



Convergence of Gamification and Machine Learning: A Systematic Literature Review

Alireza Khakpour¹ · Ricardo Colomo-Palacios¹

© The Author(s) 2020

Abstract

Recent developments in human–computer interaction technologies raised the attention towards gamification techniques, that can be defined as using game elements in a non-gaming context. Furthermore, advancement in machine learning (ML) methods and its potential to enhance other technologies, resulted in the inception of a new era where ML and gamification are combined. This new direction thrilled us to conduct a systematic literature review in order to investigate the current literature in the field, to explore the convergence of these two technologies, highlighting their influence on one another, and the reported benefits and challenges. The results of the study reflect the various usage of this confluence, mainly in, learning and educational activities, personalizing gamification to the users, behavioral change efforts, adapting the gamification context and optimizing the gamification tasks. Adding to that, data collection for machine learning by gamification technology and teaching machine learning with the help of gamification were identified. Finally, we point out their benefits and challenges towards streamlining future research endeavors.

Keywords Gamification · Machine learning · Learning · Personalization · Behavioral change · Systematic literature review

1 Introduction

During recent decades, the significance of human–computer interaction (HCI) increased exponentially, due to the devastating technological development. Gamification as one of the approaches to improve this interaction, both in terms of efficiency and effectiveness, has been one of the new trends during recent years. There are several definitions in literature regarding gamification, but one of the most agreed upon is the one provided by Deterding et al. (2011), where gamification is defined as “*The use of game design elements in non-game context*”. Gamification is progressively becoming an integral part of every

✉ Alireza Khakpour
alireza.khakpour@hiof.no

Ricardo Colomo-Palacios
ricardo.colomo-palacios@hiof.no

¹ Faculty of Computer Sciences, Østfold University College, B R A Veien 4, 1757 Halden, Norway

computer–human interaction with the goal of encouraging the engagement of individuals to further support and improve user activities (Xi and Hamari 2019).

Gamification has been applied in many different domains, one of such is the domain of software engineering where gamification is involved in software process activities to increase motivation of participants in software projects (Herranz et al. 2014, 2015). A more specific domain is its application in health-related activities (Ahn et al. 2019), where various game elements such as point-reward systems are used on physical activities for children. From a different perspective, its application in human-system interaction is also notable, when used to motivate humans to interact with the system towards the benefit of the system. This approach has been used by Konstantakopoulos et al. (2019), where a gamified framework is developed for smart building infrastructure to stimulate occupants to consider personal energy usage in order to be more environmentally friendly. Other applications are in global climate change (Nastis and Pagoni 2019), web and mobile applications (Zichermann and Cunningham 2011), to name a few.

One of the principal applications of gamification has been in learning and education, towards improving learning processes and their outcomes (Codish and Ravid 2015; Caporarello et al. 2019). Whether it is a business training, school learning or personal life learning context, gamification has been applied to improve participation, engagement, continuity, or evaluation of the learning materials. In the gamification 2020 report provided by Gartner, it is predicted that improved versions of gamification with recent technologies such as machine learning will have a significant impact on different levels of learning platforms. i.e. personal development, organization's employee learning, and higher education (Gartner 2012). In fact, The effects of gamified learning activities have shown to be promising in many studies (de Sousa Borges et al. 2014), but the challenge has always been about its level of effectiveness. As discussed by Caporarello et al. (2019) recently, the focus of effectiveness is mostly on changing the attitude, behavior and knowledge level of the audience. However, the extent to which these goals have been achieved is limited.

On the other hand, there is an increasing amount of data in learning and education systems collected through various channels and methods that need to be transformed into some type of information and analytics for future decision-making activities (Abu Saa et al. 2019; Sin and Muthu 2015). In this context, machine learning can be used as a set of techniques and practices that utilizes the power of data, in order to create machine programs that empower humans with valuable information for various decision-making and analysis tasks. Machine learning is defined in different ways by various authors. Brett Lantz in his book defined machine learning as the process of developing computer algorithms for transforming data into intelligence (Lantz 2015). Kuhn and Johnson (2013) believe that machine learning can be interchangeably called as predictive modeling which in turn can also be described as the process of developing a mathematical tool or model that generates an accurate prediction. As another definition, we can refer to an older definition given by Michalski et al. (1983), in which they stated that the study and computer modeling of learning processes in their multiple manifestations constitutes the subject matter of machine learning. It can be seen that, each of the definitions demonstrates a different viewpoint.

Machine learning, as a powerful analytical tool, has been investigated by many researchers to ameliorate various aspects of learning processes (Monterrat et al. 2015; Khoshkangini et al. 2017; Lopez and Tucker 2018), where it is also called as learning analytics (Seufert et al. 2019). In fact, it has been used extensively in various fields and approaches in the industry alike. As reported by Gartner, 37% of organizations have implemented artificial intelligence (AI) in some form, by 2019 (2019). Nevertheless, the utilization of this potential was not of much attention to the HCI researchers, because of the high level of

knowledge and skills it requires to develop a machine learning model that works efficiently and reliably (Holzinger 2013). However, an increasing interest has been noticed in using machine learning models to optimize gamified learning platforms in recent years. Despite this increasing interest, the developments have been very limited in terms of machine learning concepts and hence, a clear gap can be seen in this era.

Overall, machine learning methods have been used to improve the performance of gamified tasks. In this regard, the fact that personalized adaptive gamification has the potential to enhance individuals' motivation and performance, especially in learning platforms, raises the application of machine learning. Machine learning can tailor the gamified interactions and dynamically configure the interaction parameters. literature presents many instances of this application (Monterrat et al. 2014; Knutas et al. 2017; Lopez and Tucker 2018). Another utilization of machine learning and gamification in learning activities is developing some type of automatic tutoring of the learner through analyzing user interactions and providing proper guidance with the help of gamification (Dalmazzo and Ramirez 2017).

Furthermore, gamification and machine learning can also be used cooperatively to enhance the effect of one another towards a predefined task. For example, in the context of behavioral change, dynamically changing gamified interactions can encourage users to interact with the system in a sustainable manner (Di Lena et al. 2017). Additionally, gamification and machine learning intersect each other in different ways both in academia and industry. There have been several attempts to utilize game design elements to optimize machine learning processes. One of such attempts is gamifying the process of data labeling, where game elements are used to increase the affordances of users to participate in the process (L'Heureux et al. 2017), so-called crowdsourcing.

Screening through literature, it can be concluded that, to the best of our knowledge there are no systematic literature reviews available that investigate the intersection of gamification and machine learning. However, there are plenty of papers focused on reviewing gamification studies, specifically, reviews of the works carried out in various applications of gamification, where gamification of learning platforms has shown to be increasingly interesting because of its effectiveness. Other works in this context are the attempts toward incorporating these two approaches to achieve an improved result in a specific task, where machine learning methods and gamification have been used together in such a way that one benefits from the other. The direction towards which this incorporation took place differs per each work. Therefore, there are plenty of works that deployed machine learning methods to improve gamified tasks and in contrast, others utilized gamification elements to help machine learning practices.

Bearing in mind the aforementioned ideas, the focus of this study is to identify and analyze the convergence of gamification and machine learning, with a rigor focus of its application in learning environments, and to investigate the effect of this convergence over the two technologies. The Massachusetts Institute of Technology (MIT) defined the term convergence as the merging of distinct technologies, processing, or devices into a unified whole that creates a host of new pathway and opportunities. This is then elaborated as coming together of different fields of study, through collaboration among research groups and the integration of approaches that were originally viewed as distinct and potentially contradictory (Sharp et al. 2011). This is where, the convergence of machine learning and gamification as two distinct technologies related to two distinct research groups of Artificial Intelligence and Human-Computer Interaction is raised. MIT also referred to convergence as the blueprint for innovation, which leads to a new integrated approach for achieving advances. These advancements can emerge in both the technologies of machine

learning and gamification, transforming them from two distinct practices into a unified whole. Although many researchers and practitioners are utilizing these two technologies together, it is required to outline their convergence in order to increase coordination, productivity, and innovation.

Hence, a systematic literature review is carried out in this context. This work aims to provide a comprehensive overview of works carried out towards incorporating gamification and machine learning to benefit from the advantages of both. The scope of this research is based on systematic literature review studies in software engineering, as given in the literature.

The rest of this paper is organized as follows: Sect. 2, presents the research methodology, where it is described how the systematic literature review was planned and conducted. In Sect. 3, the threats of validity to this research activity are being presented, followed by the results and analysis of the study, answering the desired research questions in Sect. 4. The conclusions of the study are presented in Sect. 5, pointing out the potential future challenges and directions.

2 Research Methodology

2.1 Motivation

According to the literature, there are many studies incorporating gamification concepts with machine learning methods. Although, to the best of our knowledge there is no literature review available to investigate the path taken in this connection up to now and the prospective future works. Therefore, this study aims to develop a clear insight into the field, to facilitate the understanding of the current state-of-art and identify potentials for future research.

2.2 Research Methods

The research methodology of this study is based on the general guideline of a systematic literature review (SLR) in software engineering provided by Kitchenham and Charters (2004). They defined an SLR as “*A means of identifying, evaluating and interpreting all available research relevant to a particular research question, or topic area, or phenomenon of interest*”. We carried out the steps proposed by Kitchenham and Charters guideline, being (1) Identifying the need for study, (2) Defining the review protocol, (3) Identifying and selecting the primary researches, (4) Assessing the quality of studies, and (5) Conducting the data extraction. Therefore, the goal of performing a systematic literature review is to achieve an overview of the state of the question by identifying, evaluating and interpreting relevant studies in the field of interest.

2.3 Planning

Planning of the research comprises the identification of the need for the study and developing a review protocol.

2.3.1 Need for the Review

The objective of this study is to conduct a comprehensive research investigation on the convergence of gamification and machine learning and hence a systematic literature review is performed. This goal can participate in summarizing the existing evidence of the topic in order to identify any gaps and expectantly providing a framework for future research in the field.

2.3.2 Developing a Protocol

To achieve this goal, the study begins by defining an SLR protocol that includes the rationale behind the research, research questions, search strategy, study selection criteria, procedures, study quality assessment checklist and procedures, data extraction strategy and synthesis of the extracted data.

2.4 Research Questions

We have established the research questions for the study based on the main goal of this systematic literature review, which is to determine insight into the research field and research categories, along with the respected outcomes that have been provided in the topic of gamification and machine learning confluences. Furthermore, the study tends to highlight the existing evidence, gaps and future path for the field. It is then possible to formulate the research questions in the following manner:

1. What is the reported usage of machine learning in gamification?
2. What is the reported usage of gamification in Machine Learning?
3. What are the reported effects of using Gamification over Machine Learning?
4. Which aspects of gamification are affected by machine learning?
5. What are the benefits and challenges in connection with the combination of machine learning and gamification?

2.5 Search Strategy and Resources

2.5.1 Search String

In this systematic review, we first developed a search string to extract the related primary studies to the topic under consideration. In order to create the search string, we first initiated the search with a broad search string aligned with the RQs using keyword derivation. After this, we ran several pilot searches and modified some of the terms in the string. As a result, a general search string is produced which is used in every publication channel.

In this regard, we decided on two broad search terms, namely, “gamification” and “machine learning”. We created the final search string to be as follows: (“Gamification” AND “Machine Learning”).

The structure of the search term is formulated using the approach deployed by (Brereton et al. 2007), where a Boolean AND is deployed to link the major terms.

Regarding the search strategy to conduct this systematic review, the publication channels used in this research along with the inclusion and exclusion criteria under consideration during the screening of papers are as follows:

2.5.2 Search Resources

In this study, we planned to identify and investigate all available literature about the use of gamification and machine learning along with each other. In this regard, after consulting with some of the domain experts and analysis of some of the publication channels, we identified a number of electronic databases related to this research field. We have identified 4 popular publication channels, namely, IEEE Xplore, ACM Digital Library, Springer Link, and Science Direct, that have a special focus on computer science. There has been a total number of 1302 search result which is summarized in Table 1. This study is conducted during the spring 2019 and the reflected search results are until the end of March 2019.

Every paper from the search result was reviewed carefully based on its title, abstract, keywords and conclusions to identify relevant papers. Papers then have been classified into three categories of (1) Matching papers, that exactly reflect our topic (2) Somehow related papers that are relevant to the topic to some extent and (3) Excluded completely irrelevant papers. The outcoming papers then have gone through a full-text retrieval and analysis based on our inclusion and exclusion criteria.

2.5.3 Inclusion and Exclusion Criteria

In this step, we develop a set of inclusion and exclusion criteria to be applied to every research paper that has been retrieved during the first round. These criteria have been used when the full text of final retrieved papers has been analyzed, either including a paper in the research or excluding it.

2.5.3.1 Inclusion Criteria

- Study related to the utilization of gamification and machine learning.
- Study should be published in a peer-reviewed publication channel.

2.5.3.2 Exclusion Criteria

Table 1 Search resources

Source	Number of search results
IEEE Xplore	351
ACM Digital Library	294
Springer Link	479
Science Direct	178
Total	1302

- Study related to gamification topic but not to machine learning methods.
- Study related to machine learning methods but not to gamification concepts.
- Study in a language other than English.
- Study that is not identified as peer-reviewed.
- Full text of the study is not available in the respected source.

2.6 Data Extraction

Every paper retrieved from a resource was first documented and saved in a reference manager software. A piece of preliminary note is attached to each paper, containing the relevant part of the paper which made the study adhere to the inclusion criteria. Then, every paper was gone through a secondary analysis and a data extraction form have been created, in which all the findings and information of each paper was recorded. From every paper, 7 types of data were extracted and recorded in Excel, as given in Table 2.

The process of data extraction took place in three stages: Firstly, a primary analysis was carried out to collect the standard information and inclusion–exclusion criteria types of data, and secondly, a more careful examination over each of the papers were performed to collect the research question addressing level, answers to research questions and specific machine learning and gamification related data. Finally, in the third round of selection, the study quality assessment type of data was collected to identify papers that have contributed to the field positively. In the following section, we explain the quality assessment criteria.

2.7 Quality Assessment of the Literature

After identifying the studies which are relevant to our review, we evaluate the superiority of each study by passing each of them through a specific checklist. This process is carried out in order to confirm the reliability of the studies. The checklist comprises a set of conditions and questions that verifies the effectiveness of each study for the field. As stated in Kitchenham guidelines (Kitchenham 2004) based on the Center for Reviews and Dissemination (CRD) Guidelines (Cochrane Handbook for Systematic Reviews of Interventions 2019), “*the quality relates to the extent to which the study minimizes bias and maximizes*

Table 2 Types of the extracted data from studies

Type	Data
Standard information	Title, authors, publication year, journal or conference name, publisher, paper type, number of citation, average citation per year, date of extraction
Inclusion and exclusion criteria	Language, peer-reviewed, exclusion reason
Research questions	The extent to which each research question is addressed utilizing a scale from 0 to 1 (RQ1, RQ2, RQ3, RQ4, RQ5)
Answers to research questions	Direction of the study, usage of ML in gamification, usage of gamification in ML, effects of ML over gamification, effects of gamification over ML, benefits reported, challenges reported
Machine learning related	ML problem, ML algorithm, performance measurement, accuracy
Gamification related	Gamification element, improvement measurement, Index of Improvement
Study quality assessment	Checklist (study type, bias, validity, generalizability)

Table 3 Study design hierarchy for software engineering as suggested by Kitchenham (2004)

1	Evidence obtained from at least one properly designed randomized controlled trial
2	Evidence obtained from well-designed pseudo-randomized controlled trials (i.e. non- random allocation to treatment)
3-1	Evidence obtained from comparative studies with concurrent controls and allocation not randomized, cohort studies, case–control studies or interrupted time series with a control group.
3-2	Evidence obtained from comparative studies with historical control, two or more single arm studies, or interrupted time series without a parallel control group
4-1	Evidence obtained from a randomized experiment performed in an artificial setting
4-2	Evidence obtained from case series, either post-test or pre-test/post-test
4-3	Evidence obtained from a quasi-random experiment performed in an artificial setting
5	Evidence obtained from expert opinion based on theory or consensus

Table 4 Study quality assessment questions

Quality concept	Question number	Question	Yes No	
Selection bias and measurement bias	1	Does the study choose the subjects under study randomly?		
	2	Are the outcomes of the study interpreted based on the subjects under study?		
Validity	3	Is the study carried out with a scientific methodology?		
	4	Are the methods used well-defined and verifiable?		
Generalizability	5	Is there a proper use-case to test the results?		
	6	Are the results general enough to be expandable to other situations?		

internal and external validity". We first give every study a level of value based on the Kitchenham study design hierarchy to ensure a minimum level of quality. The Guideline suggests the Hierarchy of evidence, as presented in Table 3. These values indicate the level of the quality of the paper.

The next step is answering the set of questions about the quality of each study. The answer to each question is calculated on a scale of 0 to 1. Each of these questions is related to either one of the quality criteria, namely, bias, validity, and generalizability. The set of questions is presented in Table 4. The result of this stage is documented in our final papers list to demonstrate the level of work in the field.

The results of the quality assessment process are provided in appendix B. The provided results can be used to identify the quality and superiority of every paper.

3 Threats to Validity

We found that since it is important to clearly present the limitations of our research and the approaches we have employed to reduce those limitations, in this section, we discuss the threats to validity based on the following validity threats provided by Petersen and Gencel (2013), (1) descriptive validity (2) theoretical validity (3) generalizability (4) interpretive

validity (5) repeatability. Although the guideline presented by Petersen and Gencel is provided for software engineering research, we decided to report the threats to validity because we strongly believe that this guideline is most suitable and adaptable to our research, given that it provides clear understanding of the threats and practical way of addressing them. Each of these is described and discussed in the following sections.

3.1 Descriptive Validity

Based on Petersen and Gencel (2013), descriptive threats are to make sure we can describe the objective/subjective truth accurately. That measures the extent to which the observations are described accurately and more precisely, in an objective manner.

Authors have minimized this threat using two techniques: firstly, by using data extraction forms to analyze the studies in a systematic manner, without missing any significant information from any paper. Secondly, is the way that inclusion and exclusion criteria were examined for each study. To overcome this threat, authors used a checklist type of inclusion–exclusion criteria and if any of the criteria is not checked during the first round of reading the title, abstract and keyword, the paper is not excluded right away in order to lower the risk of missing a relevant study. Instead, an additional analysis carried out by screening the full text of the document and its references to find any relevant information.

3.2 Theoretical Validity

In this section, researchers investigate the theoretical validity which is to identify the confounding factors and verifying whether we seize what we intend to seize (Petersen and Gencel 2013).

Authors may select only those studies that demonstrate a satisfactory result of combining gamification and machine learning and ignore those studies that indicate a negative outcome of the combination. However, those possible negative outcomes can be important for our final research question that addresses the challenges of converging these two concepts. Furthermore, another possible threat to this validity is the way the result of each paper is evaluated. We investigated two parameters: (1) The accuracy of the respective machine learning method used in each study and, (2) The index of measuring the effect of the deployed gamification. In fact, these two factors have been reported only, without being used for inclusion and exclusion of a study, in order to minimize the threat to the theoretical validity.

3.3 Generalizability

The next validity threat is about generalizability, which deals with the degree to which the results of the study can be generalized either internally, that is within groups and communities, or externally, that is across groups and communities (Petersen and Gencel 2013).

Authors included every study that relates to the combination usage of gamification and machine learning and additionally those studies that are specific to one of the two, which is either gamification or machine learning but, the result can later be implemented in the other one. However, since the topic is to disperse in the sense of the direction towards which one is benefiting from another, there may be a risk of threat to this type of validity.

3.4 Interpretive Validity

Interpretive validity is about confirmation that the conclusions or inferences of the study are drawn correctly and in an objective way (Petersen and Gencel 2013).

To ensure the interpretive validity, the data extraction parameters are formulated in detail such that, investigating the extent to which each paper is addressing a particular research question and interpreting the exact answers that each study presents for the respective research question. In this way, the use of a systematic extraction form reduces the chance of threat to this validity. Moreover, the assessment of the quality of each paper precisely with the predefined set of parameters may also contribute to lowering the risk of interpretive validity threat.

3.5 Repeatability

Repeatability validity check verifies that the data collection and analysis approaches along with instruments used are defined neatly, in order to make repeatability and reproducibility possible (Petersen and Gencel 2013). However, Petersen and Gencel argued that repeatability or reproducibility is ensured by addressing the other four main threats to validity already mentioned above. Consequently, in this study, authors used data collection forms and reference managers to document every step of the procedure and followed the guideline for performing systematic reviews by Kitchenham (2004), which ensures the systematic walkthrough of the procedure, ensuring reproducibility of similar results to this project.

4 Results and Analysis

4.1 Number of Papers

The initial result of the research indicated a total number of 1053 papers, out of which 89 were extracted by reading the title, abstract and keywords. The refining of the paper continued by reading the full text of the studies, and a total number of 32 studies were selected and used as the basis for this SLR. Table 5 shows the number of initial results along with a final number of extracted papers. After the data extraction phase, all of the studies have been given a unique identification number to enable easier referencing for further analysis. In the rest of this paper, the studies are referred to as in the form of their identity key from (S1) to (S32), presented in appendixes.

Table 5 Number of studies selected

Source	Number of search results	Number of paper after reading abstracts	Number of papers after reading full-text
IEEE Xplore	351	29	18
ACM Digital Library	294	47	15
Springer Link	479	41	6
Science Direct	178	8	4
Total	1302	125	43

4.2 Demographics

In this section, we have present the number of papers based on the two following categories:

1. Number of papers per year.
2. Number of papers per topic.

The idea here is to understand the age of the research topic in general, and the trends of the subject, being the convergence of machine learning and gamification. This approach of analyzing the study topic trends based on literature is also presented and deployed in other studies, such as Kitchenham et al. (2009) and Sánchez-Gordón and Colomo-Palacios (2019).

4.2.1 Publishing Year

In this part, we present the number of studies distributed by their publishing year. In this systematic literature review, we have retrieved the data for all the previous years. Interestingly, the first paper was published in 2014, which shows that our topic of interest emerged recently. In 2018, researches focused more on the topic by presenting 15 studies, followed by 4 papers published in 2019. The number of studies per year is demonstrated in Fig. 1.

The above presentation demonstrated the emergence of the topic in last 6 years. After a slight attention in two years of 2014 and 2015, researches shed more light on the topic in 2016. There was a noticeable increase during this year, after which the hype was again disappeared. However in 2018 more number of studies worked around this topic that can be due to recent improvements in various technologies, such as machine learning, gamification, sensors, mobile devices, etc.

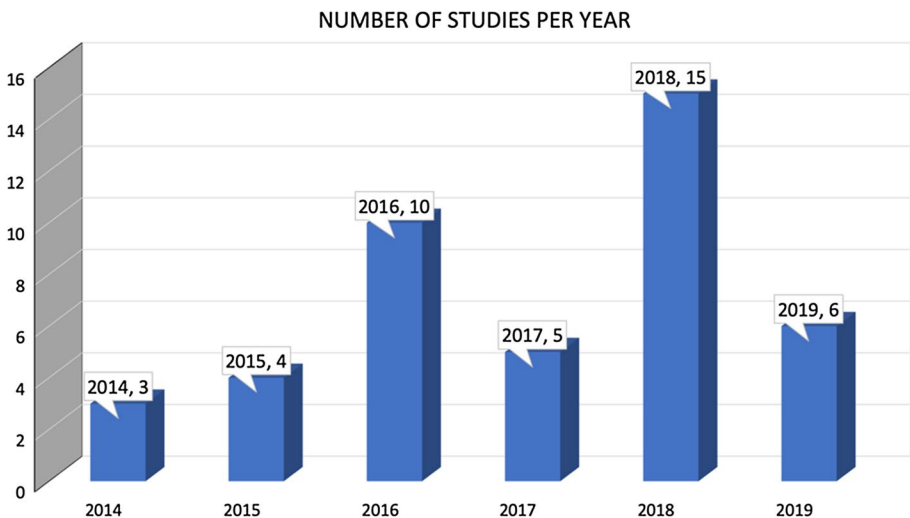


Fig. 1 Number of studies per year

4.3 Dispersion of the Topics in Literature

The next investigation carried out was on the dispersion of the topics. The results show that there is a considerable amount of studies focused on learning, with 16 papers comprising more than 40% of the literature.. This introduces the potential of the research topic in learning and education and the interest of corresponding researchers. Personalization is the next topic that has been the focus of the studies. One of the major goals of gamification tasks has been to support human behaviors. But the problem is normally the long-lasting effects of behavioral change stimulants. In this regard, machine learning has contributed to the customization of the user experience in order to encourage the continuation of user engagement. The following topic in the list includes studies that focus on behavioral change. The following interesting topic is crowdsourcing. Interestingly, in these studies the direction of the convergence was in reverse order, that is, from gamification towards the improvement of machine learning. These attempts were to improve the participation rate of the users in labeling the required data for the machine learning training step. Other efforts were summarized in affective computing, sentiment analysis, health and medical activities, and lowering energy consumption. Figure 2 shows this dispersion.

4.4 Answer to Research Questions

In this section, we perform a detailed analysis of the studies based on the research questions. Section 4.4.1 investigates the papers to answer the “RQ1: What is the reported usage of machine learning in gamification?”. Next Sect. 4.4.2, scrutinizes the answers to “RQ2: What is the reported usage of Gamification in Machine Learning?”. We then present the answer to “RQ3: What are the reported effects of using Gamification over Machine Learning?” in Sect. 4.4.3. Followed by the response to “RQ4: Which aspects of Gamification are affected by machine learning?” in 4.4.4. Finally, “RQ5: Benefits and challenges in connection with machine learning and gamification?” is addressed in Sect. 4.4.5.

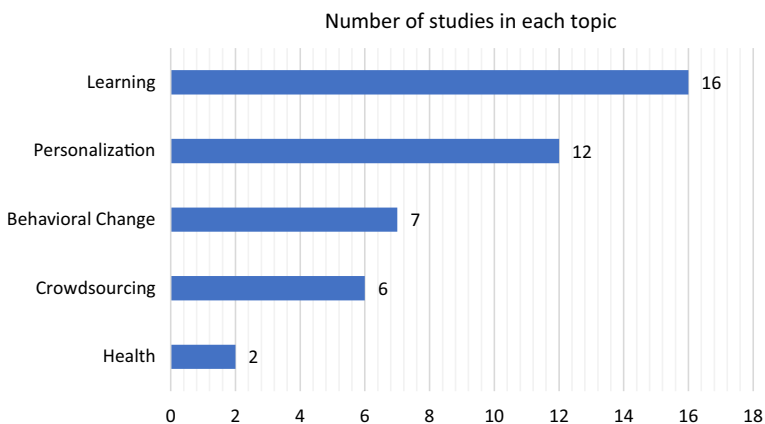


Fig. 2 Number of studies in each topic

4.4.1 RQ1. What is the Reported Usage of Machine Learning in Gamification?

To answer this research question, we identified the papers being in the direction of machine learning towards gamification. In this context, papers that attempted to use a machine learning technique or approach in order to enhance and support gamification tasks are identified. Following the identification of these studies, we further inspected each paper to find out the corresponding machine learning concept applied.

We have found 3 main areas of application based on which corresponding studies have been analyzed. These areas are presented as follows: 1. Learning, 2. Personalization, 3. Behavioral Change. Table 6 shows the goals of the reviewed papers in each category specific to the RQ1. However, there are papers in other applications that we address in the next research questions.

4.4.1.1 Learning Affective state recognition is claimed to have an influence over optimization of the learning process and its outcomes by providing learning interaction, for example in a gamified learning that can be personalized. In this regard, authors in (S10) presented a method towards retrieving the affective state of a student while interacting with a serious game learning platform by applying machine learning (Ghaleb et al. 2018). In this work, the so-called model Theory of Flow is utilized to link the affective state of the student to the user-platform interaction. This model presents three states of boredom, engagement, and frustration. A support vector machine algorithm is used to train a classifier to distinguish between different affective states. Authors have reported precision of 67%. Hence, the machine learning technique can be applied to predict the affective state of students when interacting with a gamified learning platform.

In a different scenario, machine learning and gamification can work together to facilitate learning. For instance, authors in (S11) developed an application in which deep learning and gamification are used to assist 3-4 years old children to learn generalizing objects (Suresh et al. 2018). They have used an approach called joint-embedding visual question answering, leveraging on a convolutional neural network (CNN) and a stacked recurrent neural network (RNN) called long short-term memory (LSTM). The application of machine learning in this gamified learning platform helped to extend the learning context by making the game elements dynamic and intelligent, hence optimizing the learning process.

In a similar effort towards the advancement of learning processes, authors in (S12) designed a serious game in order to assess the medical student's knowledge level (Lima et al. 2016). Various diagnostics were simulated through a gamified virtual reality assisted platform. Gamification features were used to motivate users to work with the system for a longer duration. On the other hand, a machine learning algorithm was employed to develop a disease classification model in order to assist the working of the platform. The researchers of this study used the freely available machine learning API, so-called Weka, which offers a set of machine learning algorithms.

One of the techniques to improve gamified applications is the personalization of the game elements that will be discussed in detail in the next section. With regards to the learning platforms along with the adaptation technique, personalization is claimed to be an effective solution as well. As an instance, authors in (S19) proposed an approach towards systemizing the selection of personalization strategy with the help of machine learning (Knutas et al. 2018). In their demonstration phase, they used a CN2 rule induction algorithm to model a classifier to distinguish between various situations that take

Table 6 Goals of the reviewed papers in each category

Category	Goals	No. of papers	References
Learning	<p>Personalized game, context-specific game-based application, adding information to gamification,</p> <p>Calibrate results from the gamified task,</p> <p>Emotion detection, detection of the user's affective engagement during gameplay scenarios provides forecast to motivate gameplay and give suggestions,</p> <p>Clustering student based on their performance, automatic recognition of player's engagement, adjust game content in order to increase the chances of players attaining the game's specific learning objectives concerning prosocial skill, detecting left-hand finger selection and performance in real-time.</p>	10	<p>Ghaleb et al. (2018), Suresh et al. (2018) and Lima et al. (2016)</p> <p>Psalts et al. (2018) and Knutas et al. (2018)</p> <p>Barata et al. (2015)</p> <p>Stefanidis et al. (2019)</p> <p>Palavalli et al. (2014) and Anparasanesan et al. (2019)</p>
Personalization	<p>Predicting performance by facial key point</p> <p>Personalized game context-specific game-based application adding information to gamification</p> <p>Calibrate results from the gamified task</p> <p>Emotion detection</p> <p>detection of the user's affective engagement during gameplay scenarios provides forecast to motivate gameplay and give suggestions</p> <p>Clustering student based on their performance automatic recognition of player's engagement, adjust game content in order to increase the chances of players attaining the game's specific learning objectives concerning prosocial skills predicting the future state of charge of vehicle's battery at some fixed time-offset</p> <p>Optimizing the challenge selection process and the evaluation results within an on-the-field game promoting sustainable mobility habits</p>	10	<p>Lopez and Tucker (2018)</p> <p>Ghaleb et al. (2018) and Knutas et al. (2018)</p> <p>Barata et al. (2015)</p> <p>Stefanidis et al. (2019), Di Lena et al. (2017) and Khoshkangini et al. (2017)</p> <p>Karaliopoulos et al. (2016)</p> <p>Lungu (2016)</p> <p>Schäfer et al. (2018)</p>

Table 6 (continued)

Category	Goals	No. of papers	References
Behavioral change	<p>provides forecast to motivate gameplay and give suggestions</p> <p>Personalized gamification</p> <p>Clustering student based on their performance</p> <p>automatic recognition of player's engagement, adjust game content in order to increase the chances of players attaining the game's specific learning objectives concerning prosocial skills, Uses machine learning to recognize and count stairsteps and targets to persuade people to use stairs instead of elevators</p> <p>Sounds above 60 dB were extracted from recordings for snore classification with machine learning support vector machine classifiers</p> <p>To improve forecasting performance, predicting the future state of change of the vehicle's battery at some fixed time-offset</p>	7	<p>Akasiadis et al. (2015), Cimán et al. (2016), Goswami et al. (2019) and Ortiz-Catalán et al. (2016)</p> <p>Raptis et al. (2018) and Konstantiakopoulos et al. (2019)</p> <p>Di Lena et al. (2017)</p>

place in a computer-supported collaborative learning context (CSCL), followed by the recommendation of a personalized gamification activity.

Furthermore, assessing student performance while interacting with a gamified learning environment can assist professors in dynamically changing the learning environment to adapt the condition concerning student's performance. (S20) proposed a method to achieve this task with the help of machine learning methods (Barata et al. 2015). The authors in this study used the data from a gamified learning task to perform a clustering technique grouping students' performance into different categories, namely, achievers, disheartened, underachievers, and late awakeners. They have used an algorithm called expectation-maximization (EM) and claimed that by using this student clustering technique, it is possible to predict student behaviors in the early stages of their interaction and hence, an adaptive and smart learning environment can be developed.

The above-mentioned work was focused on the performance assessment of the students while interacting with the platform to proactively improve the performance of the students, whereas, a more constructive manner of improving gamified learning platforms is to improve the performance using an adaptation mechanism. (S21) presented an approach where machine learning methods and algorithms are applied in order to adapt the game content to specific characteristics of every user, assisting with the learning process (Stefanidis et al. 2019). They have used a variation of a machine learning algorithm ϵ -greedy called, ϵ -decreasing algorithm to obtain the engagement profile of the user.

Last but not least, are some of the specific subject learning platforms that require some type of machine learning methods to implement a required application. For instance, the work presented in (S31) is an attempt towards assisting in music education, air-violin self-learning specifically, (Dalmazzo and Ramirez 2017). Authors in this work deployed two machine learning models created using decision trees and hidden Markovian and developed a fingering recognition model in a gamified virtual violin platform. The study presented in (S38) used gamification and machine learning to track child's brain development and to participate its improvement (Anparasanesan et al. 2019). Authors of this study provided an approach in which the K-means Clustering algorithm of machine learning is used to identify the initial child's brain status and suggesting a proportional brain game. The proposed solution starts with an initial evaluation of the child's brain status and assess if the current brain development is proportional to the child's age. Then, a data set is created from the initial evaluation which is then undergone the unsupervised learning algorithm of K-means clustering to identify suitable gamified tasks. Various tasks have been designed as memory games, attention games, games for concentration, and games for response time. The study demonstrates that ML can be used to provide suitable gamified tasks to particular target group based on their cognitive abilities.

4.4.1.2 Personalization One of the areas in which machine learning was used to optimize the results of gamification was personalizing the gamified tasks for each user to adapt the gamification aspects. In this regards, the study in (Teasley 2017) presented that the one-size-fit-all feedback system does not always perform satisfactory and can have mixed results, and hence a personalized feedback may moderate the negative effects. Related to gamification, the authors in (S4) used machine learning to predict the performance of each user and to adapt the complexity of the gamification task to the respective user (Lopez and Tucker 2018), they used the facial key point data in conjunction with a support vector machine algorithm and predicted the performance of each user by 76.8% accuracy.

Adding to that, another way of personalizing the game contents using the performance of the users is by understanding the performance profile of each user and customizing the game contents, accordingly. Authors in (S20) proposed a technique towards identifying the performance profile of each user by utilizing a machine learning-based clustering method to personalize the game contents based on the potential of each user (Barata et al. 2015). This approach is also an attempt towards the adaptation of the gamified platforms that can improve the engagement of the users with the gamified task.

Therefore, the adaptation of the gamified platform is another mechanism that tailors the game content with respect to the context's specific situations. As stated in Knutas et al. (2018), adaptive gamification differs from personalized gamification. In fact, adaptation refers to the gamified system react to different situations, whereas personalized gamification is modifying the game contents based on users' specific characteristics. Although, the work presented in (S21), developed a system called an adaptation manager that is capable of identifying the player characteristics and adjusts the game contents accordingly with the help of machine learning algorithms (Stefanidis et al. 2019). However, they have also created some kind of adaptable game scenarios and elements that can be selected at some specific points of interaction with the game and are offered by the adaptation manager framework of the system. The adaptation mechanism has two parts, online and offline adaptation mechanisms. The offline mechanism is used to assist the users' in-game performance, whereas the online mechanism is concerned with the level of engagement of the user.

In a different perspective, authors in (S33) considered the problem of personalizing the mobile crowdsensing processes (Karaliopoulos et al. 2016). They believe that the main concern with the crowdsourcing activities is to increase the level of contribution of the users. This contribution is subject to various types of parameters such as incentives provided to users and its proportionality to the practicality of the task. To overcome the challenge of optimizing the task proportionality to the incentives offered to the users and maximizing the chance of user contributions, authors proposed a novel approach of using the machine learning technique of logistic regression. They provided an approach of modeling the past user behaviors in previous crowdsensing applications, trying to predict the optimal pair of (task, incentive) for the current target. The proposed solution was evaluated against the real data of an online questionnaire approach to collect user preferences and the results have shown to be promising in terms of level of contributions. In another effort, (S36) proposed a machine learning based technique to automatically analyze serious games by capturing players behaviors (Palavalli et al. 2014). Authors believe that by identifying and analysis of players activities through a video based technique, it is possible to determine some of the parameters influencing the learning potentials of serious games.

(S42) carried out a study on encouraging children to do more physical activity using a personalized gamified feedback system (Schäfer et al. 2018). The machine learning techniques of SVM and random forest are utilized to classify the initial activity level of a child prior to the usage of the system. The classification models are created by learning over a pre-labeled mobile sensor data comprising children normal activities. Then, based on the users specific activity class, a personalized gamified feedback is provided to the users. The feedback system is based on visualizing the activity level of the user by means of showing an Avatar and a motivating message to the user. At the end of each day a bar chart showing the progress of the user is also provided to increase the awareness. In the experiment conducted in this study, Random Forest outperformed SVM with higher accuracy and made the personalized gamification task more engaging.

Another effort towards optimizing the engagement of the users in a gamified platform is to statically configure the game difficulty. Adjusting the difficulty of the game statically

is defined as configuring the game difficulty prior to the start of the game based on previous game play data of various players. This is against the dynamic adjustment of the game difficulty during the game based on the user performance. This approach is carried out in the study presented in (S37) by Khajah et al. (2016). The authors of this study proposed the application of a machine learning technique known as Bayesian optimization to manipulate the game difficulties. They have distinguished between two types of manipulations, namely, overt and covert. Overt manipulations are those that players can feel during the game play, in contrast, covert manipulations are those that are less visible and includes some aspects of the game that the player cannot distinguish. Authors argued that, the overt manipulation does not have any effect on user engagement compared to the covert manipulation that shown improvements in engagement in the experiments they conducted.

Rather than predicting the performance of the users in order to personalize the game content, one can understand the affect of the person interacting with the gamified platform in different stages of the game to learn the reactions of the user in different scenarios. This task is called affective computing and is discussed in a later section. But what is important here is that this approach leads to personalization, as well. In the work presented in (S10), authors performed a subject-based analysis to evaluate the adaptive nature of the learning process and highlighting the employment of interaction features towards creating a customized and personalized learning environment (Ghaleb et al. 2018). They have presented an accuracy of 74% using a support vector machine algorithm in recognition of the corresponding affective states.

(S43) is another effort to presented a framework for developing a language learning platform (Lungu 2016). The study is focusing on learning vocabulary of a new language by combining the free reading exercises and optimal repetition of learned concepts. The system comprises various modules. First, a machine learning agent that evaluates the present knowledge of the learner based on previous interactions of the user with the system. Second, a motivator agent that utilizes gamification elements to provide suitable feedbacks to the learner in order to keep the learner engaging with the system. Therefore, setting the feedback dynamically based on users performances personalizes the user experience and improves the over system efficiency.

Above all of the advantages of personalization, tailoring gamified designs and contents to each of the corresponding users is not an easy task to achieve. Hence, there have been attempts to simplify and structure the process of developing personalized gamification. For instance, (S19) proposed a technique in which personalization is carried out with the help of a machine learning algorithm-based content selection (Knutas et al. 2018). Author suggested that to overcome the difficulties of selecting personalized contents, gamified platforms can benefit from the machine learning-based algorithms to automate personalization. Furthermore, this approach may convert the process of personalized content selection into systematic and repeatable means.

Despite the difficulties of deploying personalization for gamification tasks, it has been one of the most prominent reasons for applying machine learning methods in human–computer interaction systems. Authors in (S29), proposed a system in which machine learning is used to personalize a gamified In-vehicle human–machine interface (Di Lena et al. 2017). In this system, a prototype is developed that is equipped with a dashboard offering personalized challenges to the drivers based on their estimated energy consumption that has been found with the help of machine learning methods.

Nevertheless, verifying a behavioral change success is impractical in the short-term and requires the long-term engagement of users to be able to judge on the behavioral change attainment. To achieve this goal, there have been several attempts to personalize

the game content in order to motivate users to continue using the system for a longer duration. One of which used machine learning methods is (S30) that developed a framework known as procedural content generation (PCG), which is a solution to sustain the interest of players by tailoring the game contents based on the specific users' profile and characteristics (Khoshkangini et al. 2017). The framework uses machine learning methods to improve the challenge selection based on the users' historical interactions with the system, eventually recommending new challenges that can be suitable, according to the players' profile and characteristics.

4.4.1.3 Behavioral Change Applying gamification for persuasive technologies to foster behavioral change activities have gained a lot of attention in recent years (Kappen and Orji 2017). Furthermore, promoting environmental sustainability practices taking from energy saving to pollution control has been the focus of many studies and researches as well (Akasiadis et al. 2015; Tserstou et al. 2017; Konstantakopoulos et al. 2019). As an instance, (S17) provided a gamified interface to promote renewable energy usage by residential buildings (Akasiadis et al. 2015). They used machine learning methods to provide a forecast possible electricity consumption rescheduling, hence motivating residents to take appropriate action towards the goal of the gamified task. They have trained various regression algorithms and compared them to choose the best one. They have concluded that Support Vector Regression is the best for their application, since it trains fast and is scalable compared to other regression algorithms and neural networks.

Another attempt towards achieving behavioral change is presented in (S22), where authors used machine learning methods in a serious game mobile application that identifies stairsteps and encourages people to use stairs instead of elevators (Ciman et al. 2016). This work aims to increase peoples' daily physical activity with the help of a smart serious game. Authors in this work developed a mobile application that holds a game that records and analyzes the data from smartphone sensors. The application also counts the stairsteps taken by the user to provide persuasive game elements. The task of recognizing stairsteps is a classification problem that is implemented using three machine learning algorithms, namely, decision trees, K-nearest neighbors (KNN), and kernel optimization of the margin distribution (KOMID). KOMD combined with smoothing the data, demonstrates a better result comparing to other algorithms. The algorithm shows a precision of 91%.

Machine learning-based classification of a phenomenon from the data retrieved from a gamified user interaction has been simultaneously used as a method in various studies. Serious games normally are used to provoke users to provide some specific type of data that can later be used for the desired application. This approach is also used for the intention of some behavioral changes. As an instance, (S24) used a mobile game application to deliver oropharyngeal exercises to treat snoring (Goswami et al. 2019). A machine learning support vector machine classifier is used to classify the extracted recordings of snore from the participants. The authors in this study conducted a randomized controlled trial over 16 participants with habitual snoring to play the game daily. Results are shown to be successful after 8 weeks of trial, as reported by the bed partners of all the participants.

One of the main contributions of behavioral change studies has been towards provoking buildings' human occupants to use energy more efficiently and effectively. In this regard, gamification approaches have been used by many researchers to create an interface for occupants to interact with the energy usage of buildings, hence incentivizing the energy-efficient behavior (Konstantakopoulos et al. 2019).

Another area, in which being a behavioral change influencer was the study's focus, is encouraging Eco-Driving behaviors. As an instance, (S29) fosters Eco-Driving behaviors by proposing an In-vehicle dashboard, functioning based on machine learning and gamification techniques (Di Lena et al. 2017). Researchers in this work provided various gamified challenges based on the driver's braking style. In other words, the system identifies the user's driving behavior and predicts future battery usage and offers some choices of battery saving behavioral change activities. Although the results of this work are still under investigation, it demonstrates to have noticeable outcomes. Staying in driving behavioral change, the study presented in (S39) developed a solution for recognizing drivers way of holding the steering wheel and providing them suitable feedback to provide an awareness that leads to behavioral changes (Raptis et al. 2018). They have used a support vector machine to identify drivers attentiveness by classifying them to be either as attentive or inattentive. Finally a gamified feedback system is provided at the end of the driving session to make the users aware of their steering wheel holding habits. Interviews with the users demonstrated that drivers are practicing dangerous behaviors during driving that they are not aware of. This work presented the application of ML to provide suitable gamified feedback to users towards possible behavioral change activities.

4.4.2 RQ2. What is the Reported Usage of Gamification in Machine Learning?

To answer RQ2, we first identified those papers with the direction from gamification towards machine learning. Studies that comprised of using gamification aspects and mechanics in the machine learning process were taken under inspection in this section.

Machine learning methods are well known for their power in learning from the data and predict future values for some desired target variable. However, this process requires a vast amount of labeled data, and in this regard, gamification aspects were used in combination with a crowdsourcing approach, extensively by many researchers, to produce labeled data. One of such is (S2), where a gamification framework for sensor data analytics was proposed (L'Heureux et al. 2017). They used gamification to motivate users to perform targeted action through the use of gaming mechanics. The action is in fact, labeling the sensor data which later was gone under a supervised classification problem by the K-Nearest neighbor algorithm. The authors reported improved sensor data analytics with 88.6% of accuracy.

In another data collection task, authors in (S8) used a serious game for retrieving data regarding cognitive neuroscience (Murphy et al. 2018). Their goal was to quantify cognitive aging and performance in a home situation, where they utilized a game with the purpose to make data analysis possible by collecting desired data at home. Then, they used random forest and linear regression algorithms to analyze the collected data. Other data collection tasks were (S15), where authors utilized gamification aspects by providing a web game, supporting it with a deep learning algorithm to create a facial emotion dataset (Li et al. 2016). They reported a classification accuracy of 80% by deploying the CNN algorithm over the dataset collected through the proposed game with a purpose.

The study presented by (S34) used gamification element of providing users statistics to encourage users to engage with a mobile application (Urh and Pejović 2016). This application collects and labels user's data that is later utilized by machine learning techniques. Hence, the gamification power of motivating users have been utilized to higher the quality of data collection for the purpose of machine learning. However, the authors of this study did not measured the increased level of engagement properly.

On the other hand, gamification and machine learning were used considerably in the learning and education context. (S3) utilized the power of gamification to improve the engagement of students in the learning process in teaching them machine learning. Gamifying the learning process to translate the complexity and technical knowledge of machine learning was achieved in this study (Sakulkueakulsuk et al. 2018). They reported that students had more fun, engagement and hands-on interactivity while learning ML. In the same criteria, (S5) and (S7) proposed a game with purpose with badge and leaderboards of gamification elements to teach machine learning to students through a gameplay activity. All of these studies gamified the learning process to minimize the hassle in learning technical contents, especially for non-technical students (Anderson et al. 2014; Rattadilok et al. 2018).

From a different viewpoint, authors in (S40) believe that gamification and machine learning can be used together to access users private data (Acharya et al. 2019). They have anticipated that it is possible to put users into situation of providing their private data unintentionally. To support their argument they have developed a mobile application that motivates users to play a game which is getting them into performing particular patterns to collect and create training data tailored to the user. At the same time, the authentication pattern of the mobile phone is recorded every time user log into his phone in background. A possibility that is available in Android devices. Later, ML algorithms, namely support vector machine and logistic Regression is used to predict the users lock pattern. Therefore, once again gamification has been used for data collection required for machine learning activities. However, this time a malicious target have been followed.

Until now, most of the efforts regarding the usage of gamification to assist with machine learning tasks have been around the data collection and labeling tasks which are considered to be based on the supervised machine learning problem. However, other types of machine learning problems, being unsupervised and reinforcement learning can also benefit from the advantages of gamification. Holzinger (2016) presented the concepts of interactive machine learning (iML), where the human agents or human-in-the loop can interact with the algorithm and optimize its learning process through this interaction. Adding to that, he considered the unsupervised learning as an automated ML (aML) approach, since there is no human interaction with the learning process, although he mentioned that, in an unsupervised learning task, the human expert can verify the results of the algorithms at the end of the ML-pipeline. This is exactly where gamification can participate in verifying the results and assisting the optimization of the algorithm, however, to the best of our knowledge, there have not been any studies in the literature deploying this potential. Another situation is where the gamification is used to help human intervention to label parts of the data which turns an unsupervised learning problem into another type of ML problem called semi-supervised learning. However, this is also a topic that is not addressed by the research community as well.

4.4.3 RQ3. What are the Reported Effects of Using Gamification Over Machine Learning?

To answer this research question, we investigated the studies which attempted to use gamification in order to enhance a machine learning process. There have been several studies in this regard, that are discussed in this section.

(S7) is an attempt to create a video recommender system in which authors used sentiment analysis of the user comments about every video (Mulholland et al. 2015). Comments

have been under a machine learning method using Google API to carry out a sentiment classification for video recommendation. In this context, gamification aspects, namely, badges and leaderboards have been used to encourage user interaction with the system. Authors reported that using gamification, facilitated the encouragement of user interaction, facilitating in turn the Machine learning sentiment analysis task. In the same context, (S10) reported an optimized level of engagement in an attempt to develop an affect recognition system (Ghaleb et al. 2018). They used machine learning SVM algorithm in combination with a serious game to identify and classify users' affections. Another study reporting the same effect of gamification over machine learning is (S16), where a higher engagement rate through gamification resulted in a higher detection rate of student engagement recognition in prosocial games (Psaltis et al. 2018).

Moreover, the effect of gamification in improving the prediction accuracy of the machine learning method is reported by some of the studies. For example, (S9) reported such an improvement in the application of predicting student profiles (Barata et al. 2016). They used gamification aspects such as points and leaderboards to make early student profile detection possible, which improves the accuracy of machine learning algorithms. In a different scenario, (S41) introduced a solution for people with deaf blindness that have combined issues with vision and hearing senses (Korn et al. 2018). They have developed a system that uses machine learning for object and face recognition and environmental sensing, which are then provided to the users by means of haptic communication through smart textiles. At the same time, the proposed solution is empowered by gamification techniques to be more engaging and joyful for people with these severe conditions. Here, gamification and machine learning are applied as two separate technologies to provide a beneficial solution to some specific target users.

To conclude, it can be seen in the literature that the usage of gamification provides the required data for later analysis with the help of machine learning. In fact, it is the process of gamification that made machine learning analysis possible. The same is reported by Murphy et al. (2018).

4.4.4 RQ4. Which Aspects of Gamification are Affected by Machine Learning?

This research question is in fact, contrary to the previous one and investigates the contribution of machine learning in enhancing gamification processes. There have been few studies reporting and considering the effect of machine learning in gamification that are discussed in this section.

In general, the common issue with gamification processes is the fact that one game does not fit all. This is where various users have different capabilities, preferences, potentials, and characteristics, and they require the gamification design to be configured for them accordingly. This is the case when gamification is to be used in learning and education applications. Different students have different capabilities and characteristics, making it difficult to have a gamified task that serves all of them equally. Therefore, providing a one-time configured gamified application may result in inefficient interactions with the users, making it a false effort. In this regard, the main approach that can be taken into consideration is to personalize and adapt the gamified application with respect to the user's characteristics which requires the application to understand and learn the characteristics of its users.

One of the significant influences of applying machine learning in the gamification context is to personalize the gamified tasks for the users, as well as adapting gamification with regards to the user capabilities and domain context that enhances gamification performance

to a great extent. (S4) reported an improvement in the adaptation of the gamification task with the help of machine learning (Lopez and Tucker 2018). Authors used machine learning SVM algorithm to help adapting game features and task difficulties based on user performances. They argue that since previous methods of adapting gamification tasks have not been able to obtain the user's behavior before the end of the interaction, the level of personalization is limited. Therefore, they analyze the interactions of the user dynamically and try to predict future performance with the help of machine learning. As a result, authors also mentioned that this approach can help to understand the relationship between the factors affecting user performance and gamified tasks. The authors of the mentioned study, collected the data of the users' facial expression to analyze their affective state and to predict their performances, based on which the game elements should be modified. However, the sole usage of the facial expression data may be insufficient for predicting user affective state and given the limitation of users data, it is unverifiable that into what extent the proposed technique is effective. The distinctive feature of their work is that they used the affective state indirectly to modify the game design based on user performance prediction, compared to studies that altered the game elements based on the user affective state directly. Anyhow, machine learning is shown to be useful for personalizing the game design.

Therefore, the same goal can also be achieved with the help of other techniques such as identifying the student type or their affective states. Studies reporting the same personalizing effect with these techniques are presented in (S20) and (S10) by Barata et al. (2015) and Ghaleb et al. (2018). Authors in (S10) developed a gamified engineering course to create a smarter learning environment. Moreover, they utilized the machine learning clustering task to personalize gamification by clustering students based on their specific performances. They reported that the usage of gamification proved rich responses fitted to the characteristics of every student. The study of (S10) is a well reported example for how machine learning can be used to classify students based on their performances in early stages of their interaction with a gamified learning platform and consequently adapting the game features based on that classification. The study shows that how the grouping of students based on their first year performances led to making the game components more rewarding for those who were showing less engagement in the first year, and hence the percentage of those dropped in the second year. Hence increasing the engagement of the students with the gamified platform can be achieved with the help of the machine learning.

(S20) on the other hand, used affect recognition with the help of machine learning to improve the adaptiveness parameter of a gamified task in a learning environment. In general, the goal of personalizing the game elements in a gamification strategy for learning and education purposes can be achieved with the help of machine learning. (S20) presented the application of a serious game in delivering learning contents to students. Serious games or games with purpose that employ game design elements for interacting with students are shown to be an effective solution in previous studies. Given that the user reaction to a serious game can be fluctuating based on user attributes, it is constructive to identify the user's feedback in different stages of the interaction and personalize it accordingly. The effort of this study toward identifying the three user's affective states of frustration, boredom, and engagement with the help of machine learning helps the task of personalization to be achieved effectively.

Other studies reported some specific improvements in gamified tasks with the help of machine learning as a result of their work. (S12) believes that machine learning can facilitate game content insertion during the gamification process, which can be seen as an adaptation context too (Lima et al. 2016). Authors in this study used machine learning to insert real data into the gamified platform, hence, optimizing the gamified learning experience. They proposed a method called intelligent agents to interfere in the gamified learning

activity in order to get the virtual learning environment closer to reality. (S22) presented a serious game where users are motivated to use stairs instead of elevators to increase users' physical activity (Ciman et al. 2016). They used machine learning to recognize whether users are taking stairs or not. They believe machine learning makes the process of gamification to work interactively in real-time. Therefore, one of the applications of machine learning in improvement of gamified platforms is to provide an advantageous functionality to the system. For example, the study of (S12) used gamification to classify diseases based on their symptoms in a medical application. Generally, machine learning has the ability of making the gamified platform to be intelligent and updated by feeding new data to the system continuously. This data can be inserted into the system either by a human agent or by the gamified platform itself. This is the same approach that is also used in (S22) to make the gamified platform changing its behavior dynamically and in real-time.

4.4.5 RQ5. Benefits and Challenges in Connection with Machine Learning and Gamification?

To answer the final research question, we investigated the benefits presented by the papers in connection with converging machine learning and gamification to achieve a particular task. In what follows we first provide the benefits of using machine learning for the gamification community, followed by the benefits of gamification for the machine learning community. Finally we outline some of the identified challenges regarding the combining of machine learning and gamification as a whole.

In particular, machine learning is said to help adapt game features and task difficulties as in (S4), as well as, providing early student profile detection to enhance gamification tasks as in (S9) by Barata et al. (2016) and Lopez and Tucker (2018). Adapting game features based on interaction progress and system goals helps to improve the gamified tasks which are only made possible with the help of machine learning. In other words, this helps the continuation of the user interaction with the platform that is called engagement, as well as updating interaction contents to be more realistic and beneficial. Consequently, the predicting power of machine learning can be used to analyze both the environment and the user data to predict some future events in a gamified platform and adapt the game context accordingly.

Gamification elements, such as points and badges can be used to increase the engagement level of users in a particular application. However, other attributes such as users' specific situations and their behaviors can also be positively correlated to the level of engagement. Authors in (S35) proposed a machine learning approach to collectively consider both user attributes and gaming measures to evaluate the level of engagement in an application of online emotional support system (Doran et al. 2015). They trained a random forest algorithm to predict the level of engagement over time, both for old and new members. In this study, gamification and machine learning worked cooperatively to measure the level of engagement of users in a crowdsourcing application. In general, early detection of user's future engagement level can be helpful to prevent users from leaving the gamified application. This can be achieved with the help of machine learning, by which user's activity level in the future is predicted as a supervised learning problem. The study in (S35) have deployed this application in an emotional support platform and reported positive results. They also found that gamification elements are positively correlated to the user engagement level. But, it is also required to verify how generalizable is this approach to other applications, such as a gamified learning platform that its users require a different level of motivation for continuing their engagement with the system.

Another benefit of using machine learning in gamification tasks is to automate the process of personalization, without additional overwhelming effort by operators. The same is reported in (S19), (Knutas et al. 2018). The personalization of game features based on user capabilities and preferences has always been significant to improve user experience. This is because game features are normally designed and developed once and before the exposure of the system to the real environment when much of the specific user potentials, limits, and behaviors are unknown. Therefore, it is difficult to design a system that is responsive enough to the needs of all the users.

On the other hand, there are benefits of using gamification over machine learning practices, such as improved user interaction, facilitating the machine learning-based learning materials, and higher quality data collection are reported as in (S7), (S11) and (S15), respectively (Mulholland et al. 2015; Suresh et al. 2018; Li et al. 2016). (S7) used leaderboards and badges to motivate individuals to interact with the system in order to collect their sentiments. Collected sentiments are later used for analysis in making an online recommender system. Here, the fact that gamification can add to the engagement of users is to be used in the context of machine learning-based systems that require user interactions. This is the case with every human-in-the-loop application, where human cooperatively work with the machine to either optimize the data collection task for a supervised learning problem or to verify the results of an unsupervised or reinforcement learning problem. In general, given that gamification technology basically deals with optimizing the human interaction with the machine, there is a potential to use it in any machine learning application that requires human intervention to achieve some tasks. This is where the concepts of interactive machine learning (iML) is introduced (Holzinger 2016).

Moreover, sometimes the quality of the data collected through user interaction with the system becomes important. In other words, gamification can enhance the focus of the user while interacting with the system and higher quality data can be collected especially when the data is to be collected from the user itself, such as affect recognition, as provided in (S15). In this scenario, the gamification is used to make the participating users in the data collection task to be more careful about their progress, hence enhancing the quality of the collected data. This data collection can be both a labeling task for a supervised problem and feature data collection for an unsupervised learning problem. This is a great benefit for machine learning community, given that the amount of available data is limited in many applications, such as in the case of medical applications.

In a different perspective, (S11) provides a learning platform that uses machine learning to provide dynamic content to users. The authors of this study have combined their platform with gamification to improve the overall features of the system. Gamification has been deployed to assess the performance of the user that can be used for the evaluation of the machine learning-based contents. Content dynamicity is an important feature for any learning platforms. Different students have different learning capabilities. Providing the right content to each individual student based on their interaction with a gamified learning platform improves their level of engagement, hence, improving the learning process as well.

Regarding the challenges, the papers were inspected for possible limitations of combining these technologies. There is a very limited number of challenges reported by the researchers, as they are normally focusing on the strengths of their work in the reports. (S9) argued that the constraints imposed by the limited datasets are a significant challenge since big datasets are a crucial requirement for machine learning-based gamified platforms (Barata et al. 2016). In general, most of the applications of machine learning techniques in gamified platforms, were attempts toward formulating the problem as a supervised learning, which requires a labeled target variable. Although, considering them as other types of ML problems such as semi-supervised learning can be beneficial by utilizing the available

limited amount of data. For example, in the case of clustering the users based on their performances, presented in (S10).

On the other hand, the data may be biased as mentioned in (S16), which can result in unreliable outcomes (Psaltis et al. 2018). Thus, this is important to make sure of the quality of the data. A biased dataset can result in misleading inferences and the quality of the collected data should be taken into consideration prior to any decisions. However, it is a challenging task to ensure the quality of data that is collected through a gamified human interaction.

Last but not least, (S19) reflects that the process of deploying machine learning to personalize gamification tasks requires ML experts (Knutas et al. 2018), which is a known challenge regarding ML developments in every field. Although, in recent years development of cloud based machine learning tools and services, provided the ability to deploy ML methods and techniques much easier and faster.

5 Conclusion

In recent years, gamification technology has been used to widely optimize human–computer interactions. On the other hand, the advancement of machine learning approaches results in the expansion of the studies around the prospective potential of combining these two technologies. In this study, we investigated the convergence of gamification and machine learning using a systematic literature review, to inspect every aspect of this phenomenon under scrutiny.

In this line, we formulated 5 different research questions to explore the usage of machine learning in a gamified application and vice versa, their effects on each other and the benefits and challenges regarding the combination of these two technologies. As a result, various applications of machine learning to enhance the performance of gamification have been identified. We have classified our findings in this regard into 3 different categories, being, learning and educational related works, personalization of the gamification tasks, and behavioral change efforts.

Within these categories, learning and education related application is the most used area of focus. One of the main applications of gamification has always been in the learning and education field, however, the problem with gamified tasks is that they normally do not stay in the interests of users in long-term. In this regard, there have been many attempts by the research community to overcome this issue with the help of machine learning and several solutions have been found. One of the common techniques is to personalize gamified learning platforms. This personalization is carried in different ways. One of which is to capture users' interaction with the system to understand their capabilities and preferences and then adapting the gamified tasks based on that. This is also possible to detect the affect state of the learners during the gameplay scenarios and modify the game experience respectively.

Moreover, there is the possibility of predicting future user behavior and using this prediction to provide suggestions to users as a type of motivating element. Machine learning can also help a learning system to cluster students based on the capabilities and performances so that the gamified task specific to their needs can be provided to them. At the same time, it is required to assess and evaluate the level of user engagements with the system, which is also possible with the help of machine learning and has been proved by various studies. Another important aspect of gamified learning platforms is to provide context-specific gamified contents to the users. Studies in the literature have shown that it is achievable to make the gamified applications intelligently choosing context-specific content with the help of machine learning. Last but not least, is where gamification and machine learning advantages are combined to achieve some specific goals such as adjusting the game content. For example,

increasing the chances of players attaining the game's specific learning objectives regarding their prosocial skills or detecting left-hand finger selection and performance in real-time.

On the other hand, the use of gamification to enhance machine learning workflows has been recognized in some of the studies, for instance, data collection for machine learning by gamification technology and teaching machine learning with the help of gamification.

Regarding the benefits of the convergence of machine learning and gamification, it has been revealed that machine learning can improve the dynamicity and intelligence of the gamification tasks by adding the adaptability and personalization features to the gamified elements. However, the limited availability of the required data for machine learning methods, and the complexity of developing reliable machine learning methods remain important challenges of this process.

Although previous studies related to the usage of machine learning and gamification technologies investigated many possibilities of cooperatively using these two technologies, there are many cases that are not yet explored. In this regards, many of the efforts have been recognized to be in the task of gamified data collection for the purpose of a supervised machine learning problem, however, gamification has the potential to assist with other types of ML problems, namely, unsupervised and reinforcement learning. Both of which, can benefit from a directed human intervention to verify the working of the respective algorithm. To be specific, the task of clustering is mostly an unsupervised learning problem where the results of a clustering algorithm, such as K-nearest-neighbor (KNN), can be coupled with the human agent to verify the results of clustering. Here, a gamified task can be used to control the human interaction with the machine learning algorithm to be in the right direction. In a different scenario, reinforcement learning, is the task of learning by try and error correction. Again, a gamified application can be used for utilizing the human knowledge to correct the errors and optimizing the generated ML model. This is very much useful in the case of medical image analysis, where human expertise is required to identify symptoms and diseases from medical images.

In this study, we have identified the widely used application of these two technologies in different scenarios, however, the increasing employment of these technologies for learning and educational purposes opens a new era to investigate closely as future work. Moreover, we found that in order to quantify the extent to which gamified learning platforms can benefit from machine learning based techniques, it is required to identify some reliable measures based on which the convergence of machine learning and gamification can be evaluated. This would be another prospective future work that should be investigated.

Acknowledgements Open Access funding provided by Ostfold University College.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Appendix A

See Table 7.

Table 7 Overview of the selected studies

ID	References	Title	Focus	RQ1	RQ2	RQ3	RQ4	RQ5	Relevancy
S1	Baughman et al. (2014)	Deepqa Jeopardy! gamification: a machine-learning perspective	Business	0	0	0	1	0	0.2
S2	L'Heureux et al. (2017)	A gamification framework for sensor data analytics	Crowdsourcing, IoT	0	1	0	1	0	0.4
S3	Sakulueakulsuk et al. (2018)	Kids making AI: integrating machine learning, gamification, and social context in STEM education	Learning	0	1	0	1	0	0.6
S4	Lopez and Tucker (2018)	Towards personalized adaptive gamification: a machine learning model for predicting performance	Personalization, performance improvement	1	0	1	0	1	0.6
S5	Rattadilok et al. (2018)	Teaching students about machine learning through a gamified approach	Learning	0	1	0	0	1	0.4
S6	Anderson et al. (2014)	An extensible online environment for teaching data science concepts through gamification	Learning	0	1	0	0	0	0.2
S7	Mulholland et al. (2015)	360-MAM-affect: sentiment analysis with the Google prediction API and EmoSentieNet	Sentiment analysis, recommender system	0	0	0	1	1	0.4
S8	Murphy et al. (2018)	Quantifying cognitive aging and performance with at-home gamified mobile EEG	Health	0	1	0	1	1	0.6
S9	Barata et al. (2016)	Early prediction of student profiles based on performance and gaming preferences	Learning	0	1	0	1	1	0.6
S10	Ghaleb et al. (2018)	Towards affect recognition through interactions with learning materials	Learning, personalization	1	1	1	1	0	0.8
S11	Suresh et al. (2018)	Gamification of a visual question answer system	Learning	1	0	0	0	1	0.4
S12	Lima et al. (2016)	A 3D serious game for medical students training in clinical cases	Learning	1	0	1	0	0	0.4

Table 7 (continued)

ID	References	Title	Focus	RQ1	RQ2	RQ3	RQ4	RQ5	Relevancy
S13	Falah et al. (2018)	A quantitative approach to design special purpose systems to measure hacking skills	Learning	0	1	0	1	1	0.6
S14	Tserstou et al. (2017)	SCENT: citizen sourced data in support of environmental monitoring	Crowdsourcing, IoT	1	0	1	0	1	0.6
S15	Li et al. (2016)	Towards an "in-the-wild" emotion dataset using a game-based framework	Crowdsourcing, affective computing	1	1	0	1	1	0.8
S16	Psaltis et al. (2018)	Multimodal student engagement recognition in prosocial games	Learning, engagement	1	1	0	1	1	0.8
S17	Akasiadis et al. (2015)	Incentives for rescheduling residential electricity consumption to promote renewable energy usage	Energy, behavioral change	1	0	1	0	0	0.4
S18	Roth and Kesdoğan (2018)	A privacy enhanced crowdsourcing architecture for road information mining using smartphones	Crowdsourcing, IoT, engagement	0	1	0	0	1	0.4
S19	Knutas et al. (2018)	A process for designing algorithm-based personalized gamification	Learning, personalization	1	0	0	0	1	0.4
S20	Barata et al. (2015)	Gamification for smarter learning: tales from the trenches	Learning, personalization, engagement	1	0	1	0	0	0.4
S21	Stefanidis et al. (2019)	Learning prosocial skills through multi-adaptive games: a case study	Learning, personalization, engagement	1	0	0	0	1	0.4
S22	Ciman et al. (2016)	Stairstep recognition and counting in a serious game for increasing users' physical activity	Behavioral change, activity recognition	1	0	1	0	0	0.4
S23	Kontadakis et al. (2018)	Gamified platform for rehabilitation after total knee replacement surgery employing low cost and portable inertial measurement sensor node	Engagement, rehabilitation, medical	1	0	0	0	0	0.2

Table 7 (continued)

ID	References	Title	Focus	RQ1	RQ2	RQ3	RQ4	RQ5	Relevancy
S24	Goswami et al. (2019)	Smartphone-based delivery of oropharyngeal exercises for treatment of snoring: a randomized controlled trial	Behavioral change, medical	1	0	0	0	0	0.2
S25	Ortiz-Catalan et al. (2016)	Phantom motor execution facilitated by machine learning and augmented reality as treatment for phantom limb pain: a single group, clinical trial in patients with chronic intractable phantom limb pain	Medical	0	0	0	0	0	0
S26	Konstantakopoulos et al. (2019)	A deep learning and gamification approach to improving human-building interaction and energy efficiency in smart infrastructure	Energy, behavioral change, IoT	1	0	1	0	0	0.4
S27	Xu et al. (2016)	Understanding the impact of personality traits on mobile app adoption—Insights from a large-scale field study	Affective computing	1	0	0	0	1	0.4
S28	Silva and Analide (2019)	Computational sustainability and the PHESS platform: Using affective computing as social indicators	Affective computing, IoT	0	0	0	0	0	0
S29	Di Lena et al. (2017)	In-vehicle human machine interface: an approach to enhance eco-driving behaviors	Personalization, behavioral change	1	1	0	0	0	0.4
S30	Khoshkangini et al. (2017)	Machine learning for personalized challenges in a gamified sustainable mobility scenario	Personalization, recommender systems	1	0	0	0	0	0.2
S31	Dalmazzo and Ramirez (2017)	Air violin: a machine learning approach to fingering gesture recognition	Learning	1	1	0	0	0	0.4
S32	Di Nunzio et al. (2018)	A gamified approach to NaïVe Bayes classification: a case study for newswires and systematic medical reviews	Crowdsourcing, eHealth, learning, engagement	0	1	0	0	0	0.2

Table 7 (continued)

ID	References	Title	Focus	RQ1	RQ2	RQ3	RQ4	RQ5	Relevancy
S33	Karaliopoulos et al. (2016)	First learn then earn: optimizing mobile crowdsensing campaigns through data-driven user profiling	Personalization	1	0	0	0	0	0.2
S34	Urh and Pejović (2016)	TaskyApp: inferring task engagement via smartphone sensing	Crowdsourcing	0	1	0	1	0	0.4
S35	Doran et al. (2015)	Stay A while and listen: user interactions in a crowdsourced platform offering emotional support	Crowdsourcing	0	0	0	0	1	0.2
S36	Palavalli et al. (2014)	Analyzing gaming-simulations using video based techniques	Learning	1	0	1	0	0	0.4
S37	Khajah et al. (2016)	Designing engaging games using bayesian optimization	Other	1	0	1	0	0	0.4
S38	Anparasanesan et al. (2019)	Smart monitor for tracking child's brain development	Learning	1	0	0	0	0	0.2
S39	Raptis et al. (2018)	DARA: assisting drivers to reflect on how they hold the steering wheel	Behavioral change	1	0	0	0	0	0.2
S40	Acharya et al. (2019)	Gamification of wearable data collection: a tool for both friend and foe	Other	0	1	0	0	0	0.2
S43	Lungu (2016)	Bootstrapping an ubiquitous monitoring ecosystem for accelerating vocabulary acquisition	Personalization	1	0	0	0	0	0.2

Appendix B

See Table 8.

Table 8 Quality of the papers checklist

ID	References	Bias		Validity		Generalizability		Study quality hierarchy (CRD study type)
		Does the study choose the subjects under study randomly?	Are the outcomes of the study interpreted not based on the subjects under study?	Is the study carried out with a scientific methodology?	Are the methods used well-defined and verifiable?	Is there a proper use-case to test the results?	Are the results general enough to be expandable to other situations?	
S1	Baughman et al. (2014)	–	–	Yes	Yes	–	Yes	5
S2	L'Heureux et al. (2017)	No	Yes	Yes	Yes	Yes	Yes	3-1
S3	Sakulkeakulsuk et al. (2018)	Yes	Yes	Yes	No	Yes	Yes	4-2
S4	Lopez and Tucker (2018)	No	Yes	Yes	Yes	Yes	Yes	2
S5	Rattadilok et al. (2018)	–	–	Yes	Yes	No	No	–
S6	Anderson et al. (2014)	No	Yes	Yes	Yes	Yes	Yes	5
S7	Mulholland et al. (2015)	–	–	Yes	No	No	No	–
S8	Murphy et al. (2018)	Yes	Yes	Yes	Yes	Yes	Yes	1
S9	Barata et al. (2016)	No	Yes	Yes	Yes	Yes	Yes	3-1
S10	Ghaleb et al. (2018)	No	Yes	Yes	Yes	Yes	Yes	2

Table 8 (continued)

ID	References	Bias		Validity		Generalizability		Study quality hierarchy (CRD study type)
		Does the study choose the subjects under study randomly?	Are the outcomes of the study interpreted not based on the subjects under study?	Is the study carried out with a scientific methodology?	Are the methods used well-defined and verifiable?	Is there a proper use-case to test the results?	Are the results general enough to be expandable to other situations?	
S11	Suresh et al. (2018)	-	-	Yes	No	No	No	-
S12	Lima et al. (2016)	-	-	Yes	Yes	No	Yes	-
S13	Falah et al. (2018)	No	Yes	Yes	Yes	Yes	Yes	2
S14	Tserstou et al. (2017)	-	-	-	-	No	-	-
S15	Li et al. (2016)	Yes	Yes	Yes	Yes	Yes	Yes	3-1
S16	Psaltis et al. (2018)	Yes	Yes	Yes	Yes	Yes	Yes	1
S17	Akasiadis et al. (2015)	Yes	Yes	Yes	Yes	Yes	Yes	4-1, 1
S18	Roth and Kesdoğan (2018)	No	Yes	Yes	Yes	Yes	No	4-3
S19	Knutas et al. (2018)	-	-	Yes	Yes	-	Yes	5
S20	Barata et al. (2015)	No	Yes	Yes	Yes	Yes	Yes	2
S21	Stefanidis et al. (2019)	No	Yes	Yes	Yes	Yes	Yes	2
S22	Ciman et al. (2016)	No	Yes	Yes	Yes	Yes	Yes	2
S23	Kontadakis et al. (2018)	Yes	Yes	Yes	Yes	Yes	Yes	1
S24	Goswami et al. (2019)	Yes	Yes	Yes	Yes	Yes	Yes	1

Table 8 (continued)

ID	References	Bias	Validity		Generalizability		Study quality hierarchy (CRD study type)	
			Does the study choose the subjects under study randomly?	Are the outcomes of the study interpreted not based on the subjects under study?	Is the study carried out with a scientific methodology?	Are the methods used well-defined and verifiable?		Is there a proper use-case to test the results?
S25	Ortiz-Catalan et al. (2016)	No	Yes	No	No	Yes	Yes	2
S26	Konstantakopoulos et al. (2019)	Yes	Yes	No	Yes	Yes	Yes	1
S27	Xu et al. (2016)	Yes	Yes	Yes	Yes	Yes	Yes	2
S28	Silva and Analide (2019)	Yes	Yes	No	Yes	Yes	No	1
S29	Di Lena et al. (2017)	-	-	Yes	No	No	No	-
S30	Khoshkangini et al. (2017)	Yes	Yes	No	Yes	Yes	No	1
S31	Dalmazzo and Ramirez (2017)	-	-	Yes	No	No	No	-
S32	Di Nunzio et al. (2018)	-	Yes	No	Yes	Yes	Yes	1
S33	Karaliopoulos et al. (2016)	No	Yes	Yes	Yes	Yes	Yes	1
S34	Urh and Pejović (2016)	No	Yes	Yes	No	No	No	4-3
S35	Doran et al. (2015)	Yes	Yes	Yes	Yes	Yes	Yes	1
S36	Palavalli et al. (2014)	No	Yes	Yes	No	No	Yes	4-3

Table 8 (continued)

ID	References	Bias Does the study choose the subjects under study randomly?	Validity		Generalizability		Study quality hierarchy (CRD study type)
			Are the outcomes of the study interpreted not based on the subjects under study?	Is the study carried out with a scientific methodology?	Are the methods used well-defined and verifiable?	Is there a proper use-case to test the results?	
S37	Khajah et al. (2016)	No	Yes	Yes	Yes	Yes	4-3
S38	Anparasanesan et al. (2019)	No	Yes	Yes	No	No	4-3
S39	Raptis et al. (2018)	No	Yes	Yes	Yes	Yes	4-3
S40	Acharya et al. (2019)	No	Yes	Yes	Yes	No	4-3
S41	Korn et al. (2018)	-	Yes	Yes	No	No	5
S42	Schäfer et al. (2018)	No	Yes	Yes	Yes	Yes	4-2
S43	Lungu (2016)	-	Yes	Yes	No	No	5

References

- Abu Saa, A., Al-Emran, M., & Shaalan, K. (2019). Factors affecting students' performance in higher education: A systematic review of predictive data mining techniques. *Technology, Knowledge and Learning*, 24, 567–598. <https://doi.org/10.1007/s10758-019-09408-7>.
- Acharya, S., Matovu, R., Serwadda, A. & Griswold-Steiner, I. (2019). Gamification of wearable data collection: A tool for both friend and foe. In *Proceedings of the 2019 3rd international conference on compute and data analysis* (pp. 68–77). ICCDA 2019. Kahului, HI, USA: Association for Computing Machinery. <https://doi.org/10.1145/3314545.3314572>.
- Ahn, S. J., Johnsen, K., & Ball, C. (2019). Points-based reward systems in gamification impact children's physical activity strategies and psychological needs. *Health Education & Behavior*. <https://doi.org/10.1177/1090198118818241>.
- Akasiadis, C., Panagidi, K., Panagiotou, N., Sernani, P., Morton, A., Vetsikas, I. A., Mavrouli, L., & Goutsias, K. (2015). Incentives for rescheduling residential electricity consumption to promote renewable energy usage. In *2015 SAI intelligent systems conference (IntelliSys)* (pp. 328–337). <https://doi.org/10.1109/IntelliSys.2015.7361163>.
- Anderson, P. E., Turner, C., Dierksheide, J. & McCauley, R. (2014). An extensible online environment for teaching data science concepts through gamification. In *2014 IEEE frontiers in education conference (FIE) proceedings* (pp. 1–8). <https://doi.org/10.1109/FIE.2014.7044205>.
- Anparasanesan, T., Mathangi, K., Kobikanth, S., Seyon, S. & Anjalie Gamage, M. P. A. W. (2019). Smart monitor for tracking child's brain development. In *Proceedings of the 2019 8th international conference on educational and information technology* (pp. 68–72). ICEIT 2019. Cambridge, United Kingdom: Association for Computing Machinery. <https://doi.org/10.1145/3318396.3318401>.
- Barata, G., Gama, S., Jorge, J., & Gonçalves, D. (2015). Gamification for smarter learning: Tales from the trenches. *Smart Learning Environments*, 2, 10. <https://doi.org/10.1186/s40561-015-0017-8>.
- Barata, G., Gama, S., Jorge, J., & Gonçalves, D. (2016). Early prediction of student profiles based on performance and gaming preferences. *IEEE Transactions on Learning Technologies*, 9, 272–284. <https://doi.org/10.1109/TLT.2016.2541664>.
- Baughman, A. K., Chuang, W., Dixon, K. R., Benz, Z., & Basilico, J. (2014). Deepqa jeopardy! gamification: A machine-learning perspective. *IEEE Transactions on Computational Intelligence and AI in Games*, 6, 55–66. <https://doi.org/10.1109/TCIAIG.2013.2285651>.
- Brereton, P., Kitchenham, B. A., Budgen, D., Turner, M., & Khalil, M. (2007). Lessons from applying the systematic literature review process within the software engineering domain. *Journal of Systems and Software*, 80, 571–583. <https://doi.org/10.1016/j.jss.2006.07.009>.
- Caporarello, L., Magni, M. & Pennarola, F. (2019). One game does not fit all. gamification and learning: Overview and future directions. In A. Lazazzara, R. C. D. Nacamulli, C. Rossignoli, & S. Za (Eds.), *Organizing for digital innovation* (pp. 179–188). Lecture notes in information systems and organisation. Springer International Publishing.
- Ciman, M., Donini, M., Gaggi, O., & Aiulli, F. (2016). Stairstep recognition and counting in a serious Game for increasing users' physical activity. *Personal and Ubiquitous Computing*, 20, 1015–1033. <https://doi.org/10.1007/s00779-016-0968-y>.
- Cochrane handbook for systematic reviews of interventions. 2019. <http://handbook-5-1.cochrane.org/>. Accessed Mar 6.
- Codish, D., & Ravid, G. (2015). Detecting playfulness in educational gamification through behavior patterns. *IBM Journal of Research and Development*, 59: 6: 1-6:14. <https://doi.org/10.1147/JRD.2015.2459651>.
- Dalmazzo, D., & Ramirez, R. (2017). Air violin: A machine learning approach to fingering gesture recognition. In *Proceedings of the 1st ACM SIGCHI international workshop on multimodal interaction for education* (pp. 63–66). MIE 2017. New York, NY, USA: ACM. <https://doi.org/10.1145/3139513.3139526>.
- de Sousa Borges, S., Durelli, V. H. S., Reis, H. M. & Isotani, S. (2014). A systematic mapping on gamification applied to education. In *Proceedings of the 29th annual ACM symposium on applied computing* (pp. 216–222). SAC'14. New York, NY, USA: ACM. <https://doi.org/10.1145/2554850.2554956>.
- Deterding, S., Sicart, M., Nacke, L., O'Hara, K. & Dixon, D. (2011). Gamification. using game-design elements in non-gaming contexts. In *CHI'11 extended abstracts on human factors in computing systems* (pp. 2425–2428). CHI EA'11. New York, NY, USA: ACM. <https://doi.org/10.1145/1979742.1979575>.
- Di Lena, P., Mirri, S., Prandi, C., Salomoni, P. & Delnevo, G. (2017). In-vehicle human machine interface: An approach to enhance eco-driving behaviors. In *Proceedings of the 2017 ACM workshop on*

- interacting with smart objects* (pp. 7–12). SmartObject'17. New York, NY, USA: ACM. <https://doi.org/10.1145/3038450.3038455>.
- Di Nunzio, GM, Maistro, M & Vezzani, F. (2018). A gamified approach to Naïve Bayes classification: A case study for newswires and systematic medical reviews. In *Companion proceedings of the the web conference 2018* (pp. 1139–1146). WWW'18. Republic and Canton of Geneva, Switzerland: International World Wide Web Conferences Steering Committee. <https://doi.org/10.1145/3184558.3191547>.
- Doran, D., Yelne, S., Massari, L., Calzarossa, M.-C., Jackson, L. T. & Moriarty, G. (2015). Stay awhile and listen: User interactions in a crowdsourced platform offering emotional support. In *Proceedings of the 2015 IEEE/ACM international conference on advances in social networks analysis and mining 2015* (pp. 667–674). ASONAM'15. Paris, France: Association for Computing Machinery. <https://doi.org/10.1145/2808797.2809311>.
- Falah, A., Pan, L. & Chen, F. (2018). A quantitative approach to design special purpose systems to measure hacking skills. In *2018 IEEE international conference on teaching, assessment, and learning for engineering (TALE)* (pp. 54–61). <https://doi.org/10.1109/TALE.2018.8615431>.
- Gartner. (2012). *Gamification 2020: What is the future of gamification?*. Stamford: Gartner.
- Gartner. (2019). *Gartner survey shows 37 percent of organizations have implemented AI in some form*. Stamford: Gartner.
- Ghaleb, E., Popa, M., Hortal, E., Asteriadis, S. & Weiss, G. (2018). Towards affect recognition through interactions with learning materials. In *2018 17th IEEE international conference on machine learning and applications (ICMLA)* (pp. 372–379). <https://doi.org/10.1109/ICMLA.2018.00062>.
- Goswami, U., Black, A., Krohn, B., Meyers, W., & Iber, C. (2019). Smartphone-based delivery of oropharyngeal exercises for treatment of snoring: A randomized controlled trial. *Sleep and Breathing*, 23, 243–250. <https://doi.org/10.1007/s11325-018-1690-y>.
- Herranz, E., Colomo-Palacios, R., & de Amescua Seco, A. (2015). Gamiware: A gamification platform for software process improvement. In R. V. O'Connor, M. U. Akkaya, K. Kemaneci, M. Yilmaz, A. Poth, & R. Messnarz (Eds.), *Systems, software and services process improvement* (pp. 127–139)., Communications in computer and information science Springer: Springer.
- Herranz, E., Palacios, R. C., de Amescua Seco, A., & Yilmaz, M. (2014). Gamification as a disruptive factor in software process improvement initiatives. *Journal of Universal Computer Science*, 20, 885–906.
- Holzinger, A. (2013). Human-computer interaction and knowledge discovery (HCI-KDD): What is the benefit of bringing those two fields to work together? In *International conference on availability, reliability, and security* (pp. 319–328). Springer.
- Holzinger, A. (2016). Interactive machine learning for health informatics: When do we need the human-in-the-loop? *Brain Informatics*, 3, 119–131. <https://doi.org/10.1007/s40708-016-0042-6>.
- Kappen, D. L., & Orji, R. (2017). Gamified and persuasive systems as behavior change agents for health and wellness. *XRDS*, 24, 52–55. <https://doi.org/10.1145/3123750>.
- Karaliopoulos, M., Koutsopoulos, I. & Titsias, M. (2016). First learn then earn: Optimizing mobile crowdsensing campaigns through data-driven user profiling. In *Proceedings of the 17th ACM international symposium on mobile ad hoc networking and computing* (pp. 271–280). MobiHoc'16. Paderborn, Germany: Association for Computing Machinery. <https://doi.org/10.1145/2942358.2942369>.
- Khajah, M. M., Roads, B. D., Lindsey, R. V., Liu, Y.-E., & Mozer, M. C. (2016). Designing engaging games using bayesian optimization. In *Proceedings of the 2016 CHI conference on human factors in computing systems* (pp. 5571–5582). CHI'16. San Jose, California, USA: Association for Computing Machinery. <https://doi.org/10.1145/2858036.2858253>.
- Khoshkingini, R., Marconi, A., & Valetto, G. (2017). Machine learning for personalized challenges in a gamified sustainable mobility scenario. In *Extended abstracts publication of the annual symposium on computer-human interaction in play* (pp. 361–368). CHI PLAY'17 extended abstracts. New York, NY, USA: ACM. <https://doi.org/10.1145/3130859.3131321>.
- Kitchenham, B. (2004). Procedures for performing systematic reviews. *Keele, UK, Keele University*, 33(2004), 1–26.
- Kitchenham, B., Pearl Brereton, O., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering: A systematic literature review. In *Information and software technology. Special section—most cited articles in 2002 and regular research papers* (Vol. 51, pp. 7–15). <https://doi.org/10.1016/j.infsof.2008.09.009>.
- Knutas, A., van Roy, R., Hynninen, T., Granato, M., Kasurinen, J., & Ikonen, J. (2017). Profile-based algorithm for personalized gamification in computer-supported collaborative learning environments. In *GHITALY@ CHIItaly*.
- Knutas, A., van Roy, R., Hynninen, T., Granato, M., Kasurinen, J., & Ikonen, J. (2018). A process for designing algorithm-based personalized gamification. *Multimedia Tools and Applications*. <https://doi.org/10.1007/s11042-018-6913-5>.

- Konstantakopoulos, I. C., Barkan, A. R., He, S., Veeravalli, T., Liu, H., & Spanos, C. (2019). A deep learning and gamification approach to improving human-building interaction and energy efficiency in smart infrastructure. *Applied Energy*, 237, 810–821. <https://doi.org/10.1016/j.apenergy.2018.12.065>.
- Kontadakis, G., Chasiouras, D., Proimaki, D., Halkiadakis, M., Fyntikaki, M., & Mania, K. (2018). Gamified platform for rehabilitation after total knee replacement surgery employing low cost and portable inertial measurement sensor node. *Multimedia Tools and Applications*. <https://doi.org/10.1007/s11042-018-6572-6>.
- Korn, O., Holt, R., Kontopoulos, E., Kappers, A. M. L. Persson, N.-K. & Olson, N. (2018). Empowering persons with deafblindness: Designing an intelligent assistive wearable in the SUITCEYES project. In *Proceedings of the 11th PErvasive technologies related to assistive environments conference* (pp. 545–551). PETRA'18. Corfu, Greece: Association for Computing Machinery. <https://doi.org/10.1145/3197768.3201541>.
- Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling*. New York: Springer. <https://doi.org/10.1007/978-1-4614-6849-3>.
- L'Heureux, A., Grolinger, K., Higashino, W. A. & Capretz, M. A. M. (2017). A gamification framework for sensor data analytics. In *2017 IEEE international congress on internet of things (ICIOT)*, 74–81. <https://doi.org/10.1109/IEEE.ICIOT.2017.18>.
- Lantz, B. (2015). *Machine learning with R* (2nd ed.). Birmingham: Packt Publishing.
- Li, W., Abtahi, F., Tsangouri, C. & Zhu, Z. (2016). Towards an “In-the-wild” emotion dataset using a game-based framework. In *2016 IEEE conference on computer vision and pattern recognition workshops (CVPRW)* (pp. 1526–1534). <https://doi.org/10.1109/CVPRW.2016.190>.
- Lopez, C., & Tucker, C. (2018). Towards personalized adaptive gamification: A machine learning model for predicting performance. *IEEE Transactions on Games*. <https://doi.org/10.1109/TG.2018.2883661>.
- Lungu, M. F. (2016). Bootstrapping an ubiquitous monitoring ecosystem for accelerating vocabulary acquisition. In *Proceedings of the 10th European conference on software architecture workshops* (pp. 1–4). ECSAW'16. Copenhagen: Association for Computing Machinery. <https://doi.org/10.1145/2993412.3003389>.
- Michalski, R. S., Carbonell, J. G., & Mitchell, T. M. (1983). *Machine learning: An artificial intelligence approach*. Berlin: Springer, Berlin Heidelberg.
- Monterrat, B., Lavoué, E., & George, S. (2014). Motivation for Learning: Adaptive Gamification for Web-based Learning Environments. 6th International Conference on Computer Supported Education (CSEDU 2014) (pp. 117–125). <https://hal.archives-ouvertes.fr/hal-01130791>.
- Monterrat, B., Lavoué, É., & George, S. (2015). Toward an adaptive gamification system for learning environments. In S. Zvacek, M. T. Restivo, J. Uhoimobhi, & M. Helfert, (Eds.), *Computer supported education* (pp. 115–129). Communications in computer and information science Berlin: Springer International Publishing.
- Mulholland, E., Mc Kevitt, P., Lunney, T., Farren, J., & Wilson, J. (2015). 360-MAM-Affect: Sentiment analysis with the Google prediction API and EmoSenticNet. In *2015 7th international conference on intelligent technologies for interactive entertainment (INTETAIN)* (pp. 217–221).
- Murphy, B., Aleni, A., Belaoucha, B., Dyer, J., & Nolan, H. (2018). Quantifying cognitive aging and performance with at-home gamified mobile EEG. In *2018 international workshop on pattern recognition in neuroimaging (PRNI)* (pp. 1–4). <https://doi.org/10.1109/PRNI.2018.8423954>.
- Nastis, S. A., & Pagoni, E. G. (2019). Gamification of global climate change: An experimental analysis. *Studies in Nonlinear Dynamics & Econometrics*. <https://doi.org/10.1515/snde-2017-0105>.
- Ortiz-Catalan, M., Guðmundsdóttir, R. A., Kristoffersen, M. B., Zepeda-Echavarria, A., Caine-Winterberger, K., Kulbacka-Ortiz, K., et al. (2016). Phantom motor execution facilitated by machine learning and augmented reality as treatment for phantom limb pain: A single group, clinical trial in patients with chronic intractable phantom limb pain. *The Lancet*, 388, 2885–2894. [https://doi.org/10.1016/S0140-6736\(16\)31598-7](https://doi.org/10.1016/S0140-6736(16)31598-7).
- Palavalli, B. M., Krishna, H., & Jayagopi, D. (2014). Analyzing gaming-simulations using video based techniques. In *Proceedings of the second international conference on technological ecosystems for enhancing multiculturalism* (pp. 235–239). TEEM'14. Salamanca, Spain: Association for Computing Machinery. <https://doi.org/10.1145/2669711.2669905>.
- Petersen, K., & Gencel, C. (2013). Worldviews, research methods, and their relationship to validity in empirical software engineering research. In *2013 joint conference of the 23rd international workshop on software measurement and the 8th international conference on software process and product measurement* (pp. 81–89). <https://doi.org/10.1109/IWSM-Mensura.2013.22>.
- Psaltis, A., Apostolakis, K. C., Dimitropoulos, K., & Daras, P. (2018). Multimodal student engagement recognition in prosocial games. *IEEE Transactions on Games*, 10, 292–303. <https://doi.org/10.1109/TCAIG.2017.2743341>.

- Lima, R. de, M., de Medeiros Santos, A., Neto, F. M. M., de Sousa Neto, A. F., Leão, F. C. P., de Macedo, F. T. & de Paula Canuto, A. M. (2016). A 3D serious game for medical students training in clinical cases. In *2016 IEEE international conference on serious games and applications for health (SeGAH)*, 1–9. <https://doi.org/10.1109/SeGAH.2016.7586255>.
- Raptis, D., Iversen, J., Mølbak, T. H. & Skov, M. B. (2018). DARA: Assisting drivers to reflect on how they hold the steering wheel. In *Proceedings of the 10th Nordic conference on human-computer interaction* (pp. 1–12). NordiCHI'18. Oslo, Norway: Association for Computing Machinery. <https://doi.org/10.1145/3240167.3240186>.
- Rattadilok, P., Roadknight, C. & Li, L. (2018). Teaching students about machine learning through a gamified approach. In *2018 IEEE international conference on teaching, assessment, and learning for engineering (TALE)* (pp. 1011–1015). <https://doi.org/10.1109/TALE.2018.8615279>.
- Roth, C., & Kesdoğan, D. (2018). A privacy enhanced crowdsourcing architecture for road information mining using smartphones. In *2018 IEEE 11th conference on service-oriented computing and applications (SOCA)* (pp. 17–24). <https://doi.org/10.1109/SOCA.2018.00010>.
- Sakulkueakulsuk, B., Witoon, S., Ngarmkajornwiwat, P., Pataranutaporn, P., Surareungchai, W., Pataranutaporn, P. & Subsoontorn, P. (2018). Kids making AI: Integrating machine learning, gamification, and social context in STEM education. In *2018 IEEE international conference on teaching, assessment, and learning for engineering (TALE)* (pp. 1005–1010). <https://doi.org/10.1109/TALE.2018.8615249>.
- Sánchez-Gordón, M., & Colomo-Palacios, R. (2019). Taking the emotional pulse of software engineering: A systematic literature review of empirical studies. *Information and Software Technology*, 115, 23–43. <https://doi.org/10.1016/j.infsof.2019.08.002>.
- Schäfer, H., Bachner, J., Pretscher, S., Groh, G & Demetriou, Y. (2018). Study on motivating physical activity in children with personalized gamified feedback. In *Adjunct publication of the 26th conference on user modeling, adaptation and personalization* (pp. 221–226). UMAP'18. Singapore, Singapore: Association for Computing Machinery. <https://doi.org/10.1145/3213586.3225227>.
- Seufert, S., Meier, C., Soellner, M., & Rietsche, R. (2019). A pedagogical perspective on big data and learning analytics: A conceptual model for digital learning support. *Technology, Knowledge and Learning*, 24, 599–619. <https://doi.org/10.1007/s10758-019-09399-5>.
- Sharp, P. A., Cooney, C. L., Kastner, M. A., Lees, J., Sasisekharan, R., Yaffe, M. B. & Sur, M. (2011). The third revolution: The convergence of the life sciences. *Physical Sciences, and Engineering. [White paper on convergence, Massachusetts Institute of Technology (MIT), Washington, DC]*.
- Silva, F., & Analide, C. (2019). Computational sustainability and the PHESS platform: Using affective computing as social indicators. *Future Generation Computer Systems*, 92, 329–341. <https://doi.org/10.1016/j.future.2018.10.006>.
- Sin, K., & Muthu, L. (2015). Application of big data in education data mining and learning analytics: A literature review. *ICTACT Journal on Soft Computing*, 5(4), 1035–1049.
- Stefanidis, K., Psaltis, A., Apostolakis, K. C., Dimitropoulos, K., & Daras, P. (2019). Learning prosocial skills through multiadaptive games: A case study. *Journal of Computers in Education*, 6, 167–190. <https://doi.org/10.1007/s40692-019-00134-8>.
- Suresh, S., Nagaraj Rao, V., & Srinivasa, G. (2018). Gamification of a visual question answer system. In *2018 IEEE tenth international conference on technology for education (T4E)* (pp. 41–44). <https://doi.org/10.1109/T4E.2018.00016>.
- Teasley, S. D. (2017). Student facing dashboards: One size fits all? *Technology, Knowledge and Learning*, 22, 377–384. <https://doi.org/10.1007/s10758-017-9314-3>.
- Tserstou, A., Jonoski, A., Popescu, I., Asumpcao, T. H., Athanasiou, G., Kallioras, A. & Nichersu, I. (2017). SCENT: Citizen sourced data in support of environmental monitoring. In *2017 21st international conference on control systems and computer science (CSCS)* (pp. 612–616). <https://doi.org/10.1109/CSCS.2017.93>.
- Urh, G., & Pejović, V. (2016). TaskyApp: Inferring task engagement via smartphone sensing. In *Proceedings of the 2016 ACM international joint conference on pervasive and ubiquitous computing: Adjunct* (pp. 1548–1553). UbiComp'16. Heidelberg, Germany: Association for Computing Machinery. <https://doi.org/10.1145/2968219.2968547>.
- Xi, N., & Hamari, J. (2019). Does gamification satisfy needs? A study on the relationship between gamification features and intrinsic need satisfaction. *International Journal of Information Management*, 46, 210–221. <https://doi.org/10.1016/j.ijinfomgt.2018.12.002>.
- Xu, R., Frey, R. M., Fleisch, E., & Ilic, A. (2016). Understanding the impact of personality traits on mobile app adoption: Insights from a large-scale field study. *Computers in Human Behavior*, 62, 244–256. <https://doi.org/10.1016/j.chb.2016.04.011>.

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

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