration. Therefore, game design was adopted as an approach for making non-game activities more enjoyable and motivating. While gamification strives at its core to increase motivation, yet motivation is not a unitary phenomenon - different people may have different types and amounts of motivation, which can be shaped by the activity they are undertaking (Gooch et al., 2016). Additionally, success in one educational context does not guarantee that the same mechanism will be motivationally successful in another educational context.

An important distinction in the motivation research is that between intrinsic and extrinsic motivation (Ryan & Deci, 2000). While extrinsic motivation relies on incentives or expected consequences of an action, intrinsic motivation stems from fulfilling the action itself. According to the Self Determination Theory (Ryan & Deci, 2000), humans seek out activities to satisfy intrinsic motivational needs, such as competence, autonomy, or relatedness. More specifically, (Ryan et al., 2006) argue that the intrinsic appeal of games is due to their ability to satisfy the basic psychological needs for competence, autonomy, and relatedness. While self-determination theory provides a good theoretical starting point for studying the motivational dynamics of 'gamified' educational activities, further research is needed to bridge motivation to a more granular level of game elements and learners' personalities. Although the connection between motivation and gamification design is demonstrated by a number of the reviewed studies, they do not add persuasive evidence confirming the effect of gamification as a motivational tool. The papers claiming to examine the motivational effects of gamification often report effects on learning outcomes instead on motivation.

The reviewed collection of empirical studies on gamifying education is very diverse with respect to the focus of the studies and the reported outcomes. This makes it difficult to find categorization that organizes the reviewed works in logical categories, captures the diversity and puts at the same time every work in a separate category. We selected a categorization with a focus on the effects of gamification on learners. It includes four categories: affective (A), behavioral (B), cognitive (C), and others. The intention with this grouping was to use it as an organizational structure for connecting outcomes with game elements and gamified activities. As under this categorization many outcomes fall into two categories, we extended it with behavioral and cognitive (B + C), affective and cognitive (A + C), and affective and behavioral (A + B) groupings. Table 6 presents the studies falling into a single category, organized in three sections: behavioral, affective, and cognitive, and connecting their outcomes with the corresponding game elements and gamified activities. Table 7 presents the studies falling into two categories, organized in the same way.

The two tables provide a more compact view, capturing the links between three key categorizing variables: game elements, gamified activities and reported outcomes. The more focused information extracted in the tables explicates data relevant to the questions guiding the study. Although the empirical work on applying gamification in educational contexts continues to grow, there is not sufficient evidence indicating noticeable progress based on collating and synthesizing the previous experiences. While

Table 6 Categorization of the studies falling into a single category (Behavioral (B), Affective (A) or Cognitive (C))

Cat	Paper	Game elements	Gamified activity	Reported Outcome
В	(Amriani et al., 2014)	Points, badges, leaderboard, status, levels, unlockable content, customization	Learning interactions in Virtual Realty	Improved participation and engagement?
	(Hew et al., 2016)	Points, badges, leaderboard	Project activities	The effect of gamification on engagement?
	(Holman et al., 2013)	Badges, leveling, autonomy, leaderboard, grade predictor	Overall class activities	The impact of grade predictor on planning the work over the course
	(Knutas et al., 2014)	Points, badges, up-vote, down-vote	Collaborative learning	Improved student collaboration?
	(Lambruschini & Pizarro, 2015)	Points, badges, leaderboard	Course participation	Boosted communication, participation, punctuality?
	(Landers & Landers, 2015)	Leaderboards	Course project activities	Improved time on task
	(Latulipe et al., 2015)	Stamps, tokens, leaderboard	In-class course activities	Encouraged harder work and engagement?
	(Lehtonen et al., 2015)	Points, badges, leaderboard	Online Java exercises	Increased use of an open learning environment
	(Smith et al., 2014)	Merit points, badges, voting	Online discussions	Improved participation and quality of online discussions
	(Utomo & Santoso, 2015)	Badges, progress bar	Online learning activities	Fostered learning activities?
Α	(Auvinen et al., 2015)	Badges, heatmap	Online exercises	Differences in reacting to gamification vs. feedback?
	(Boticki et al., 2015)	Badges	Mobile driven learning activities	Motivated a specific category of students?
	(Christy & Fox, 2014)	Leaderboard	Assessment	Created stereotype threat in specific circumstances?
	(Codish & Ravid, 2015)	Points, badges, riddles	Interactions with a Learning Management System (LMS)	Evidence that gamification behavior patterns predict playfulness
	(Hakulinen et al., 2015)	Badges	Homework exercises	Improved motivation
	(Hasegawa et al., 2015)	Points, trials, character, ranking, progress	Vocabulary learning	Motivated continuous learning
	(Morschheuser et al., 2014)	Points, badges, personas	Interactions with a Personal Learning Environment (PLE)	Increased intention to use the PLE?
	(Shi et al., 2014)	Leaderboard, progress, feedback, social status	Interactions with a learning environment	Increased intrinsic motivation?
	(Sillaots, 2014)	Points, scoreboard, goals, avatar, feedback, levels, luck, competition	In-class activities	Mixed acceptance of game elements
	(Simoes et al., 2015)	Points, badges, leaderboard	Homework in the Schooooools LMS	Increased disposition to the experience flow?
C	(Anderson et al., 2015)	Points, badges, leaderboard, competition	Assessment	Improved performance?
	(Attali & Arieli-Attali, 2015)	Points	Assessment	No effect on performance
	(Bernik et al., 2015)	Points, badges, leaderboard, progress	Learning activities within a course module	Improved performance?

Table 6 Categorization of the studies falling into a single category (Behavioral (B), Affective (A) or Cognitive (C)) (Continued)

_				
	(Jang et al., 2015)	Point, levels, life points, avatars, feedbacks, time pressure	Tutorial driven learning activities	Improved learning performance
	(Long & Aleven, 2014)	Stars/badges	Problem solving with re-practicing	No improvement of learning
	(Paiva et al., 2015)	Points, badges	Interaction with an Intelligent Tutoring System (ITS)	Correlation between points, badges and learning

Note. The first column (Cat) contains the categories, while the second column contains the papers reporting the corresponding outcomes. The question mark "?" placed after a statement in the outcomes column marks the corresponding results as inconclusive

the range of gamifying strategies is expanding, they are scattered across many different educational contexts and the aggregated information cannot confirm any emerging systematic approach yet. As it can be seen from the tables, the empirical research on gamified learning is quite fragmented. It covers studies on different configurations of game elements, used to gamify different activities and resulting in different outcomes, without any identifiable pattern of distribution. For example, the points-badgesleaderboard configuration is dominating, with 6 works studying its effect. However, the activities gamified with this configuration vary widely: project activities, course participation, online Java exercises, homework in high school LMS assessment and overall course activities. Within the category "Gamified activity" dominating is "Overall class activities" but again the configurations of game elements used to gamify it are very different: badges, leveling, autonomy, leaderboard, grade predictor; stamps, tokens, leaderboard; points, scoreboard, goals, avatar, feedback, levels, luck, competition; points, badges, leaderboard; points, leaderboard; badges, leaderboard, virtual coins, pseudonyms. The dearth of studies that build on the previous ones or parallel their efforts on exploring particular aspects of the effect of gamification on engagement and learning suggests a piecemeal approach. In the current studies that mix together points, badges, leaderboard, progress, status, etc. without a discernible systematic experimental approach, it is difficult to identify which game elements or configurations are most effective in promoting engagement and supporting learning for given activity and group of learners.

What types of goals?

We noticed that in addition to the heterogeneous nature of the empirical research, the stated goals of the studies were not always in line with the reported outcomes. To provide an additional dimension for organizing and examining the links between the corresponding categorizing variables we further grouped the studies according to their stated goal (see Appendix 2, which lists the reviewed studies along with their goals). The two top categories for grouping the studies based on the study goals are: learner-centric and platform-centric (see Table 8). The bulk of works which expands and differentiates the earlier research on the effect of gamification on learners (e.g. (Dicheva et al., 2015)) falls in the first category (44 papers). This category includes 4 subcategories grouping further the studies as follows:

Table 7 Categorization of studies falling in two categories: Behavioral and Cognitive (B + C), Affective and Cognitive (A + C), Affective and Behavioral (A + B)

	Paper	Game elements	Gamified activity	Reported Outcome
B +	(Boskic & Hu, 2015)	Choice, role playing, feedback	Course assignments	Increased engagement understanding?
С	(Ibanez et al., 2014)	Badges, leaderboard	Course activities	Confirmation of the effect of gamification on engagement and learning?
	(Krause et al., 2015)	Points, achievements, leaderboards, avatars	Online course activities	Improved retention period and learning performance
	(Laskowski & Badurowicz, 2014)	Points, badges, leaderboard	Overall course activities	Improved engagement and performance?
	(Lehtonen et al., 2015)	Points, badges,	Online class activities	Increased online activities and learning performance?
	(Nevin et al., 2014)	Badges, levels, feedback, leaderboard, voluntary participation	Interactions with a learning environment	Increased knowledge retention, reduced attrition
	(Pedro et al., 2015b)	Points, badges, levels, feedback, ranking	Interactions with a Virtual Learning Environment (VLE)	Improved performance, reduced undesirable behaviors in VLE?
	(Pettit et al., 2015)	Challenge, competition, progress, status, achievement, prizes, chance, surprise, anticipation, humor	Interactions with an Audience Response System (ARS)	Increased engagement and learning
	(Poole et al., 2014)	Points, leaderboards	Class activities	Increased engagement, fosters learning?
A + C	(Barrio et al., 2015)	Points	Learning activities using a Student Response System (SRS)	Improved motivation, attention, learning performance?
	(Bonde et al., 2014)	Simulation, narrative, fictional characters	Lab activities	Increased learning outcomes and motivation
	(Hanus & Fox, 2015)	Badges, leaderboard, virtual coins, pseudonyms	In-class and out-of-class activities	Improving satisfaction, empowerment, academic performance not confirmed
	(Mekler et al., 2015)	Points, levels, leaderboard	Image annotation	Increased competence need and performance?
	(Su & Cheng, 2015)	Badges, leaderboard, missions	Field activities	Increased motivation and learning?
A + B	(Perry, 2015)	Points, badges	Interactions with a language learning system	Increased playfulness and engagement in learning?

- A. Behavioral and cognitive results: focusing on behavioral and cognitive effects caused by gamification.
- B. Categories of learners: focusing on the effects of gamification on different groups of
- C. Learners' perception: focusing on the learner's perception of different game mechanics and principles.
- D. Measures: focusing on the measures used for assessing the outcomes.

Table 8 Paper distribution according to the type of study

Category	Type of study	Nu (%)	Papers
	Behavioral and cognitive results	31 (61%)	(Amriani et al., 2014; Anderson et al., 2015; Attali & Arieli-Attali, 2015; Auvinen et al., 2015; Bernik et al., 2015; Bonde et al., 2014; Boskic & Hu, 2015; Boticki et al., 2015; Christy & Fox, 2014; Hakulinen et al., 2015; Hanus & Fox, 2015; Hasegawa et al., 2015; Ibanez et al., 2014; Jang et al., 2015; Knutas et al., 2014b; Krause et al., 2015; Landers & Landers, 2015; Laskowski & Badurowicz, 2014; Latulipe et al., 2015; Leach et al., 2014; Lehtonen et al., 2015; Long & Aleven, 2014; Morschheuser et al., 2014; Nevin et al., 2014; Pedro et al., 2015; Perry, 2015; Poole et al., 2014; Shi et al., 2014; Sillaots, 2014; Smith et al., 2014; Utomo & Santoso, 2015)
Learner-Centric	Categories of learners	6 (12%)	(Barata et al., 2014; Codish & Ravid, 2014; Herbert et al., 2014; Hew et al., 2016; Mekler et al., 2015; Tu et al., 2015)
	Learners' perception	4 (8%)	(Davis & Klein, 2015; Knutas et al., 2014a; Paiva et al., 2015; Sillaots, 2015)
	Measures	3 (6%)	(Codish & Ravid, 2015; Holman et al., 2015; Simoes et al., 2015)
Platform-Centric	Game elements and gamified platforms	7 (13%)	(Barrio et al., 2015; Chang & Wei, 2015; Lambruschini & Pizarro, 2015; Pedro et al., 2015a; Pettit et al., 2015; Su & Cheng, 2015; Tvarozek & Brza, 2014)

These four groups cover a wide variety of goals. Group A includes studies of the effectiveness of gamification in the classroom longitudinally (Hanus & Fox, 2015); the impact of gamification on retention and learning success (Jang et al., 2015; Krause et al., 2015), on participation and quality of online discussions (Smith et al., 2014), on reducing undesirable behaviors and increasing performance in virtual learning environments (Pedro et al., 2015b) and in personal learning environments (Lehtonen et al., 2015; Morschheuser et al., 2014); the effect of badges on student behavior (Hakulinen et al., 2015) and how they predict the student exam success (Boticki et al., 2015); the causal effect of gamifying a course project with leaderboards (Landers & Landers, 2015); the learning effectiveness of a gamified simulation (Bonde et al., 2014) and the effect of transforming a traditional course into a role-playing game (Boskic & Hu, 2015).

A progress has been made within the learner-centric category with explorations of psychological effects of gamification which can be summarized by the question: How students with different personalities, dispositions and learning styles are influenced by game elements? While in our first review the question shared between most of the papers was "Is gamification effective?", now it appears in a more extended version, in combination with the questions "for what?" or "to whom?".

Group B includes papers identifying learner types based on how students experience gamified courses (Barata et al., 2014) and how different learners perceive playfulness (Codish & Ravid, 2014), on the variation in motivation between learners with different gamification typologies (Herbert et al., 2014), on exploring whether points, leader-boards, and levels increase performance, competence, need satisfaction, and intrinsic motivation (Mekler et al., 2015), on involving Asian students in gamified course activities (Hew et al., 2016) and on the predictive effect of gaming personality on their game dynamic preferences (Tu et al., 2015). Even though the amount of papers addressing the question "to whom" is still limited, an emerging shared message particularly relevant to instructional designers recognizes that what one learner values, another may not, what one learner believes is achievable, another may not. Understanding

differences in learners' drivers, what they value and what they dislike is important to the design of reward, progress, and feedback systems with potential to achieve desired outcomes for the intended groups of learners.

Group C includes papers on students' perceptions of simple game elements such as badges (Davis & Klein, 2015) or combination of points and badges (Paiva et al., 2015). It also includes studies on how students perceive a game-like course (Sillaots, 2015) and on profiling learners based on their gamification preferences (Knutas et al., 2014).

Another emerging topic in this category groups the works on measuring the impact of gamification (Group D). This group includes papers on the impact of gamification on students' engagement and how to measure that impact (Simoes et al., 2015), papers on the effectiveness of gamification behavior patterns as a measure of playfulness (Codish & Ravid, 2015), and how predictive measurements can help students plan their pathways in gamified courses (Holman et al., 2015). While gamification is promoted as a motivational instrument, studies measuring its motivational effects are still limited.

In the second category we have placed 7 articles, which study the effect of incorporating selected game elements or game principles into specific learning platforms or experiment with conventional game elements by assigning them new roles. This category includes studies of employing gamification in audience response systems (Barrio et al., 2015; Pettit et al., 2015), in mobile learning systems (Su & Cheng, 2015), in Learning Management Systems (Lambruschini & Pizarro, 2015) and in MOOCs (Chang & Wei, 2015). Two papers explore creating badges as a tool for measuring students' interest (Tvarozek & Brza, 2014) and the effect of collaborative badge creation on engagement and motivation (Pedro et al., 2015a). The papers listed in the platform-centric category do not cover all gamified platforms proposed in the reviewed papers. When according to our judgment the focus of a paper was on behavioral effects, as for example in (Krause et al., 2015), that paper was included in the first category. The availability of successful gamified platforms will help widen the scope of gamified educational activities and create a ground for broadening experimental studies towards developing evidence based practices.

How conclusive are the reported results?

One of the evolving goals of this review was to take a closer look at the supporting evidence for the 'positive' or 'negative' results of the empirical studies as reported by their authors. This was provoked by the fact that some of the papers studying the effect of gamification on learners reported a mix of positive and negative results, other were inconclusive, and yet other expressed a degree of caution, while the strength of the evidence backing the positive and negative results were varying significantly.

A common pattern observed in most studies is to design and develop a particular gamified course/activity/environment, test it in a pilot and assess users' approvals and gains in performance. The reported outcome often concludes that the gamification produced the pursued learning gains and that the users appreciated the added gamification features. Irrespective of the goals of the studies, the works on gamifying education should be subject to the same level of skepticism and scrutiny that is applied to any other areas of empirical research. In order to improve our understanding and to offer a

more realistic picture of the nature of the effects of using gamification in education, consistent with the presented evidence, we undertook a more in-depth examination of the reviewed papers with a focus on both the reported outcomes and how they have been obtained. The primary aim of this effort was twofold: (i) to provide a critical review, questioning the validity of some reported outcomes, and (ii) to offer a picture that avoids the harmful effects of an one-sided viewpoint.

Our decision on the validity of the gamification studies was guided by the following factors: the sample size, the number of study groups, the length of the study, how the data was collected, how the variables were controlled, how and by what statistical procedures the data was analyzed, how well the conclusions are supported by the data, and does the study give enough information to convince the reader in the correctness of the evaluation conclusions. The examination of the selected papers indicated that the empirical studies tended to use surveys and quasi-experimental designs, while the randomized controlled trials were less common. According to the nature of the empirical study, the papers were partitioned into two major categories: ABC studies, which target Affective/Behavioral/Cognitive outcomes, and non-ABC studies. The ABC studies were further partitioned into three subcategories: positive, negative and inconclusive, based on the reliability of the evidence for the reported ABC outcome. The outcomes were marked as "positive" if valid evidence confirms the claim and marked as "negative" if the evidence confirms its negation. The studies were marked as "inconclusive" if the presented evidence was judged as insufficient based on inadequacies, such as small sample sizes, lack of comparison groups, use of purely descriptive statistics, short experiment timeframes, and unreliable statistical evidence. For example, reported positive effects of gamification based on a two-week study could be attributed to the 'novelty effect' of the used tool or approach rather than to the added gamification features. In the inconclusive category we also included papers studying gamification in combination with some other factors, which make uncertain whether the observed effects can be attributed to the gamification or to the other variables, as well as papers where no positive effect was found but negative effect was not discernible either.

The classification of papers in accordance with our judgment of the degree of validity of the reported results is presented in Table 9 and the proportions of the resulting grouping of the ABC papers in Fig. 1.

The paper grouping, based on the strengths of the presented evidence, reveals that the high expectations for positive outcomes from gamified learning are not confirmed by the results of the reviewed empirical studies (see Fig. 1).

The examination of the papers shows that from the 41 ABC empirical studies only 15 present conclusive evidences for the reported outcomes. In those 15 papers, the findings related to the benefits of gamification are mixed: 12 studies present evidence for positive effects of gamification in educational settings, while 3 present evidence for negative effects. A surprising fact is that the vast majority of the empirical works (25 studies) report inconclusive outcomes, which means that there is no basis for confidence in the reported results. Such outcomes obscure the level of progress in the area of educational gamification. Table 10 and Table 11 below are obtained from Table 6 and Table 7, correspondingly, by eliminating the studies marked as inconclusive. With this relatively small number of (15) papers and a diverse specter

Table 9 Categorizations of the Studies based on the outcomes and presented degree of evidence

Results	Nu (%)	Papers
ABC studies	41 (79%)	
Positive	12 (23%)	(Bonde et al., 2014; Hakulinen et al., 2015; Hasegawa et al., 2015; Holman et al., 2013; Holman et al., 2015; Krause et al., 2015; Landers & Landers, 2015; Lehtonen et al., 2015; Nevin et al., 2014; Paiva et al., 2015; Pettit et al., 2015; Smith et al., 2014)
Negative	3 (6%)	(Attali & Arieli-Attali, 2015; Hanus & Fox, 2015; Long & Aleven, 2014)
Inconclusive	26 (50%)	(Amriani et al., 2014; Anderson et al., 2015; Auvinen et al., 2015; Barrio et al., 2015; Bernik et al., 2015; Boskic & Hu, 2015; Boticki et al., 2015; Christy & Fox, 2014; Codish & Ravid, 2015; Hew et al., 2016; Ibanez et al., 2014; Knutas et al., 2014b; Lambruschini & Pizarro, 2015; Laskowski & Badurowicz, 2014; Latulipe et al., 2015; Leach et al., 2014; Mekler et al., 2015; Morschheuser et al., 2014; Pedro et al., 2015; Perry, 2015; Poole et al., 2014; Shi et al., 2014; Sillaots, 2014; Simoes et al., 2015; Su & Cheng, 2015; Utomo & Santoso, 2015)
Non-ABC	11 (21%)	(Barata et al., 2014; Chang & Wei, 2015; Codish & Ravid, 2014; Davis & Klein, 2015; Herbert et al., 2014; Holman et al., 2015; Knutas et al., 2014a; Pedro et al., 2015a; Sillaots, 2015; Tu et al., 2015; Tvarozek & Brza, 2014)

of game elements and activities, the presented outcomes are insufficient to draw definitive conclusions on the effectiveness of gamification on students' engagement, learning or participation. This judgement can be interpreted as an answer to the first guiding question about the existing empirical evidence for the impact of gamification on motivational processes and effectiveness of learning. Currently, there is a dearth of quality empirical evidence to support general claims of the impact of gamification on student' learning and motivation. Whilst 12 studies report encouraging outcomes, they cover a range of specific combinations of game elements, specific activities and outcomes and thus do not support practical generalization. It would be short-sighted to assume that gamified implementations with the same configurations of game elements would function similarly across different educational contexts. For example, (Hakulinen et al., 2015) present convincing evidence that points, badges and leaderboard incorporated in online Java exercises increases the use of an open learning environment. However, with the current understanding of the motivational mechanisms afforded by gamification, we cannot generalize this claim to other activities, game element combinations or academic subjects. In general, studies reporting positive results from using a specific combination of game elements do not promote the understanding of the causal effect of the combination, as it is unclear whether the combination or a particular element led to the positive outcome

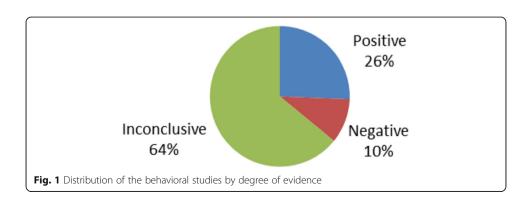


Table 10 Categorization of studies falling into a single category: Behavioral (B), Affective (A) or Cognitive (C)

Cat	Paper	Game elements	Gamified activity	Reported Outcome
В	(Holman et al., 2013)	Badges, leveling, autonomy, leaderboard, grade predictor	Overall class activities	The impact of grade predictor on planning the work over the course
	(Landers & Landers, 2015)	Leaderboards	Course project activities	Improves time on task
	(Lehtonen et al., 2015)	Points, badges, leaderboard	Online Java exercises	Increases the use of an open learning environment
	(Smith et al., 2014)	Merit points, badges, voting	Online discussions	Improves participation and quality of online discussions
Α	(Hakulinen et al., 2015)	Badges	Homework exercises	Improves motivation
	(Hasegawa et al., 2015)	Points, trials, character, ranking, progress	Vocabulary learning	Motivates continuous learning
C	(Attali & Arieli-Attali, 2015)	Points	Assessment	No effect on performance
	(Jang et al., 2015)	Points, levels, life points, avatars, feedbacks, time pressure	Tutorial driven learning activities	Improves learning performance
	(Long & Aleven, 2014)	Stars/badges	Problem solving with re-practicing	Does not improve learning
	(Paiva et al., 2015)	Points, badges	Interaction with ITS	Correlation between points, badges and learning

(e.g. Bonde et al., 2014; Jang et al., 2015). Negative results such as those of Hanus and Fox (2015), who reported that badges, leaderboard, virtual coins, and pseudonyms incorporated in a communication course can have a detrimental effect on students' motivation, satisfaction, and empowerment, help understand the limits of gamification. Again, the results obtained from such studies should be interpreted in a restricted manner, for the specific combinations of game elements, gamified activity, academic subject, and age group. The piecemeal approach observed in the reviewed studies slows down the advancement in the understanding of the effect of incorporating game elements in learning activities. From the 14 studies listed in the two tables, with 14 different combinations of game elements and 15 different gamified activities, it is

Table 11 Categorization of studies falling in two categories: Behavioral and Cognitive (B + C), Affective and Cognitive (A + C), Affective and Behavioral (A + B)

	Paper	Paper Game elements		Reported Outcome
B + C	(Krause et al., 2015)	Points, achievements, leaderboards, avatars	Online course activities	Improves retention period and learning performance
	(Nevin et al., 2014)	Badges, levels, feedback, leaderboard, voluntary participation	Interactions with a learning environment	Increases knowledge retention, reduces attrition
	(Pettit et al., 2015)	Challenge, competition, progress, status, achievement, prizes, chance, surprise, anticipation, humor	ARS interactions	Increases engagement and learning
A + C	(Bonde et al., 2014)	Simulation, narrative, fictional characters	Lab activity	Increases learning outcomes and motivation
	(Hanus & Fox, 2015)	Badges, leaderboard, virtual coins, pseudonyms	In-class and out-of-class activities	No improvement of satisfaction, empowerment, academic performance

difficult to derive useful information on how to gamify a new (different) activity with predictable outcomes. For example, two papers (Hakulinen et al., 2015) and (Landers & Landers, 2015) report positive outcomes for using single game elements, but one is for badges and the other one for leaderboards. On the other hand, two of the studies reporting negative results deal with Mathematics (Attali & Arieli-Attali, 2015; Long & Aleven, 2014). But, in these two cases, the game elements, the learning activities, the student level and the gender vary. In addition, the mix of badges, levels, leaderboards, progress, feedback, status and avatars used in the conclusive studies makes it hard to know which of these elements actually worked. Furthermore, the fundamental differences in the studied educational contexts hamper the transfer of experimented practices from one learning situation to another. All this suggests a need for a more systematic program of experimental studies.

We note that our judgment in studying inconclusiveness can be viewed as rather subjective. Therefore, Table 12 presents the papers judged as "inconclusive" along with a short explanation for placing them in this group. In several cases our judgment simply conveys the paper's conclusion where the authors themselves acknowledge that the results of the study should be interpreted with caution.

While it seems apparent that gamification has the potential to create enhanced learning environments, there is still insufficient evidence that it (1) produces reliable, valid and long-lasting educational outcomes, or (2) does so better than traditional educational models. There is still insufficient empirical work that investigates the educational potential of gamification in a rigorous manner. Increasing the number of studies that use randomized controlled trials or quasi-experimental designs will increase the scientific robustness. The continued (and coordinated) collection of evidence, that is, data that substantiate the successes and failures of gamification, remains crucial for building an empirical knowledge base and consolidating best practices, extracting guidelines and eventually developing predictive theories. It is necessary to strengthen the methodical base of gamified learning and systematically enlarge the body of evidence that explains what factors and conditions produce desirable outcomes. The empirical research should thereby not just be fixated on the pros of gamified learning, but also be open to the cons and the conditions when gamification for learning should be avoided (Linehan et al., 2011; Westera, 2015).

Indirectly related to the conclusiveness of the reported results are the measurements used. A significant number of the studies (15) are using performance as a measure of the effect that gamification has on the studied activities. This is understandable for several reasons. First, the driving criterion for adopting any technology in education is whether and how much it can improve learning. Second, one can argue that high learner performance provides evidence of learners motivation since performance has been shown to correlate with learner's motivation. However, such an approach is imperfect. Performance is an indirect measure of motivation that is influenced by many non-motivational factors such as ability, prior knowledge, and quality of instruction, while motivation is the actual driving force which makes individuals want to do something and help them continue doing it. Therefore, it is beneficial to understand the motivational triggers that engage learners. This suggests a need of studies that utilize more reliable measures of motivation and characterize better how gamification influences learner motivation and consequently how it improves learner engagement and

Table 12 Inconclusive studies with reasons for this categorization

Amriani et al., 2014) Small population size (38 students); Short period of study (2 weeks) Study several variables - unclear which one contributed to the observed outcomes (Auvinen et al., 2015) Comparison of two approaches - a gamified version and a version with heatmaps Barrio et al., 2015) Short study period (four 90 min sessions) (Bernik et al., 2015) Short study period (two weeks) and limited context (Boskic & Hu, 2015) Reported statistical results show medium effect size and low R squared values (Boticki et al., 2015) Codish & Ravid, 2015) Limited sample size and context (acknowledged by the authors) (Hew et al., 2016) Short study period (23+ 19 days) with small population size (22 + 43) Evidence are inconclusive as acknowledged by the authors + Small population size (22 students) (Knutas et al. 2014) Small population size; No comparison group Lambruschini & Pizarro, 2015) Laskowski & Badurowicz, 2014) Reported positive and negative results (the gamified group had lower average grade) Latulipe et al., 2014) Positive and negative results with no evidence of improved performance (Leach et al., 2014) Positive and negative results with no evidence of improved learning Mekler et al., 2015) Small population size (16 students) for statistical significance; Short study period (20 days) Small population size (11 students). No control group. Weak evidence. Short study period (3 weeks) Short study period (two weeks); Small population size (20 students)		<u> </u>
Anderson et al., 2015) Study several variables - unclear which one contributed to the observed outcomes Cauvinen et al., 2015) Comparison of two approaches - a gamified version and a version with heatmaps Short study period (four 90 min sessions) Short study period (two weeks) and limited context Weak evaluation, based on interviews with the instructor who taught all 7 classes and one student Reported statistical results show medium effect size and low R squared values (Codish & Ravid, 2015) Chew et al., 2016) Short study period (23+ 19 days) with small population size (22+43) (Ibanez et al., 2014) Evidence are inconclusive as acknowledged by the authors + Small population size (22 students) Knutas et al. 2014b) Small population size; No comparison group (Laskowski & Badurowicz, 2014) Reported positive and negative results (the gamified group had lower average grade) (Latulipe et al., 2015) Study 3 factors at the same time, unclear which one led to improved performance (Leach et al., 2015) Short study session (22 min) with self-reported measurement Mekler et al., 2015) Short study session (22 min) with self-reported measurement Short study period (1 h) Perry, 2015) Small population size (16 students) for statistical significance; Short study period (1 h) Short study period (3 weeks) Short study period (two weeks); Small population size (20 students)	Paper No	Reason (data collection/statistical evidence)
Outcomes (Auvinen et al., 2015) Comparison of two approaches - a gamified version and a version with heatmaps (Barrio et al., 2015) Short study period (four 90 min sessions) (Bernik et al., 2015) Short study period (two weeks) and limited context (Weak evaluation, based on interviews with the instructor who taught all 7 classes and one student (Boticki et al., 2015) Reported statistical results show medium effect size and low R squared values (Codish & Ravid, 2015) Limited sample size and context (acknowledged by the authors) (Hew et al., 2016) Short study period (23+ 19 days) with small population size (22 + 43) (Ibanez et al., 2014) Evidence are inconclusive as acknowledged by the authors + Small population size (22 students) (Knutas et al. 2014b) Small population size; No comparison group (Laskowski & Badurowicz, 2014) Reported positive and negative results (the gamified group had lower average grade) (Latulipe et al., 2015) Study 3 factors at the same time, unclear which one led to improved performance (Leach et al., 2014) Positive and negative results with no evidence of improved learning (Mekler et al., 2015) Short study session (22 min) with self-reported measurement (Morschheuser et al., 2014) Short study period (20 days) (Pedro et al., 2015b) Small population size (16 students) for statistical significance; Short study period (1 h) (Perry, 2015) Small population size (11 students). No control group. Weak evidence. (Poole et al., 2014) Short study period (3 weeks) Short study period (1 two weeks); Small population size (20 students)	(Amriani et al., 2014)	Small population size (38 students); Short period of study (2 weeks)
heatmaps Short study period (four 90 min sessions) Short study period (two weeks) and limited context Weak evaluation, based on interviews with the instructor who taught all 7 classes and one student Boskic & Hu, 2015) Reported statistical results show medium effect size and low R squared values Codish & Ravid, 2015) Limited sample size and context (acknowledged by the authors) Short study period (23+ 19 days) with small population size (22 + 43) Evidence are inconclusive as acknowledged by the authors + Small population size (22 students) Knutas et al. 2014b) Small population size; No comparison group Lambruschini & Pizarro, 2015) Small population size, No comparison group Laskowski & Badurowicz, 2014) Reported positive and negative results (the gamified group had lower average grade) Latulipe et al., 2015) Study 3 factors at the same time, unclear which one led to improved performance Leach et al., 2014) Positive and negative results with no evidence of improved learning Mekler et al., 2015) Short study session (22 min) with self-reported measurement Morschheuser et al., 2014) Short study period (16 students) for statistical significance; Short study period (1 h) Perry, 2015) Small population size (11 students). No control group. Weak evidence. Poole et al., 2014) Short study period (3 weeks) Short study period (10 weeks); Small population size (20 students)	(Anderson et al., 2015)	
Short study period (two weeks) and limited context Weak evaluation, based on interviews with the instructor who taught all 7 classes and one student Boticki et al., 2015) Reported statistical results show medium effect size and low R squared values (Codish & Ravid, 2015) Limited sample size and context (acknowledged by the authors) Short study period (23+ 19 days) with small population size (22 + 43) (Ibanez et al., 2014) Evidence are inconclusive as acknowledged by the authors + Small population size (22 students) Knutas et al. 2014b) Small population size; No comparison group (Laskowski & Badurowicz, 2014) Reported positive and negative results (the gamified group had lower average grade) (Latulipe et al., 2015) Study 3 factors at the same time, unclear which one led to improved performance (Leach et al., 2014) Positive and negative results with no evidence of improved learning (Mekler et al., 2015) Short study session (22 min) with self-reported measurement Morschheuser et al., 2014) Short study period (20 days) Pedro et al., 2015) Small population size (16 students) for statistical significance; Short study period (1 h) Perry, 2015) Small population size (11 students). No control group. Weak evidence. Poole et al., 2014) Short study period (3 weeks) Short study period (two weeks); Small population size (20 students)	(Auvinen et al., 2015)	
Weak evaluation, based on interviews with the instructor who taught all 7 classes and one student Reported statistical results show medium effect size and low R squared values (Codish & Ravid, 2015) Limited sample size and context (acknowledged by the authors) (Hew et al., 2016) Short study period (23+ 19 days) with small population size (22 + 43) Evidence are inconclusive as acknowledged by the authors + Small population size (22 students) (Knutas et al. 2014b) Small population size, No comparison group (Laskowski & Badurowicz, 2014) Reported positive and negative results (the gamified group had lower average grade) (Latulipe et al., 2015) Study 3 factors at the same time, unclear which one led to improved performance (Leach et al., 2014) Positive and negative results with no evidence of improved learning (Mekler et al., 2015) Short study session (22 min) with self-reported measurement (Morschheuser et al., 2014) Short study period (20 days) Pedro et al., 2015) Small population size (16 students) for statistical significance; Short study period (1 h) Perry, 2015) Small population size (11 students). No control group. Weak evidence. Poole et al., 2014) Short study period (3 weeks) Short study period (two weeks); Small population size (20 students)	(Barrio et al., 2015)	Short study period (four 90 min sessions)
all 7 classes and one student Reported statistical results show medium effect size and low R squared values (Codish & Ravid, 2015) Limited sample size and context (acknowledged by the authors) (Hew et al., 2016) Short study period (23+ 19 days) with small population size (22 + 43) (Ibanez et al., 2014) Evidence are inconclusive as acknowledged by the authors + Small population size (22 students) (Knutas et al. 2014b) Small population size; No comparison group (Laskowski & Badurowicz, 2014) Reported positive and negative results (the gamified group had lower average grade) (Latulipe et al., 2015) Study 3 factors at the same time, unclear which one led to improved performance (Leach et al., 2014) Positive and negative results with no evidence of improved learning (Mekler et al., 2015) Short study session (22 min) with self-reported measurement (Morschheuser et al., 2014) Short study period (20 days) (Pedro et al., 2015) Small population size (16 students) for statistical significance; Short study period (1 h) (Perry, 2015) Small population size (11 students). No control group. Weak evidence. (Poole et al., 2014) Short study period (3 weeks) (Shi et al., 2014) Short study period (two weeks); Small population size (20 students)	(Bernik et al., 2015)	Short study period (two weeks) and limited context
Limited sample size and context (acknowledged by the authors) Short study period (23+ 19 days) with small population size (22 + 43) Evidence are inconclusive as acknowledged by the authors + Small population size (22 students) Small population size; No comparison group Clambruschini & Pizarro, 2015) Small population size. No comparison group Claskowski & Badurowicz, 2014) Reported positive and negative results (the gamified group had lower average grade) Clatulipe et al., 2015) Study 3 factors at the same time, unclear which one led to improved performance Cleach et al., 2014) Positive and negative results with no evidence of improved learning Mekler et al., 2015) Short study session (22 min) with self-reported measurement Morschheuser et al., 2014) Short study period (20 days) Pedro et al., 2015) Small population size (16 students) for statistical significance; Short study period (1 h) Perry, 2015) Small population size (11 students). No control group. Weak evidence. Poole et al., 2014) Short study period (3 weeks) Shi et al., 2014) Short study period (two weeks); Small population size (20 students)	(Boskic & Hu, 2015)	
(Hew et al., 2016) Short study period (23+ 19 days) with small population size (22 + 43) (Ibanez et al., 2014) Evidence are inconclusive as acknowledged by the authors + Small population size (22 students) (Knutas et al. 2014b) Small population size; No comparison group (Lambruschini & Pizarro, 2015) Small population size. No comparison group (Laskowski & Badurowicz, 2014) Reported positive and negative results (the gamified group had lower average grade) (Latulipe et al., 2015) Study 3 factors at the same time, unclear which one led to improved performance (Leach et al., 2014) Positive and negative results with no evidence of improved learning (Mekler et al., 2015) Short study session (22 min) with self-reported measurement (Morschheuser et al., 2014) Short study period (20 days) (Pedro et al., 2015) Small population size (16 students) for statistical significance; Short study period (1 h) (Perry, 2015) Small population size (11 students). No control group. Weak evidence. (Poole et al., 2014) Short study period (3 weeks) (Shi et al., 2014) Short study period (two weeks); Small population size (20 students)	(Boticki et al., 2015)	Reported statistical results show medium effect size and low R squared values
Evidence are inconclusive as acknowledged by the authors + Small population size (22 students) Small population size; No comparison group (Lambruschini & Pizarro, 2015) Small population size. No comparison group (Laskowski & Badurowicz, 2014) Reported positive and negative results (the gamified group had lower average grade) (Latulipe et al., 2015) Study 3 factors at the same time, unclear which one led to improved performance (Leach et al., 2014) Positive and negative results with no evidence of improved learning (Mekler et al., 2015) Short study session (22 min) with self-reported measurement (Morschheuser et al., 2014) Short study period (20 days) (Pedro et al., 2015b) Small population size (16 students) for statistical significance; Short study period (1 h) (Perry, 2015) Small population size (11 students). No control group. Weak evidence. (Poole et al., 2014) Short study period (3 weeks) Shi et al., 2014) Short study period (two weeks); Small population size (20 students)	(Codish & Ravid, 2015)	Limited sample size and context (acknowledged by the authors)
population size (22 students) Knutas et al. 2014b) Small population size; No comparison group (Lambruschini & Pizarro, 2015) Small population size. No comparison group (Laskowski & Badurowicz, 2014) Reported positive and negative results (the gamified group had lower average grade) Study 3 factors at the same time, unclear which one led to improved performance (Leach et al., 2015) Short study session (22 min) with self-reported measurement (Morschheuser et al., 2014) Short study period (20 days) (Pedro et al., 2015) Small population size (16 students) for statistical significance; Short study period (1 h) Small population size (11 students). No control group. Weak evidence. (Poole et al., 2014) Short study period (3 weeks) Shi et al., 2014) Short study period (two weeks); Small population size (20 students)	(Hew et al., 2016)	Short study period (23+ 19 days) with small population size (22+43)
(Lambruschini & Pizarro, 2015) Small population size. No comparison group (Laskowski & Badurowicz, 2014) Reported positive and negative results (the gamified group had lower average grade) (Latulipe et al., 2015) Study 3 factors at the same time, unclear which one led to improved performance (Leach et al., 2014) Positive and negative results with no evidence of improved learning (Mekler et al., 2015) Short study session (22 min) with self-reported measurement (Morschheuser et al., 2014) Short study period (20 days) (Pedro et al., 2015) Small population size (16 students) for statistical significance; Short study period (1 h) (Perry, 2015) Small population size (11 students). No control group. Weak evidence. (Poole et al., 2014) Short study period (3 weeks) (Shi et al., 2014) Short study period (two weeks); Small population size (20 students)	(Ibanez et al., 2014)	
(Laskowski & Badurowicz, 2014) Reported positive and negative results (the gamified group had lower average grade) Study 3 factors at the same time, unclear which one led to improved performance (Leach et al., 2014) Positive and negative results with no evidence of improved learning (Mekler et al., 2015) Short study session (22 min) with self-reported measurement (Morschheuser et al., 2014) Short study period (20 days) (Pedro et al., 2015) Small population size (16 students) for statistical significance; Short study period (1 h) (Perry, 2015) Small population size (11 students). No control group. Weak evidence. (Poole et al., 2014) Short study period (3 weeks) (Shi et al., 2014) Short study period (two weeks); Small population size (20 students)	(Knutas et al. 2014b)	Small population size; No comparison group
lower average grade) Study 3 factors at the same time, unclear which one led to improved performance (Leach et al., 2014) Positive and negative results with no evidence of improved learning (Mekler et al., 2015) Short study session (22 min) with self-reported measurement (Morschheuser et al., 2014) Short study period (20 days) (Pedro et al., 2015b) Small population size (16 students) for statistical significance; Short study period (1 h) (Perry, 2015) Small population size (11 students). No control group. Weak evidence. (Poole et al., 2014) Short study period (3 weeks) Shi et al., 2014) Short study period (two weeks); Small population size (20 students)	(Lambruschini & Pizarro, 2015)	Small population size. No comparison group
performance (Leach et al., 2014) Positive and negative results with no evidence of improved learning (Mekler et al., 2015) Short study session (22 min) with self-reported measurement (Morschheuser et al., 2014) Short study period (20 days) (Pedro et al., 2015b) Small population size (16 students) for statistical significance; Short study period (1 h) (Perry, 2015) Small population size (11 students). No control group. Weak evidence. (Poole et al., 2014) Short study period (3 weeks) (Shi et al., 2014) Short study period (two weeks); Small population size (20 students)	(Laskowski & Badurowicz, 2014)	
Mekler et al., 2015) Short study session (22 min) with self-reported measurement Morschheuser et al., 2014) Short study period (20 days) Small population size (16 students) for statistical significance; Short study period (1 h) Small population size (11 students). No control group. Weak evidence. (Poole et al., 2014) Short study period (3 weeks) Shi et al., 2014) Short study period (two weeks); Small population size (20 students)	(Latulipe et al., 2015)	
Morschheuser et al., 2014) Short study period (20 days) Small population size (16 students) for statistical significance; Short study period (1 h) Small population size (11 students). No control group. Weak evidence. Short study period (3 weeks) Shi et al., 2014) Short study period (two weeks); Small population size (20 students)	(Leach et al., 2014)	Positive and negative results with no evidence of improved learning
(Pedro et al., 2015b) Small population size (16 students) for statistical significance; Short study period (1 h) Small population size (11 students). No control group. Weak evidence. (Poole et al., 2014) Short study period (3 weeks) Shi et al., 2014) Short study period (two weeks); Small population size (20 students)	(Mekler et al., 2015)	Short study session (22 min) with self-reported measurement
Short study period (1 h) (Perry, 2015) Small population size (11 students). No control group. Weak evidence. (Poole et al., 2014) Short study period (3 weeks) (Shi et al., 2014) Short study period (two weeks); Small population size (20 students)	(Morschheuser et al., 2014)	Short study period (20 days)
(Poole et al., 2014) Short study period (3 weeks) (Shi et al., 2014) Short study period (two weeks); Small population size (20 students)	(Pedro et al., 2015b)	
(Shi et al., 2014) Short study period (two weeks); Small population size (20 students)	(Perry, 2015)	Small population size (11 students). No control group. Weak evidence.
	(Poole et al., 2014)	Short study period (3 weeks)
(Sillaots, 2014) Small population size (28 students)	(Shi et al., 2014)	Short study period (two weeks); Small population size (20 students)
	(Sillaots, 2014)	Small population size (28 students)
Simoes et al., 2015) Small population size (26 students); Single survey based on homework activity with no control group.	(Simoes et al., 2015)	
(Su & Cheng, 2015) Unclear measurements for the motivational outcomes	(Su & Cheng, 2015)	Unclear measurements for the motivational outcomes
Tu et al., 2015) Results do not support the predictive relationship between gaming personality and game dynamics	(Tu et al., 2015)	
(Utomo & Santoso, 2015) Small population size (31 students); Short period of study (1 week)	(Utomo & Santoso, 2015)	Small population size (31 students); Short period of study (1 week)

outcomes. Motivation is associated with a number of learning related concepts such as engagement, effort, goals, focus of attention, self-efficacy, confidence, achievement, interest, etc. Improving our understanding of motivational aspects of gamification will enable us to predict its effect on the related concepts. In addition, it will help improve the gamification design, in particular, how to design an appropriate gamified experience that strengthens the motivation of a given population of learners and leads to desirable learning outcomes.

Theoretical perspective

Gamification is growing as an area of both practice and research. The majority of the studies reviewed in the previous sections lack a theoretical underpinning that can help

understand the researchers' motivation and the justification for how their gamification approach is supported by a theoretical framework. For completeness of the review, in this section we outline theoretical works underpinning the use of gamification in education, published within the review period. Overall, the bulk of theoretical research addressing gamification maintains that focusing on points and rewards rather than on play and intrinsic engagement cannot always meet the goal of desired behavior change by catering to the intrinsic values of learners (Hansch et al., 2015; Songer & Miyata, 2014; Tomaselli et al., 2015). This suggests a user-centered approach in the design of gamified systems, characterized by a focus on the needs and desires of learners. A new line of research is taking steps towards developing a theory of educational gamification by combining motivational and learning theories aimed at linking gamification to practical education (Landers, 2015; Landers, Bauer, Callan, & Armstrong, 2015) or by developing a framework for integrating gamification with pedagogy (Tulloch, 2014) or psychology of games (Lieberoth, 2015).

Tulloch (2014) maintains that gamification is a product of an overlooked history of pedagogic refinement, a history of training that is effective, but largely ignored, namely the process of games teaching players how to play. He challenges the evolving concept of gamification, conceptualizing it not as a simple set of techniques and mechanics, but as a pedagogic heritage and an alternative framework for training and shaping participant behavior that has at its core the concepts of entertainment and engagement. Yet, Biro (2014) considers gamification as a new educational theory, alongside of behaviorism, cognitivism, constructivism and connectivism.

Songer and Miyata (2014) propose to deviate from using simple game elements often found in gamification approaches and move to a "gameful" experience that fosters intrinsic motivation of players. The authors address the issue of gamifying educational contexts with discussions about gamer motivations, the relationship between games and play, and designs for optimal learning within games. Based on the theoretical foundations of behavioral psychology, anthropology and game studies, the authors propose a model for the design and evaluation of playful experiences in learning environments inspired by game design.

With related concerns, (Tomaselli et al., 2015) attempts to analyze the most engaging factors for gamers in the context of gamification by questioning the relevance of some of the most used gamification strategies like attributing points and badges or simple reputation elements to users. The authors explore how engagement is associated with a variety of types of contemporary digital games. The results show that although there is support for the importance of competition against peers (contrary to the current prevailing understanding), the challenge of overcoming the game's obstacles and mastering them is what matters the most to players, regardless of the type of the game. The takeaway message is that the gamified system designers should not be so concerned with rankings and online comparisons to encourage users to compete against each other, but with their use as a personal reference, creating challenging environments and guidance for users to achieve their mastery interests.

Landers (2015) advocates that no single theory is able to explain gamification. Accordingly, he presents a set of theories organized in two categories, motivational and learning theories that are most likely to explain the effects of gamification when it is implemented as an instructional intervention. Among learning theories, Landers

identifies two major frameworks to describe the learning outcomes of gamification: the theory of gamified instructional design and classic conditioning theories of learning. He also identifies three major types of motivational theories: expectancy-based theories, goal-setting theory, and the self-determination theory.

The theory of gamified learning proposed by (Landers et al., 2015) provides two specific causal pathways by which gamification can affect learning and a framework for testing these pathways. Their theory identifies two specific processes by which gamification can affect learning. In both processes the gamification is aimed at affecting learning-related behavior. In the first one, this behavior moderates the relationship between instructional quality and learning. In the second, this behavior mediates the relationship between game elements and learning. Critically, one or both of these processes may be involved in any particular gamification effort. For gamification to be effective, it must successfully alter an intermediary learner behavior or learner attitude. That behavior or attitude must then itself cause changes in learning directly or must strengthen the effectiveness of existing instruction.

In their explorative study, Hansch et al. (2015) examine the motivational potential of gamification in online learning. Through reviewing ten platforms and an in-depth analysis, they explore how the motivational potential of gamification mechanics and the social and interactive elements in online learning can be effectively combined to build a community of engaged learners. The authors conclude that the starting point in gamifying online education should be learners' needs, motivations and goals, rather than a platform-centric approach that strives to use technical features to hit some pre-defined performance metrics.

According to Lieberoth (2015), it might not be the game itself that stimulates individuals, but rather the packaging: the fact that an activity resembles a game. The simple framing of an activity as a "game" can potentially alter an individual's behavior. To demonstrate this insight Lieberoth designed an experiment focusing on the psychological effects of framing tasks as games versus including game mechanics. The outcomes indicate that engagement and enjoyment increased significantly due to the psychological effects of framing the task as a game. Furthermore, no actual increased interest or enjoyment was measured by adding actual game mechanics to the task, when it was already framed as a game. This study reveals an interesting psychological perspective of gamification in educational environments: merely making an activity seem like a game impacts learners' engagement.

In addition to the gamification works with theoretical, conceptual or methodological orientation, five literature reviews (Borges, Durelli, Reis, & Isotani, 2014; Caponetto et al., 2014; Dicheva & Dichev, 2015; Faiella & Ricciardi, 2015; Gerber, 2014) have been published over the last two years. While these reviews synthesize the empirical research on gamification in education, neither of them provides a critical analysis of the strengths and weaknesses of the research findings of the reviewed studies. The present review addresses this gap by evaluating analytically the validity of the reported results.

The research on gamification frameworks, platforms, and toolsets that help making the design and development of gamification applications easier, faster, and cheaper has also been showing progress in the last few years. Since the current research on gamification specific frameworks is not explicitly driven by educational objectives, we refer the interested readers to a corresponding literature review on this topic (Mora et al., 2015).

While the reviewed theoretical studies are touching interesting points, the covered topics are insufficient for complete understanding of the motivational mechanisms of gamification in educational context. Without a theoretical framework backing the design of the studies and the interpretation of their results, it is problematic to select an appropriate gamification structure or to differentiate which of the employed game mechanisms and principles were essential for arriving at successful outcomes. Hence, there is a need of theoretical and empirical studies capable of mutually advancing each other. This will allow bridging the identified gaps in order to understand how gamification in education works, when it works best, and its limits and key strategies.

Conclusion

Gamification in education is an approach for encouraging learners' motivation and engagement by incorporating game design principles in the learning environment. The importance of sustaining students' motivation has been a long-standing challenge to education. This explains the significant attention that gamification has gained in educational context - its potential to motivate students. However, the process of integrating game design principles within varying educational experiences appears challenging and there are currently no practical guidelines for how to do so in a coherent and efficient manner. The discussion in the present review has been structured based on the combinations of the game elements used, the gamified subjects, the type of learning activities, and the identified goals, ending with a thorough discussion on the reliability and validity of the reported outcomes. The review confirmed that the research on gamification is very diverse with respect to the focus of the studies, the reported outcomes and methodological approaches. It also indicates that the research focus at present is mainly on empirical studies with less attention to the theoretical considerations. Moreover, the majority of the studies target college students. A number of gamification approaches, driven by specific objectives, have been applied to support learning and related activities in a variety of educational contexts. Studies on how distinct categories of learners are affected by gamification, what to measure as an outcome, and how to add a gamified layer to a core activity are also emerging. Despite the fact that gamification in education is still growing phenomenon, the reviewed studies indicate that (i) The practice of gamifying learning has outpaced researchers' understanding of its mechanisms and methods, (ii) Insufficient high-quality evidence exists to support the long-term benefits of gamification in educational context, and (iii) The understanding of how to gamify an activity depending on the specifics of the educational context is still limited.

We have identified a growing number of studies reporting empirical evidences for the effectiveness of gamification in educational context. At the same time, it is noticeable that a growing body of reported results is backed by inconclusive and insufficient evidence for making valid claims about the efficacy of gamification in education. Possible reasons for this are from one side the hype to publish on gamification and from another, the addressing of an overly broad research question based on limited supportive evidence. Whether gamification motivates students, improves learning or increases participation, are too broad questions. Instead, the

focus should be narrowed to questions of the type: whether game design elements G are effective for learners of type L participating in activity of type A. All these indicate a need of a systematic program of experimental studies mapping game elements to the learning and motivational specifics of individual (groups of) learners. Another grey area that deserves attention is how to avoid gamification scenarios that can harm learning.

Gamification is a psychologically driven approach targeting motivation—the desire and willingness to do something. From technical perspective, it is a motivational design problem. While the majority of the reviewed studies do analyze specific educational effects of gamification (on learning, attainment, participation), their focus is often aside from motivation. When motivation is targeted, it is typically examined through observable indicators, such as grades, attendance, etc. that are not always directly linked to it. As a result, the educational benefits of gamification in terms of increasing student motivation or linking this motivation to learning outcomes are still not well understood.

While the effort to understand the effects of gamification on learning is expanding, there is a need for exploring the effect of game design elements in its broad sense including game mechanics and game dynamics and across learning contexts. The observed emphasis on points, badges, and leaderboards is too narrow to address the relevant motivational factors. It is also crucial to understand the target population of a gamified system in order to gamify a learning activity successfully. Specifically, the unique needs and preferences of each group of learners, along with the particular learning objectives relevant to that group must inform the choice of game elements.

A comparison of the results of this survey with the previous ones, which marked the climb to the inflated expectation, indicates a trend of decline of the expectations. The rise and fall of expectations for applying gamification in educational contexts is nothing out of the ordinary. Most emerging technologies and the accompanying research go through an initial period of hype as described by the Gartner's Hype Cycle, before evolving for a second period of measured popularity, in which it attains maturity and meets the expectations (Naik, 2015). There are several assumptions underlying the usefulness of gamification in educational context, such as gamification is motivating, gamification is engaging, gamification can improve attendance and participation. However, research remains inconclusive on these assumptions. Educational contexts in which gamification may be particularly useful have not been confirmed yet. This does not mean though that gamification cannot be used with success in a learning context. It simply means that the educational benefits of gamification have not been scientifically confirmed yet. Only continued theoretical and rigorous systematic empirical work in varying gamification settings and across contexts will enable us to establish a practical, comprehensive, and methodical understanding of the benefits of applying gamification in educational contexts.

Endnotes

¹This terminology has been popularized through the book "For the win: How game thinking can revolutionize your business" by Werbach and Hunter and a series of Coursera's MOOCs.

Appendix 1

Table 13 Characterization of Studies Based on Game Mechanics, Learning Context, Measurements and Outcomes

Paper	Game elements Subject Gamified activity Sample size	Subject	Gamified activity	Sample size	Duration	Data collection	Outcome
(Amriani et al., 2014)	Points, badges, leaderboard, status, levels, unlockable content, customization	Web development, Game development	Learning interactions in Virtual Realty	93 learners	12 weeks for each module	Exploratory data analysis: log data from VR + Questionnaire	Improved participation + engagement
(Anderson et al., 2015)	Badges, points, competition, leaderboard	Intro to Data Science	Assessment	37 students randomized into gamified and control groups	1 class period	Performance on test problems + comments on the system	Improved performance
(Attali & Arieli-Attali, 2015)	Points	Mathematics	Assessment	1218 Adults + 693 grades 6–8 students randomized into gamified and control groups	1 test session	Recorded accuracy and speed in test performance	The effect of points on performance (accuracy + speed)
(Auvinen et al., 2015)	Badges, heatmap	Data Structures and Algorithms	Online exercises	Gamified group ($N = 254$), heatmap group ($N = 109$)	2 semesters	System logs + numerical feedback	Differences in reacting to gamification feedback
(Barata et al., 2014)	XPs, badges, levels, leaderboard, challenges	Multimedia Content Production	Course activities	35 enrolled students (12 + 23)	1 semester	Student performance + accumulated XP	Effects of gamification on different student types
(Barrio et al., 2015)	Points	Unspecified	Learning activities using SRS	131 (62 + 69) students	4 sessions (90 min each)	Surveys + posttest	Improved motivation, attention, learning performance
(Bernik et al., 2015)	Points, badges, leaderboard, progress	3D Modeling	Learning activities in a course module	28 + (27 control group) students	2 weeks	Pre-post-test + Questionnaire	Improved performance
(Bonde et al., 2014)	Lab simulation + narrative + fictional characters	Biology	Lab activities	91 students (gamified + control groups)	40 days	Learning outcomes	Increase in learning outcomes and motivation
(Boskic & Hu, 2015)	Choice, role playing, feedback	Education course	Course assignments	25 students	13 weeks	Interview (an instructor + a student)	Increased engagement understanding
(Boticki et al., 2015)	Badges	N/A	Mobile app- based learning activities	305 primary school students	1 school year	Systems logs, observations, feedback	Motivated a specific category of students

Table 13 Characterization of Studies Based on Game Mechanics, Learning Context, Measurements and Outcomes (Continued)

(Chang & Wei 2015)	40 gamification mechanisms	Unspecified	Coursera edX Udacity	35 frequent users +	Unspecified	Oriectionnaire	Gamification mechanics
			based activities	5,020 learners			identified as engaging
(Christy & Fox, 2014)	Leaderboards	Mathematics	Assessment	80 women randomized into 3 groups maledominated leaderboard, female-dominated leaderboard and control groups	1 experimental test	Tests score + self-assessment of math and academic skills	Leaderboards create stereotype threat
(Codish & Ravid, 2014)	Points, badges, leaderboards, narrative	Software analysis and design	In-class and course project activities	Two course (n = 102 and n = 58)	Two consecutive courses	2 Quasi-experiments + 4 surveys	Personal differences in perceived game mechanics
(Codish & Ravid, 2015)	Points, badges, riddles	Software analysis and design	LMS - interactions	38 undergraduates students	One semester	Questionnaire + log analysis	Evidence for GBP to predict playfulness
(Davis & Klein, 2015)	Badges	N/A	Afterschool learning	High school students		Pre-, during and post- focus groups, usability tests	Evidence for students' perceptions of badges (contextual)
(Hakulinen et al., 2015)	Badges	Data Structures and Algorithms	Homework exercises	281 randomized - control and treatment groups	1 semester	System logs and students' attitudes from feedback	Improved motivation
(Hanus & Fox, 2015)	Badges, virtual coins, leaderboard, pseudonyms	Communication course	In-class and out-of- class activities	80 students divided in gamified and non-gamified courses	One semester (16 W)	Four surveys + Exam scores	Satisfaction, empowerment, academic performance
(Hasegawa et al., 2015)	Points, trials, character, ranking, progress	English	Vocabulary learning	27 undergraduate/ graduate	Unspecified	Survey	Motivated continuous learning
(Herbert et al., 2014)	Points, badges, leaderboard, status, levels, unlockable content, customization	Web development, Game development	Learning interactions in Virtual Realty	93 learners	12 weeks for each module	Log data from VR + questionnaire + exploratory data analysis	Relation between learner gamification typology and learner interactive behavior
(Hew et al., 2016)	Points, badges, leaderboard	Designing Questionnaire	Project activities	22 (11 + 11) 43 (21 + 22) students	2 studies 3 + 19 days	Log analysis +test scores	The effect of gamification on engagement

Table 13 Characterization of Studies Based on Game Mechanics, Learning Context, Measurements and Outcomes (Continued)

					\ i \ i \ i \ i \ i \ i \ i \ i \ i \ i		
(Holman et al., 2013)	Badges, leveling, leaderboard, autonomy, grade predictor	Information studies, Political theory	Class activities	292 + 231 students	1 semester	Analyzing system logs	The impact of grade predictor on planning the work over the course
(Ibanez et al., 2014)	Badges, leaderboards	Operating systems	Course activities	22 undergraduate students	1 semester	Log files + pre-post-tests	The effect of gamification on engagement and learning.
(Jang et al., 2015)	Levels, point, life points, avatars, feedbacks, time pressure	Photoshop	Learning activities	114 students randomized into gamified and control groups	Unspecified	Pre- and post-tests	Improved learning performance
Knutas et al. (2014a)	N/A	Software engineering	Collaborative teamwork	17 students	5 day period	Classroom interactions + interviews + survey	Profiling based on gamification preference
Knutas et al. (2014b)	Points, badges, up-vote or down-vote	Introductory programming	Collaborative learning	249 students	1 semester	System logs + survey	Positive effect on student collaboration
(Krause et al., 2015)	Achievements, points, leaderboards, avatars	Python for statistical analysis	Online course activities	71 control, 67 game condition	1 semester	Tests and quizzes performance	Improved retention period, and learning performance
(Lambruschini & Pizarro, 2015)	Points, badges, leaderboard	Process management	Course participation	94 students	1 semester	Single survey - 13% response	Increased communication, participation, punctuality
(Landers & Landers, 2015)	Leaderboards	Industrial/organizational psychology	Course project activities	109 students randomized into gamified and control groups	1 semester	Recorded project activities	Improved time on task
(Laskowski & Badurowicz, 2014)	Points, leaderboards, badges	Service-Oriented Architectures	All course activities	62 master students randomized into gamified and control groups	1 semester	Attendance + performance on tests, assignments, projects	Improved engagement and performance
(Latulipe et al., 2015)	Stamps, tokens, leaderboards	CS1	In-class course activities	92 + 65 students	2 semesters	Surveys – twice per semester	Encouraged harder work and engagement
(Leach et al., 2014)	Badges, points	Technology and society	Online class activities	50 students randomized into gamified and control groups	85 days	Recorded student online activities	Increased online activities and learning performance

Table 13 Characterization of Studies Based on Game Mechanics, Learning Context, Measurements and Outcomes (Continued)

	TOTAL STREET, CHARACTERISTICS OF STATES OF CALL STREET, THE STATES OF CONTRACT AND ACCOUNTS AND CONTRACTORS OF	ימודול זעוללו ומודולט, בלמודווו	וש כטוונכאני ואוכשפעוכוו		ינווימכט/		
(Lehtonen et al., 2015)	Points, badges, leaderboard	Open online learning - Java	Online Java exercises	13 577 users	9 years	Log analysis	Increased usage of the open learning environment
(Long & Aleven, 2014)	Stars/badges	Middle school mathematics	Problem solving with re-practicing	Problem solving with 267 students randomized re-practicing into gamified and control groups	5 class periods	5 class periods Pre- and post-tests + recorded data	Does not improve learning
(Mekler et al., 2015)	Points, levels and leaderboard	Unspecified	Image annotation	273 students –130 autonomy, 143 control oriented	22 min session	22 min session Tracking the amount of tags generated	Increased competence need (negative) and performance (positive)
(Morschheuser et al., 2014)	Points, badges personas	N/A	Interaction with PLE	70 participants	20 days	Questionnaire	Increased intention to use the PLE
(Nevin et al., 2014)	Badges, Ievels, feedback, Ieaderboard, voluntary participation	Medical education	Interactions with a learning environment	152 participants + focus group (n = 17)	1 academic year	Log data from the platform and from the focus group	Increased knowledge retention, reduce attrition
(Paiva et al., 2015)	Points, badges	High school mathematics	Interactions with ITS	100 students	Unspecified	Log data (100 students)	Correlation between XPs, badges and learning - positive
(Pedro et al., 2015a)	User-generated badges	N/A	Badges creation, attribution	Learning platform users	1 academic year	System logs	Participation badge creation and attribution
(Pedro et al., 2015b)	Points, badges, levels, feedback, ranking	Middle school mathematics	Interactions with VLE	7 girls and 9 boys divided randomly in two groups	ر م	The score with the gamified system	Improved performance, reduced undesirable behaviors in VLE
(Perry, 2015)	Points, badges	French language	Interactions with Language learning system	11 students	1 semester	Pre- and post- questionnaires, focus groups, recordings of the sessions	Increased playfulness and engagement in learning
(Pettit et al., 2015)	Challenge, progress, competition, status, achievement, prizes, chance, surprise, anticipation, humor	medical microbiology	ARS interactions	91 students (86%)	2/3 semester	Questionnaire	Increased engagement and learning

Table 13 Characterization of Studies Based on Game Mechanics, Learning Context, Measurements and Outcomes (Continued)

(Poole et al., 2014)	Points, leaderboards	Accounting	Class activity	55 students in gamified and traditional classes	3 weeks	Test score, survey	Increased engagement, foster learning
(Shi et al., 2014)	Leaderboard, progress, feedback, social status,	Unspecified	Interactions via a learning environment	20 students	2 weeks	Survey + system logs	Increased intrinsic motivation
(Sillaots, 2014)	Goals, avatar, XPs, feedback scoreboard, levels, luck, competition	Research methods	In-class activities	28 IT students	1 semester	questionnaire	Mixed acceptance of game elements
(Sillaots, 2015)	Goals, avatar, XPs, feedback scoreboard, levels, luck, competition, game vocabulary	Computer games	Game design activities	23 students	1 semester	Observation, group interview, questionnaire	Perception of game elements incorporated in a course
(Simoes et al., 2015)	Badges, points, leaderboard	Unspecified	Homework in Schooooools	26 primary school students	1 homework	Psychometric survey	Increased disposition to the experience flow
(Smith et al., 2014)	Merit points, badges, voting	Engineering/IT	Online discussions	1586 students	6 weeks	Participation data and survey	Improved participation and quality of online discussion
(Su & Cheng, 2015)	Badges, leaderboard, missions	Botanics	Field activities	3 fourth-grade classes (1 + 2)	3 weeks	Questionnaire + pre- and post-test	Increased motivation and learning
(Tu et al., 2015)	Badges	Unspecified	Test activities	71 master students	14 weeks	Survey	Relation between gaming personality and game dynamics
(Tvarozek & Brza, 2014) Interactive badges	Interactive badges	Introductory programming + Advanced algorithms	Programming exercises	186 students	1 semester for each course	Surveys + badge interactions	Relation between badge interaction and engagement
(Utomo & Santoso, 2015) Badges, progress bar	Badges, progress bar	Student-centered learning	Online learning activities	31 students	1 week	Focus group + feedback forms	Fostered learning activities

Appendix 2

Table 14 Goals of the Studies

Table 14 Goals of the Studie	?\$
Paper	Stated Goals
(Amriani et al., 2014)	The effect on ^a learning participation
(Anderson et al., 2015)	The effect of gamifying Learn2Mine system on student performance
(Attali & Arieli-Attali, 2015)	The effect on test performance
(Auvinen et al., 2015)	To compare analytically visualizations vs. gamification on student's performance with different goal orientations
(Barata et al., 2014)	To identify student types based on how they experience gamified courses
(Barrio et al., 2015)	The effect of gamifying Student Response Systems
(Bernik et al., 2015)	The effect on student achievement
(Bonde et al., 2014)	The learning effectiveness of gamified simulation
(Boskic & Hu, 2015)	To study the effect of transforming a traditional course into a role-playing game
(Boticki et al., 2015)	To study how badge score predicts the student exam success
(Chang & Wei, 2015)	To study the array of game mechanics used in MOOCs and their engaging effect
(Christy & Fox, 2014)	The impact of leaderboards on performance re: the stereotype threat and social comparison
(Codish & Ravid, 2014)	To study how learners perceive playfulness
(Codish & Ravid, 2015)	To study the effectiveness of gamification behavior patterns as a measure of playfulness
(Davis & Klein, 2015)	To study high school students' perceptions of badges
(Hakulinen et al., 2015)	The effect of badges on student behavior
(Hanus & Fox, 2015)	The effectiveness of gamification longitudinally
(Hasegawa et al., 2015)	The effect on motivating learners reluctant to continue learning
(Herbert et al., 2014)	To understand the variation in motivation between learners with different gamification typologies
(Hew et al., 2016)	The effect on Asian students performance
(Holman et al., 2013)	The effect on coursework planning
(Ibanez et al., 2014)	The effect on engagement and learning performance
(Jang et al., 2015)	The effect on student learning
Knutas et al. (2014a)	To identify learners' gamification preferences
Knutas et al. (2014b)	The effect on student collaboration in online learning
(Krause et al., 2015)	The effect on retention and learning
(Lambruschini & Pizarro, 2015)	The effect on LMS
(Landers & Landers, 2015)	To understand the causal effect of gamifying a course project with leaderboards
(Laskowski & Badurowicz, 2014)	The effect on engagement and quality of learning
(Latulipe et al., 2015)	The effect on improving student engagement
(Leach et al., 2014)	The effect on the results of graded assignments
(Lehtonen et al., 2015)	The effect on the usage of the learning environment
(Long & Aleven, 2014)	The effect on shared student/system control in a linear equation tutor
(Mekler et al., 2015)	To study whether points, leaderboards and levels increase performance, competence need, satisfaction and intrinsic motivation
(Morschheuser et al., 2014)	To study how to enhance user's engagement with PLE
(Nevin et al., 2014)	To understand how gamification affects acceptance and use of medical knowledge software
(Paiva et al., 2015)	The effect on students' learning in the MeuTutor
(Pedro et al., 2015a)	The effect of badges co-creation on engagement and motivation
(Pedro et al., 2015b)	The effect on reducing undesirable behaviors and increasing performance in VLE

Table 14 Goals of the Studies (Continued)

(Perry, 2015)	The effect on learning French as a second language
(Pettit et al., 2015)	The effect on audience response systems
(Poole et al., 2014)	The effect on engaging and learning generation Y students
(Shi et al., 2014)	To identify the factors increasing intrinsic motivation in social e-learning environments
(Sillaots, 2014)	To study the learners' acceptance of game mechanics
(Sillaots, 2015)	To study the learners' perception of game elements
(Simoes et al., 2015)	To study the impact of gamification on students' engagement and how to measure that impact
(Smith et al., 2014)	The effect on participation and quality of online discussions
(Su & Cheng, 2015)	The effect on mobile learning systems
(Tu et al., 2015)	To study the predictive effect of gaming personality on their game dynamic preferences
(Tvarozek & Brza, 2014)	To study the creation of a badge as a tool for measuring students' interest
(Utomo & Santoso, 2015)	The effect on students' motivation

^aNote: The phrase "The effect on" should be interpreted as "The effect of gamification on"

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant No. HRD 1623236 "Targeted Infusion Project: Increasing Student Motivation and Engagement in STEM Courses through Gamification".

Authors' contributions

Both authors contributed equally towards this article. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Received: 2 November 2016 Accepted: 2 November 2016 Published online: 20 February 2017

References

Amriani, A., Aji, A., Utomo, A. Y., Wahidah, F., & Junus, K. (2014). Gamified E-learning model based on community of inquiry. In 2014 IEEE International Conference on Advanced Computer Science and Information Systems, Jakarta, 474– 480. doi: 10.1109/ICACSIS.2014.7065830.

Anderson, P. E., Nash, T., & McCauley, R. (2015). Facilitating programming success in data science courses through gamified scaffolding and Learn2Mine, ITICSE '15, July 04–08, 2015, Vilnius, Lithuania.

Attali, Y., & Arieli-Attali, M. (2015). Gamification in assessment: do points affect test performance? *Computers & Education*, 83, 57–63.

Auvinen, T., Hakulinen, L., & Malmi, L. (2015). Increasing Students' Awareness of their Behavior in Online Learning Environments with Visualizations and Achievement Badges, *IEEE Transactions on Learning Technologies*, 8(3), 261–273, doi: 10.1109/TLT.2015.2441718.

Barata, G., Gama, S., Jorge, J., & Gonçalves, D. (2014). Identifying student types in a gamified learning experience. International Journal of Game-Based Learning (IJGBL), 4(4), 19–36.

Barrio, C.M., Organero M.M., & Soriano, J. S. (2015). Can Gamification Improve the Benefits of Student Response Systems in Learning? An Experimental Study. *IEEE Transactions on Emerging Topics in Computing, PP*(99). doi: 10.1109/TETC. 2015.2497459.

Bernik, A., Bubaš, G., & Radošević, D. (2015). A pilot study of the influence of gamification on the effectiveness of an e-Learning Course. In *26th Central European Conference on Information and Intelligent Systems (CECIIS 2015)*. (pp. 73–79). Varazdin: Faculty of Organization and Informatics.

Biro, G. I. (2014). Didactics 2.0: a pedagogical analysis of gamification theory from a comparative perspective with special view to the components of learning. *Procedia - Social and Behavioral Sciences, 141,* 148–151.

Bonde, M. T., Makransky, G., Wandall, J., Larsen, M. V., Morsing, M., Jarmer, H., & Sommer, M. O. (2014). Improving biotech education through gamified laboratory simulations. *Nature Biotechnology*, 32(7), 694–697.

Borges, S. S., Durelli, V. H. S., Reis, H. M., & Isotani, S. (2014). A systematic mapping on gamification applied to education. In ACM SAC'14 Conference, Gyeongju, South Korea (pp. 216–222).

Boskic, N., & Hu, S. (2015). Gamification in higher education: how we changed roles. *European Conference on Games Based Learning*, (pp. 741–748). Reading, UK: Academic Conferences and Publishing International Limited.

Boticki, I., Baksa, J., Seow, P., & Looi, C. K. (2015). Usage of a mobile social learning platform with virtual badges in a primary school. *Computers & Education*, 86, 120–136.

Broer, J. (2014). Gamification and the trough of disillusionment. In A. Butz, M. Koch, & J. Schlichter (Eds.), Mensch & Computer 2014 - Workshopband (pp. 389–395). Berlin: De Gruyter Oldenbourg.

- Caponetto, I., Earp, J., & Ott, M. (2014). Gamification and education: a literature review. In 8th European Conference on Games Based Learning (pp. 50–57). Germany: ECGBL. ISBN 978-1-910309-55-1.
- Chang, J.W., & Wei, H.Y. (2015). Exploring Engaging Gamification Mechanics in Massive Online Open Courses, Educational Technology & Society, 12/2015 (in print).
- Christy, K. R., & Fox, J. (2014). Leaderboards in a virtual classroom: a test of stereotype threat and social comparison explanations for women's math performance. *Computers & Education*, 78, 66–77.
- Codish, D., & Ravid, G. (2014). Academic course gamification: the art of perceived playfulness. *Interdisciplinary Journal of E-Learning and Learning Objects*, 10, 131–151.
- Codish, D., & Ravid, G. (2015). Detecting playfulness in educational gamification through behavior patterns. *IBM Journal of Research and Development*, 59(6), 1–14. doi:10.1147/JRD.2015.2459651
- Davis, K., & Klein, E. (2015). Investigating high school students' perceptions of digital badges in afterschool learning. In ACM Conference on Human Factors in Computing Systems (CHI '15), (pp. 4043–4046). New York, NY: ACM
- Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: defining gamification. In 15th International Academic MindTrek Conference: Envisioning Future Media Environments (pp. 9–15). New York, NY: ACM.
- Dichev, C., Dicheva, D., Angelova, G., & Agre, G. (2014). From gamification to gameful design and gameful experience in learning. *Journal of Cybernetics and Information Technologies*, 14(4), 80–100. doi:10.1515/cait-2014-0007
- Dicheva, D., & Dichev, C. (2015). Gamification in education: where are we in 2015? World Conference on E-Learning (ELEARN 2015), Kona, Hawaii, October 19–22 (pp. 1276–1284). Chesapeake: AACE.
- Dicheva, D., Dichev, C., Agre, G., & Angelova, G. (2015). Gamification in education: a systematic mapping study. Educational Technology & Society, 18(3), 75–88.
- Enders, B., & Kapp, K. (2013). Gamification, Games, and Learning: What Managers and Practitioners Need to Know, Hot Topics, *The eLearning Guild Research*, 1–7.
- Faiella, F., & Ricciardi, M. (2015). Gamification and learning: a review of issues and research. *Journal of e-Learning and Knowledge Society*, 11(3), 13–21.
- Gartner. (2013). Hype cycle for emerging technologies, 2013. Retrieved from http://www.gartner.com/document/2571624. Accessed 15 Apr 2016.
- Gerber, H. (2014). Problems and Possibilities of Gamifying Learning: A Conceptual Review. *Internet Learning Journal*, 3 (2), Article 5. Retrieved from http://digitalcommons.apus.edu/internetlearning/vol3/iss2/5. Accessed 16 Jan 2017.
- Gooch, D., Vasalou, A., Benton, L., & Khaled, R. (2016). *Using gamification to motivate students with dyslexia*. CHI 2016, May 7–12 2016. San Jose.
- Hakulinen, L., Auvinen, T., & Korhonen, A. (2015). The effect of achievement badges on students' behavior: an empirical study in a university-level computer science course. *International Journal of Emerging Technologies in Learning (IJET)*, 10(1): 18–29
- Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does gamification work? A literature review of empirical studies on gamification. In 47th Hawaii International Conference on System Sciences, Hawaii, USA (pp. 3025–3034). doi:10.1109/HICSS.2014.377.
- Hansch, A., Newman, C., & Schildhauer, T. (2015). Fostering Engagement with Gamification: Review of Current Practices on Online Learning Platforms. (November 23, 2015). HIIG Discussion Paper Series No. 2015–04. Retrieved from http://dx.doi.org/10.2139/ssrn.2694736.
- 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
- Hasegawa, T., Koshino, M., & Ban, H. (2015). An English Vocabulary Learning Support System for the Learner's Sustainable Motivation. Springer Plus: Innovative Cloud Application in Computer Intelligence, 4 (99). doi: 10.1186/s40064-015-0792-2.
- Herbert, B., Charles, D., Moore, A., & Charles, T. (2014). An investigation of gamification typologies for enhancing learner motivation. In *International Conference on Interactive Technologies and Games, UK* (pp. 71–78).
- Hew, K. F., Huang, B., Chu, K. W. S., & Chiu, D. K. W. (2016). Engaging Asian students through game mechanics: findings from two experiment studies. *Computers & Education*, 92–93, 221–236.
- Holman, C., Aguilar, S., & Fishman, B. (2013). GradeCraft: what can we learn from a game-inspired learning management system? *Third International Conference on Learning Analytics and Knowledge, 2013,* (pp. 260–264). New York, NY: ACM.
- Holman, C., Aguilar, S. J., Levick, A., Stern, J., Plummer, B., & Fishman, B. (2015). Planning for success: how students use a grade prediction tool to win their classes. In *2015 Third International Conference on Learning Analytics and Knowledge*, (pp. 260–264). New York, NY: ACM.
- Ibanez, M., Di Serio, A., & Delgado-Kloos, C. (2014). Gamification for engaging computer science students in learning activities: a case study. IEEE Transactions on Learning Technologies, 7(3), 291–301.
- Jang, J., Park, J., & Yi, M. Y. (2015). Gamification of online learning. In 17th International Conference on Artificial Intelligence in Education (AIED) (pp. 646–649). Switzerland: Springer International Publishing.
- Knutas, A., Ikonen, J., Maggiorini, D., Ripomonti, L., & Porras, J. (2014a). Creating software engineering student interaction profiles for discovering gamification approaches to improve collaboration. In 15th International Conference on Computer Systems and Technologies (CompSysTech'14), (pp. 378–385). New York, NY: ACM.
- Knutas, A., Ikonen J., Nikula, U., & Porras, J. (2014). Increasing Collaborative Communications in a Programming Course with Gamification: A Case Study. 15th Int. Conference on Computer Systems and Technologies (CompSysTech'14), (pp. 370–377). New York, NY: ACM
- Koivisto, J., & Hamari, J. (2014). Demographic differences in perceived benefits from gamification. *Computers in Human Behavior*, 35, 179–188.
- Krause, M., Mogalle, M., Pohl, H., & Williams, J. J. (2015). A playful game changer: fostering student retention in online education with social gamification. In L@S'15 Proc. of Learning@ Scale Conference (pp. 95–102). Vancouver: ACM Press.

- Lambruschini, B. B., & Pizarro, W. G. (2015). Tech Gamification in University Engineering Education. In 10th International Conference on Computer Science & Education (ICCSE 2015) (pp. 295–299) IEEE Conference Publications, doi: 10.1109/ ICCSE.2015.7250259.
- Landers, R. N. (2015). Developing a theory of gamified learning: linking serious games and gamification of learning. Simulation & Gaming, 45, 752–768.
- Landers, R. N., Bauer, K. N., Callan, R. C., & Armstrong, M. B. (2015). Psychological theory and the gamification of learning. In T. Reiners & L. Wood (Eds.), Gamification in education and business (pp. 165–186). Cham: Springer.
- Landers, R. N., & Landers, A. K. (2015). An empirical test of the theory of gamified learning: the effect of leaderboards on time-on-task and academic performance. *Simulation & Gaming*, 45, 769–785.
- Laskowski, M., & Badurowicz, M. (2014). Gamification in Higher Education: A Case Study. In MakeLearn 2014: Human Capital without Borders: Knowledge and Learning for Quality of Life. Management, Knowledge and Learning International Conference, 971–975. (pp. 971–975). Portorož: ToKnowPress.
- Latulipe, C., Long, N. B., & Seminario, C. E. (2015). Structuring flipped classes with lightweight teams and gamification. SIGCSE, 2015, Proceedings of the 46th ACM Technical Symposium on Computer Science Education (pp. 392–397). New York. NY: ACM
- Leach, D., Laur, B., Code, J., Bebbington, T., & Broome, D. (2014). *Gamification for online engagement in higher education:* a randomized controlled trial (pp. 151–159). Madison: Games Learning + Society Conference.
- Lehtonen, T., Aho, T., Isohanni, E., & Mikkonen, T. (2015). On the role of gamification and localization in an open online learning environment: Javala experiences. In 15th Koli Calling Conference on Computing Education Research (pp. 50–59). New York, NY: ACM.
- Lieberoth, A. (2015). Shallow gamification psychological effects of framing an activity as a game. *Games and Culture*, 10(3), 249–268.
- Linehan, C., Kirman, B., Lawson, S., & Chan, G. (2011). Practical, appropriate, empirically-validated guidelines for designing educational games. In ACM Annual Conference on Human Factors in Computing Systems, May 7–12 (pp. 1979–1988). Canada: Vancouver.
- Long, Y., & Aleven, V. (2014). Gamification of joint student/system control over problem selection in a linear equation tutor. In S. Trausan-Matu, K. E. Boyer, M. Crosby, & K. Panourgia (Eds.), 12th International Conference on Intelligent Tutoring Systems, ITS 2014 (pp. 378–387). Honolulu: Springer International Publishing.
- Mahnič, N. (2014). Gamification of politics: start a new game! Teorija in Praksa, 51(1), 143-161.
- McGonigal, J. (2011). Reality is broken: why games make us better and how they can change the world. New York: Penguin.
- Mekler, E. D., Brühlmann, F., Tuch, A. N., & Opwis, K. (2015). Towards Understanding the Effects of Individual Gamification Elements on Intrinsic Motivation and Performance. *Computers in Human Behavior*. doi:10.1016/j.chb.2015.08.048.
- Mora, A., Riera, D., & Gonzalez, C. (2015). A literature review of gamification design frameworks. In Seventh International conference on Virtual Worlds and Games for Serious Applications: VS-Games, Sweden (pp. 1–8). doi:10.1109/VS-GAMES.2015.7295760.
- Morrison, B., & DiSalvo, B. (2014). Khan Academy Gamifies Computer Science. In *Proceedings of SIGCSE'2014* 45th ACM Technical Symposium on Computer Science Education, ACM (pp. 39–44).
- Morschheuser, B. S., Rivera-Pelayo, V., Mazarakis, A., & Zacharias, V. (2014). Interaction and reflection with quantified self and gamification: an experimental study. *Journal of Literacy and Technology*, 15(2), 136–156.
- Naik, L. (2015). Gamification Isn't Dead, It's Just Very Misunderstood. *Pulse, July 2015*. Available at https://www.linkedin.com/pulse/gamification-isnt-dead-its-just-very-misunderstood-lee-naik.
- Nevin, C. R., Westfall, A. O., Rodriguez, J. M., Dempsey, D. M., Cherrington, A., Roy, B., Patel, M., & Willig, J. H. (2014). Gamification as a tool for enhancing graduate medical education. *Postgrad Med Journal*, 90(1070), 685–693.
- Nicholson, S. (2015). A RECIPE for meaningful gamification. In L. Wood & T. Reiners (Eds.), *Gamification in education and business* (pp. 1–20). New York: Springer.
- Paiva, R., Barbosa, A., Batista, E., Pimentel, D., & Bittencourt, I. (2015). Badges and XP: an observational study about learning. In *Frontiers in Education Conference (FIE)* (pp. 1–8). doi:10.1109/FIE.2015.7344074.
- Pedro, L.Z., Lopes, A.M.Z., Prates, B.G., Vassileva, J., & Isotani, S. (2015b). Does Gamification Work for Boys and Girls? An Exploratory Study with a Virtual Learning Environment. In *Proceedings of the 30th Annual ACM Symposium on Applied Computing (SAC'15)*, 214–219, doi:10.1145/2695664.2695752.
- Pedro, L., Santos, C., Aresta, M., & Almeida, S. (2015a). Peer-Supported Badge Attribution in a Collaborative Learning Platform: The SAPO Campus case. *Computers in Human Behavior*, 51, 562–567, doi:10.1016/j.chb.2015.03.024.
- Perry, B. (2015). Gamifying French language learning: a case study examining a quest-based, augmented reality mobile learning tool. Social and Behavioral Sciences, 174, 2308–2315.
- Pettit, R.K., McCoy, L., Kinney, M., & Schwartz, F.N., (2015). Student Perceptions of Gamified Audience Response System Interactions in Large Group Lectures and via Lecture Capture Technology Approaches to Teaching and Learning. BMC Medical Education, 15(92). doi: 10.1186/s12909-015-0373-7.
- Poole, S., Kemp, E., Patterson, L., & Williams, K. (2014). Get your head in the game: using gamification in business education to connect with generation Y. *Journal for Excellence in Business Education*, 3(2), 1–9.
- Ryan, R. M., Rigby, S. C., & Przybylski, A. (2006). The motivational pull of video games: a self-determination theory approach, *Motivation and Emotion*, 30, 347–363. Springer Science+Business Media. doi: 10.1007/s11031-006-9051-8
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55, 68–78.
- Seaborn, K., & Fels, D. I. (2015). Gamification in theory and action: a survey. International Journal of Human Computer Studies, 74, 14–31.
- Shi, L., Cristea, A. I., Hadzidedic, S., & Dervishalidovic, N. (2014). Contextual gamification of social interaction towards increasing motivation in social e-learning. In 13th International Conference on Web-based Learning (ICWL2014) (pp. 116–122). Tallinn: Springer. 14–17 August, LNCS 8613.
- Sillaots, M. (2014). Gamification of higher education by the example of course of research methods. *Advances in Web-Based Learning ICWL, Springer Lecture Notes in Computer Science, 8613*, 106–115. Tallinn: Springer Lecture Notes in Computer Science.

- Sillaots, M. (2015). Gamification of higher education by the example of computer games course. In *The Seventh International Conference on Mobile, Hybrid, and On-line Learning (eLmL)* (pp. 62–58). Lisbon: IARIA. ISBN 978-1-61208-385-8. Available at http://www.thinkmind.org/index.php?view=article&articleid=elml_2015_4_20_50048. Accessed 17 Jan 2017.
- Simoes, J., Mateus, S., Redondo, R., & Vilas, A. (2015). An Experiment to Assess Students' Engagement in a Gamified Social Learning Environment, *elearning Papers*, 43, July 2015, DOI: 10.13140/RG.2.1.2384.0488.
- Smith, E., Herbert, J., Kavanagh, L. & Reidsema, C. (2014). The Effects of Gamification on Student Learning through the Use of Reputation and Rewards within Community Moderated Discussion Boards. *AAEE: 24th Annual Conference of the Australasian Association for Engineering Education*, Australia, ISBN: 9780992409906. Available at https://espace.library.ug.edu.au/view/UQ:326086/UQ326086.pdf. Accessed 16 Jan 2017.
- Songer, R. W., & Miyata, K. (2014). A playful affordances model for gameful learning. In TEEM '14 2nd International Conference on Technological Ecosystems for Enhancing Multiculturality, Salamanca, Spain (pp. 205–213). doi:10.1145/2669711.2669901.
- Su, C. H., & Cheng, C. H. (2015). A mobile gamification learning system for improving the learning motivation and achievements. *Journal of Computer Assisted Learning*, 31(3), 268–286.
- Tomaselli, F., Sanchez, O., & Brown, S. (2015). How to engage users through gamification: the prevalent effects of playing and mastering over competing. In Thirty Sixth International Conference on Information Systems, Fort Worth (pp. 1–16).
- Tu, C. H., Yen, C. J., Sujo-Montes, L., & Roberts, G. A. (2015). Gaming personality and game dynamics in online discussion instructions. *Educational Media International*, 52(3), 155–172. doi:10.1080/09523987.2015.1075099
- Tulloch, R. (2014). Reconceptualising gamification: play and pedagogy. Digital Culture & Education, 6(4), 317-333.
- Tvarozek, J., & Brza, T. (2014). Engaging students in online courses through interactive badges. In 2014 International Conference on e-Learning, September 2014, Spain (pp. 89–95). Retrieved from http://elearning-conf.eu/docs/cp14/paper-12.pdf.
- Utomo, A. Y., & Santoso, H. B. (2015). Development of gamification-enriched pedagogical agent for e-learning system based on community of inquiry. In *Proceedings of the International HCI and UX Conference (CHIuXiD'15), Indonesia* (pp. 1–9). doi:10.1145/2742032.2742033.
- Werbach, K. (2014). (Re) Defining gamification: a process approach, persuasive technology. *Lecture Notes in Computer Science*, 8462, 266–272.
- Werbach, K., & Hunter, D. (2012). For the win: how game thinking can revolutionize your business. Philadelphia: Wharton Digital Press.
- Westera, W. (2015). Games are motivating, aren't they? Disputing the arguments for digital game-based learning. International Journal of Serious Games, 2(2), 3–15. http://dx.doi.org/10.17083/ijsg.v2i2.58.

Submit your manuscript to a SpringerOpen journal and benefit from:

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

Submit your next manuscript at ▶ springeropen.com