Int. J. Human–Computer Studies 161 (2022) 102774

Contents lists available at ScienceDirect

International Journal of Human - Computer Studies

journal homepage: www.elsevier.com/locate/ijhcs

Evaluating a gamification proposal for learning usability heuristics: Heureka

Raimel Sobrino-Duque, Noelia Martínez-Rojo, Juan Manuel Carrillo-de-Gea, Juan José López-Jiménez, Joaquín Nicolás [*], José Luis Fernández-Alemán

Faculty of Computer Science, Department of Informatics and System, University of Murcia, Spain

ARTICLE INFO

Keywords:
Usability
Nielsen heuristics
Gamification
Student experiment

ABSTRACT

This paper presents the results of an educational experiment conducted to determine whether an automated, card-based gamification strategy has an impact on the learning of Jakob Nielsen's 10 heuristic usability rules. The participants in the experiment were 55 students enrolled on a human-computer interaction course. According to the results of the experiment and the hypothesis tests performed to compare both traditional and gamified approaches, there were no significant differences (t (53) = 0.66, $p = 0.52$), although the scores attained by the students who used the gamification strategy were slightly better when evaluated one week later ($M = 6.29$ and $M = 6.57$ out of 10, respectively). Moreover, the students' perceptions reflect that the proposed tool is easy to use (MD = 4.00 out of 5) and useful as regards learning (MD = 4.00 out of 5). Further research is needed to determine whether incorporating other gamification elements, such as rankings, difficulty levels, and game modes, would have a positive impact on student motivation, engagement and performance.

1. Introduction

Usability is a product attribute that influences the quality of a software system. According to ISO 9241 (2018), usability is "the extent to which a system, product, or service can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use". Usability is important because it addresses pragmatic aspects of a product related to behavioral goals that the software must achieve (Guerino and Valentim, 2020).

Educators are continuously seeking ways in which to innovate the learning process in order to decrease the dropout rate and, in general, make this process more effective in terms of a better understanding of the subjects and consequently increase pass rates (Sharma and Sharma, 2021). One means employed to achieve this objective is that of adapting traditional learning methods to new pedagogical theories. These new approaches include the application of digital strategies by taking advantage of emerging technologies, such as the Internet, multimedia, and lately social networks and video games (Crittenden et al., 2019). Digital strategies concern not only technologies, but also the ways in which devices and software are used in order to enrich learning, whether inside or outside of the classroom. Moreover, these strategies can be used in both formal and informal learning, and can transform conventional learning ideas and activities into something new and meaningful (Johnson et al., 2016).

The concept of gamification originated from the digital media industry (Rodrigues et al., 2019). Several attempts have been made to define gamification. Barber (2021) defines gamification as "the application of gaming elements to non-gaming contexts", while Huotari and Hamari (2012) define it as "a process of enhancing a service with affordances for gameful experiences in order to support user's overall value creation." Other definitions include "the use of game thinking and game mechanics to engage users and solve problems" (Dale, 2014). Gamification has already been successfully used in education, marketing, organizational, health and environmental initiatives to achieve desirable outcomes by influencing user behavior (Bai et al., 2020). Interest in applying gamification to education is increasing, given its capacity to capture and sustain students' attention, which is a prerequisite for students' success in educational environments (Khalil et al., 2018).

A review of an academic bibliography carried out by (Carrion et al., 2019) showed that the terms gamification and serious game share characteristics, although they are different. A serious game is defined as "any kind of interactive computer application that incorporates gamification principles and serves an educational purpose, or aims to achieve a predefined goal" (Meijer et al., 2021). It is widely recognized that

* Corresponding author at: Faculty of Computer Science, Campus of Espinardo, Murcia, Spain.
E-mail address: jnr@um.es (J. Nicolás).

https://doi.org/10.1016/j.ijhcs.2022.102774
Received 13 April 2021; Received in revised form 11 January 2022; Accepted 18 January 2022
Available online 25 January 2022

serious games provide numerous benefits, including improved student participation, timely assessment and feedback, and ultimately, the achievement of learning outcomes (Bai et al., 2020). Games of this nature have been proposed for many Computer Science domains, including Computer fundamentals (Sindre et al., 2009), Programming (Muratet et al., 2011); (Haaranen et al., 2014); (Hakulinen et al., 2013); (O'Donovan et al., 2013), Operating systems (Hill et al., 2003), Information systems and Computer engineering (Barata et al., 2013), Information and computer technologies (Domínguez et al., 2013), Mathematics (Gordon et al., 2013), Computer organization and Cloud computing (E. Rodrigues et al., 2007), Software engineering, and Design & Usability (Labrador and Villegas, 2014).

The aim of this paper is to determine the impact of the use of gamification on the learning process of a specific well-known usability asset, specifically Jakob Nielsen's 10 usability heuristic rules (Nielsen, 2020). In this vein, the specific contributions of this paper are the following: (1) a systematic gamification process based on rigorous learning theories that underpins the game design for learning usability heuristics; (2) a user-centered, self-built tool called Heureka to support that learning process; and (3) an empirical evaluation of both process and tool, involving students enrolled on the User Interfaces course, a Human-Computer Interaction (HCI) subject on the Degree in Computer Engineering at the University of OMITTED FOR REVIEW (OMITTED FOR REVIEW).

The remainder of this paper is organized as follows. Section 2 introduces the background to the theory of gamification, while Section 3 reviews previous work on gamification experiences for usability learning. Section 4 introduces the process used to gamify Heureka, a web application whose intention is to allow users to practice the 10 heuristics for user interface design by Jakob Nielsen. The experiment methodology is presented in Section 5, and Section 6 shows the results of the statistical analysis carried out on the data obtained. Section 7 discusses the study findings and lessons learned considering the key objectives of the experiment carried out, while threats to validity are highlighted in Section 8. Finally, Section 9 provides some concluding remarks and shows an outline of future work.

2. Background

2.1. Gamification and serious games

One important element in the present-day learning process is the use of Information and Communication Technologies (ICT) owing to their acceptance by both students and teachers. Of the various ICT tools for education, there are resources with which to create work environments, or to share files, interactive games, and others (Carrion et al., 2019). Serious games are, along with gamification, currently growing among interactive games, and will likely continue to do so because of four factors: (1) the number of domains in which serious games are used is growing; (2) the serious games industry is closely related to the video game industry, the latter of which is in a state of permanent growth; (3) videogames are now part of our culture, and (4) serious games are increasingly open to the teaching of adults (Gounaridou et al., 2021).

Game elements can be defined as elements that are characteristic of games (Deterding et al., 2011). The game elements deliver information to the players and usually appear in forms of a user interface. Game elements can be classified on various levels of abstraction. A systematic mapping found 27 game elements distributed throughout 43 serious games (Dos Santos et al., 2018). These game elements were classified in three categories: dynamics, components, and mechanics. There were four elements in the dynamics group, of which Fantasy was the most frequently used element (17 games in total), while there were nine elements in the components group. The most frequently used elements in that group were Level (36 games), Quest (16 games) and Avatar (14 games). Lastly, 14 elements were collected in the mechanics group, in which Goal (21 games) and Point System (16 games) were the most frequently used elements.

Researchers have defined game mechanics in various ways, and there is no consensus as to a unified taxonomy, thus making it difficult to distinguish how game mechanics work analytically speaking (Schell, 2019). Game mechanics help build a narrative with which to keep users curious and make them look forward to the evolution of the gamified environment as they get involved with activities. Several game mechanics have been proposed, such as points, badges, levels, challenges, leaderboards, on-boarding, social engagement loops, and feedback (Zichermann and Cunningham, 2011).

A gamification strategy is defined as using game-based mechanics, aesthetics, and game thinking to engage people, motivate action, promote learning, and solve problems (Kapp, 2012). A gamification strategy is an integrated effort that provokes the user's proactive participation and action by applying game mechanics and elements to the non-game field (Kim, 2021). Previous studies have presented numerous gamification strategies. However, many of them consider neither the users nor the context, or are duplicates because of the inconsistent definitions suggested by several researchers (Kim, 2021). Several gamification strategies have been reported in literature (Zichermann and Cunningham, 2011): (1) Competition, (2) Relationship, (3) Challenge, (4) Compensation, (5) Achievement, (6) Self-expression, and (7) Usability.

2.2. Gamification-motivation theories

Despite the fact that gamification has been employed in different academic disciplines (Bai et al., 2020), motivational theories with which to define and analyze the foundations of gamification are underdeveloped, as claimed in a literature survey by Seaborn and Fels (2015). Some gamification frameworks have been defined in order to ease the application of these theories into the development of both rigorous and engaging gamification approaches. Seaborn and Fels (2015) evidence the existence of several theories and frameworks. The theories underpinning the gamification frameworks are the following: (1) Self-Determination Theory, (2) Intrinsic and Extrinsic Motivation, (3) Situational Relevance, (4) Situated Motivational Affordance, (5) Transtheoretical Model of Behavior Change, (6) Universal Design for Learning, and (7) User-Centered Design.

- The Self-Determination Theory, which specifically concerns the concepts of autonomy, competence and relatedness (Ryan and Deci, 2000), is the basis of the framework defined by Aparicio et al. (2012). This framework is divided into the following four steps: (1) identification of the main objective, stating the reasons for the use of gamification; (2) identification of the cross-cutting objective, defining what intrinsically motivating factors the system is intended to provide; (3) determination of the game mechanics, determining which game mechanics will be used on the basis of their relationship with the self-determination concepts, and finally, (4) the evaluation of the framework in the system applied, indicating how to evaluate the framework in those systems.
- Intrinsic and extrinsic motivation theories (Blohm and Leimeister 2013) (Nicholson 2012) support user-centered frameworks for the so-called meaningful gamification, i.e., gamification based on intrinsic (or internal) motivation rather than extrinsic (or external) motivation. These user-centered frameworks aim to elucidate how gamification can operate on intrinsic and extrinsic motivators to elicit behavioral change and reframe activities such as learning. These frameworks specifically suggest that meaningful game elements are intrinsically motivating, regardless of the external rewards that may be associated with them. In this vein, a value-based gamification framework for designers who aim to foster and leverage intrinsic motivation was defined (Sakamoto et al., 2012). The five values that comprise the framework are: (1) information, as necessary and immediately available; (2) empathic values, drawn from

virtual characters and social involvement; (3) persuasive values, that is, information that provides a forward-looking perspective based on behaviors, actions, and outcomes; (4) economic values, related to charging and ownership; and (5) ideological values, defined as implicit beliefs through storytelling and a variety of message types. This framework is designed to be complementary to other gamification frameworks.

- The Theory of Situational Relevance (Wilson, 1973) is based on the importance that an individual places on a particular situation, although on the situation as s/he perceives it, and not as other people perceive it. Situational Relevance frameworks (Nicholson, 2012) imply that the user should make decisions concerning what is meaningful. The Theory of Situational Relevance should be circumscribed to the details of the situation that affect the individual.
- The Theory of Situated Motivational Affordance (Deterding, 2011) states that motivation is afforded when the relationship between the features of an object and the ability of a subject allows the subject to experience the satisfaction of such needs when interacting with the object. People perform activities if those activities promise to satisfy their motivational needs, such as competence, autonomy, or relatedness. Nicholson (2012) introduces a framework, stressing the need for a correspondence between the user's background and the game setting. In this framework, affordances (e.g., perceived opportunities for action on the elements of the user interface) are mapped onto motivational needs drawn from satisfaction theories of motivation (specifically from the Self-Determination Theory).
- The Transtheoretical Model of Behavior Change (TTM) (Sakamoto et al., 2012) is a theory of intentional change that focuses on the decision-making abilities of the individual rather than social and biological influences on behavior. This model grew from the systematic integration of more than 300 theories of psychotherapy, along with the analysis of the leading theories of behavior change.
- The Theory of Universal Design for Learning (UDL) addresses the need to provide educational resources for a wide spectrum of users, taking into account presentation, activity types and learning paths (Rose and Meyer, 2007).
- The User-Centered Design Theory (Norman, 1989) is a design philosophy that places the user at the center of the experience, and designs iteratively with the user's needs and desires in mind. User feedback is thus essential to specify and refine requirements and designs. Communication between the user and the system is also important, as long as the user should be aware at all times of the system status and the actions that can be taken.

In conclusion, all gamification frameworks share a core consisting of motivational theories, behavioral change and engagement (Kim, 2021). Furthermore, User-Centered Design is common to all theories (Nicholson, 2012), as it places the user in the foreground of the gamification strategy and can serve as a basis on which to combine other motivation theories.

3. Related work

In a systematic literature review, Wangenheim and Shull (2009) refer to 12 games, in which computer-based simulations led the list of the most used game type for learning. A breakdown of the studies by topics and learning domains reveals that most of them were developed in order to teach software project management skills, but usability is one of the disciplines for which less empirical evidence was found. The usability life cycle, the heuristic evaluation in the software projects development, the evaluation of the usability of a software system and the use of various techniques to address usability in systems are the main usability education related issues addressed by the scientific community (Barreto et al., 2015).

Carrol and Rosson (2006) used case-based learning as an instructional resource for the teaching of usability engineering. They analyzed the proposal that cases can be a minimalist-information design technique, that is, a design technique that (1) focuses on information in order to facilitate user action; (2) anchors information in the activity; (3) prevents, mitigates and leverages errors; and (4) develops user autonomy. These authors discuss a case study in which students performed all the steps related to the usability lifecycle.

With regard to usability audits, Tao (2005) presents an approach with which to integrate usability evaluation into behavior modeling for interactive systems in order to help students introduce usability concepts from the early stages of software development. Furthermore, Wahl (2000) presents a set of steps in which students were separated into groups and asked to develop a library automation system and evaluate the usability of the software developed by another group. Furthermore, Ludi (2005) proposed a process using a hands-on approach that allowed students to apply various techniques to address usability in systems. Students followed a usability testing process that gave them the opportunity to plan the process and methodology, recruit participants, conduct tests, and analyze test results. Moreover, Faulker and Culwin (2001) used students to conduct a usability study involving 124 users who analyzed a set of web pages and answered a set of questions related to the usability of these pages.

Despite the widespread use of game-playing elements to teach computer literacy, only two serious games, called UsabilityGame (Barreto et al., 2015) and UsabilityCity (Ferreira et al., 2014), were found to teach Jakob Nielsen's 10 usability heuristic rules.

UsabilityGame is a web application designed to complement university teaching in the field of HCI and to explain two processes: (1) the life cycle of usability engineering and (2) Jakob Nielsen's heuristic evaluation from a procedural point of view. The game experience involves two types of roles: teacher and student. The student takes on the role of Usability Engineer, whereas the teacher plays the role of Usability Engineer Leader. The process of evaluation and monitoring of students is guided by the instructor through a specific environment.

UsabilityGame is structured in three stages: (1) Requirements Analysis (the players analyze all scenarios, and from this analysis, define the requirements); (2) Design, Testing and Development of a prototype (given the requirements specification document, the students prototype an interface that meets that specification with a level of fidelity (low, medium or high) to be configured by the teacher using a prototyping tool embedded in the game); and (3) Heuristic Evaluation (given the captures of real interfaces with the problem areas highlighted visually, the students choose from the list of Nielsen heuristics that are being violated in each case).

The elements of gamification that UsabilityGame uses are: (1) interactivity, since the player can move between the game screens and check boxes to indicate their answers using a prototyping tool within the game itself; (2) narrative, because it is developed in videos and images with text balloons and cartoon drawings; (3) levels, i.e., the division of content in phases, and the need to overcome one in order to access the next phase, and finally (4) points and rankings, since the game updates the player's score at the end of each level. The winner is the person who accumulates the most points at the end of the three stages.

UsabilityCity is a web application (available in Portuguese) whose objective is to allow HCI students to learn the Nielsen heuristic evaluation method. The game does not require the intervention of a teacher and is structured in 5 phases, each one for two heuristics. If the two exercises in each phase are correct, the game moves on to the next phase.

Some of the elements of gamification presented in UsabilityCity are interactivity, narrative, and levels. Firstly, (1) interactivity is supported, since the players can move between game screens and click onto characters in order to select their answers (each character represents a Nielsen heuristic). Secondly, (2) the narrative is shown by means of images with text balloons and cartoon drawings. The player takes on the role of an inspector (represented by a character with a magnifying glass) who must identify the problems of the so-called "UsabilityCity" in order to improve the lives of its residents (called Users). The heuristics themselves are personified so as to draw the student's attention. Finally, (3) the levels divide the presentation of the content into unlockable phases.

The consulted bibliography allowed us to conclude that there is little evidence on the use of gamification elements in usability learning. This indicates a need for empirical research related to this topic and more specifically to Jakob Nielsen's heuristic rules.

4. Implementing gamification in Heureka

Alhammad and Moreno (2018) state that there are no systematic approaches with which to gamify software engineering education. Most primary studies analyzed in their systematic mapping reported no formal or structured approach. Seaborn and Fels (2015) point out that 87% of applied gamification research is not based on any theoretical foundation. To avoid this issue, the gamification process proposed by Herzig et al. (2015) was followed in the development of Heureka. These authors believe that gamification can be understood as a software development process. User-Centered Design (Don Norman, 2013) was another of the key methodologies employed when developing Heureka. This process is characterized by the fact that the users play a central role and that the entire design revolves around their interests and needs. Moreover, Nicholson (2012) argues that User-Centered Design can act as a nexus between the theoretical models underlying gamification systems. In this vein, Heureka is based on three theories that have been previously applied in interactive systems gamification: (1) the Self-Determination Theory; (2) the User-Centered Design Theory; and (3) the Operant Conditioning psychological theory (Skinner, 1965). Self-Determination Theory and User-Centered Design provide formal support for the interplay between individual psychological needs and self-motivation, whereas Operant Conditioning is an associative learning process consisting of the development of new behaviors based on their positive or negative consequences.

The gamification process presented by Herzig et al. (2015) is organized into eight workflows comprising numerous tasks and roles. For the sake of simplicity, the four high-level phases shown in Calderón et al. (2018) are used instead to articulate the explanation: (1) Business Modeling and Requirements, in which the application context is analyzed and business goals are documented; (2) Design, in which the gamification design is developed and playtested; (3) Implementation, during which the design is implemented as software artifacts and functionally tested; and (4) Monitoring and Adaptation, during which business goal achievement is measured, and subsequent design adaptations are conducted.

4.1. Business modeling and requirements

As stated by Herzig et al. (2015), several stakeholders and roles participate in gamification at some point of the process: *gamification experts, domain experts, business experts, IT experts* and *end users*. The authors of the present work played all roles except that of end users. An agile software development philosophy steered project management activities in Heureka. A Kanban board was used to visualize the workflow and task progression, using cards that move along "swimlanes" (i.e., *pending, in progress, done*) during the software development process (see Fig. 1).

The first step of the gamification process is a key aspect as regards allowing all the stakeholders —with the exception of the end users—to share a common ground in the business processes and to identify the general project goals and end users, who are the final recipients of the gamification effort. In this case, the project goal is to help players learn Jakob Nielsen's 10 heuristic rules of usability (described in Table 1)—although the resulting gamified system should be easy to adapt to different contexts in usability learning—. The end users are students of a User Interfaces subject.

The 10 Usability Heuristics created by Jakob Nielsen for HCI design were proposed in 1990 and can be considered the gold standard when evaluating usability heuristics. These heuristics are rules of thumb that must be adapted to specific interface types. Other sets of usability heuristics have been proposed in literature, which include modifications to Nielsen's heuristics and/or the addition of new heuristics in order to optimally design or evaluate specific aspects of user interfaces not covered by other heuristics. According to one systematic literature review (Quiñones and Rusu, 2017), a total of 68 usability heuristics have been created for specific domains, signifying that Nielsen's heuristics is a must-have material on an HCI course. However, learning such general rules proves to be a real challenge. This learning requires showing students many examples, thus enabling them to relate concrete situations, and move from general guidelines to particular scenarios. These rules have traditionally been learned by using a variety of resources: text, figures and videos (see e.g. Harley (2019)). Heureka was designed to be used without the intervention of an instructor and to require as few resources from the user as possible. Moreover, its gamification approach aims to motivate students by taking into account their user profile: consumer, exploiter, achiever and free spirit user types (cf. SubSection 4.2).

The end users and their use cases were then analyzed to elicit and document requirements that are aligned with the project goals. At this point, the end users were involved, and the participation of all other roles except the IT expert was also necessary. Epics and user stories were used to construct the Heureka requirements specification. These user stories were distributed in iterations on the basis of their priority,

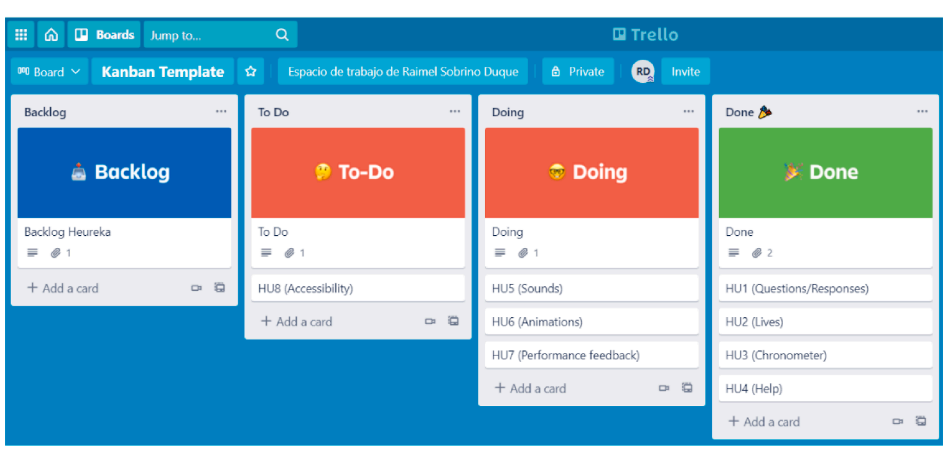
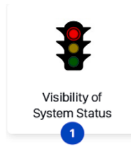
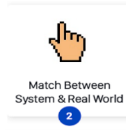
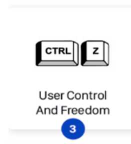
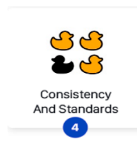
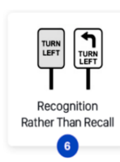
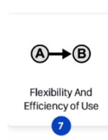
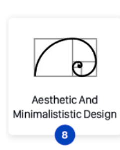
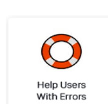
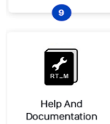


Fig. 1. An example of visualization of workflow and task progression in iteration 3 using the Kanban Board.

Table 1
Description of Jakob Nielsen's 10 Usability Heuristics taken from Nielsen (2020).

Heuristics	Description	
Visibility of system status	The design should always keep users informed about what is going on, through appropriate feedback within a reasonable amount of time.	
Match between system and the real world	The design should speak the users' language. Use words, phrases, and concepts familiar to the user, rather than internal jargon. Follow real-world conventions, making information appear in a natural and logical order.	
User control and freedom	Users often perform actions by mistake. They need a clearly marked "emergency exit" to leave the unwanted action without having to go through an extended process.	
Consistency and standards	Users should not have to wonder whether different words, situations or actions mean the same thing. Follow platform and industry conventions.	
Error prevention	Good error messages are important, but the best designs carefully prevent problems from occurring in the first place. Either eliminate error-prone conditions or check for them and present users with a confirmation option before they commit to the action.	
Recognition rather than recall	Minimize the user's memory load by making elements, actions, and options visible. The user should not have to remember information from one part of the interface to another. Information required to use the design (e.g., field labels or menu items) should be visible or easily retrievable when needed.	
Flexibility and efficiency of use	Shortcuts — hidden from novice users — may speed up the interaction for the expert user such that the design can cater to both inexperienced and experienced users. Allow users to tailor frequent actions.	
esthetic and minimalist design	Interfaces should not contain information which is irrelevant or rarely needed. Every extra unit of information in an interface competes with the relevant units of information and diminishes their relative visibility.	
Help users recognize, diagnose, and recover from errors	Error messages should be expressed in plain language (no error codes), precisely indicate the problem, and constructively suggest a solution.	
Help and documentation	It is best if the system doesn't need any additional explanation. However, it may be necessary to provide documentation to help users understand how to complete their tasks.	

simplicity and dependencies, in consonance with the user-centered design approach carried out in Heureka, in order to obtain incremental releases (a total of 8 user stories were defined).

4.2. Design

The outcome of this phase is the specification of a meaningful gamification design and requires all the stakeholders to agree on the adoption of a particular design alternative. It is worth noting that the implementation of gamification techniques arises as an important challenge during gamification (Sardi et al., 2017).While gamification does not necessarily require the implementation of costly tools and frameworks (Alhammad and Moreno, 2018), it has been pointed out that the lack of software resources and tools that help experts create gamification strategies is a major hindrance to the application of gamification elements in educational environments (Dicheva et al., 2015). The design of gamification strategies requires the support of resources that help gamification designers focus only on what is crucial for the success of the strategy. A small number of specific tools with which to support the process of gamification design currently exist, and although they are useful for designers, they provide only partial support as regards designing meaningful gamification strategies (Calderón et al., 2018).

Gamicards (Ferro, 2021) was used in the design of the gamification strategy proposed in this paper. The tool is a prototype deck of cards that aims to cover the fundamental and basic considerations for the development of more meaningful and personalized gamification strategies. Gamicards provides designers with five types of cards: (1) *Game elements and Mechanics*, (2) *User considerations*, (3) *Context*, (4) *Motivation* and (5) *Mystery Cards* (Ferro et al., 2014).

Firstly, the designers selected the User considerations cards in order to characterize the target audience by using the information extracted about the end users (i.e., male and female undergraduate students between the ages of 20 and 21). Secondly, the Context card was used to identify the educational environment in which the gamification strategy was going to take place (i.e., students enrolled in a 4th year, first-term User Interfaces subject on a BSc in Computer Engineering). The application of the Gamicards approach up to this point could be closely integrated with the activities in the Business Modeling and Requirements phase. Thirdly, the designers selected the Motivation card and determined that the reason for the design of this gamification strategy would be the evaluation and improvement of usability learning. Finally, the gamification elements and mechanics to be considered in the design were identified on the basis of the Game elements and Mechanics cards. In this regard, an exhaustive list of game elements and mechanics can be found in (Ferro, 2021). Heureka uses the following game elements: (1) *bars*; (2) *chance*; (3) *timer*; (4) *feedback*; (5) *permadeath*; (6) *quests*, and (7) *rewards*. It also includes the following game mechanics: (8) *celebrate*; (9) *punish*; and (10) *win*. All these items are detailed below:

1 Bars. The player has limited number of lives. An indicator —i.e., a lives bar—of the number of lives is included.
2 Chance. Each game is different, since the questions included in a specific quiz are randomly chosen from the pool of available questions: 11 for heuristic 1, 13 for heuristic 2, 3 for heuristic 3, 4 for heuristic 4, 9 for heuristic 5, 6 for heuristic 6, 8 for heuristic 7, 1 for heuristic 8, 7 for heuristic 9, and 5 for heuristic 10. This feature adds replay value to the experience.
3 Timer. A timer is used to monitor the time spent playing the game and to motivate the player to get better and better.
4 Feedback. The system makes use of aesthetics, animations and sound effects to provide the look and feel of a game. Animations and sound effects linked to key events are included, such as clicking on answers, losing lives or finishing the game.
5 Permadeath. The players lose if they make more than two mistakes. In this case, if the players wish to continue playing, they must start a new game from the beginning.

6 Quests. The most prominent game element in Heureka is the quiz (i.e., questions and answers); quizzes represent the challenges that the player is required to overcome. During a game, the player is presented with a set of 10 questions that must be answered by clicking onto the right alternative (see Fig. 2).
7 Rewards. The system provides positive feedback when the player performs well in the game.
8 Celebrate. There are both minor (i.e., correct answer) and major (i.e., a complete quiz) achievements in the game, which represent the outcomes that the player will celebrate.
9 Punish. Punishment is applied in two ways: (1) when a wrong answer to a question is provided, the system sends the players negative feedback; and (2) when the players lose all their lives, the game ends and the players must start again from the beginning.
10 Win. The players win when they complete the quiz without losing all their lives. Fig. 3 depicts a player's states and transitions in the game.

Of the variety of gamification strategies discussed in literature (Kim, 2021), the categories (1) *challenge*, (2) *compensation* and (3) *usability* are applicable to Heureka. As stated above, (1) the player must undertake a quest within Heureka, which consists of meeting a knowledge challenge presented in the context of the gamified system. The Self-Determination Theory emphasizes the existence of a link between motivation and behavior. The gamification of the system aims to promote the players' motivation, signifying that the players should be motivated by the gamified experience and consequently modify their behavior in order to carry out the planned learning task. Challenge strategies should allow users to carry out achievable tasks while maintaining their motivation (Kim, 2021). With regard to (2) compensation, the player's interest and satisfaction are stimulated by means of a reinforcement schema based on the Operant Conditioning theory. This simple schema facilitates learning by integrating positive (e.g., celebrate, rewards) and negative (e.g., punish, permadeath) game mechanics and elements according to the player's inputs. What is more, in the case of (3) usability, Heureka offers explanations about the usability heuristics if the player needs help. The usability strategy, therefore, enhances the player's adaptability, flexibility of use, and usage behavior (Kim, 2021).

The user type was also part of the design of the gamification strategy employed for Heureka. Marczewski's typology (Klock et al., 2018) was considered, as it describes players according to their motivations for using the gamified system. More specifically, the extended version of this typology was used (Herbert et al., 2014). As a result, the *player* (especially the *consumer* and *exploiter* subtypes), *achiever* and *free spirit* user types (Marczewski, 2015; Tondello et al., 2016) were found to fit the proposed gamification strategy. Most users who enter a gamified system do so initially for the extrinsic rewards they can attain. The rewards are, therefore, an important design element in Heureka that meet the players' user-type needs. The fundamental idea is to attempt to convert them from reward-oriented users to intrinsically motivated users, as defined by the self-determination theory. Furthermore, the achiever user type is motivated by mastery: the design elements implemented in Heureka to suit achievers are challenges, learning new skills and quests. Finally, the free spirits are motivated by autonomy, act outside the control of others and enjoy exploration. The exploratory tasks are enabled by the replayability of Heureka, and Heureka's design element is more focused on the requirements of this user type, together with the fact that the system is conceived to be used independently by players.

The design of the gamification strategy does not explicitly refer to the philanthropic and socializer player types, although elements of these player types are addressed implicitly. This is because Heureka was developed with the aim of supporting the learning of Jakob Nielsen's 10 heuristics, a specific content within the subject of User Interfaces, and elements of these player types are, therefore, addressed outside that context. For example, in User Interfaces classes, the emphasis is on the students being altruistic and helping each other without expecting a reward for it (philanthropic). Furthermore, in class, the students are intrinsically motivated by relationships, enjoying interacting with other students and creating social connections (socializers). These types of players are not mutually exclusive.

4.3. Implementation

Once a gamification strategy has been designed, the execution of several technical activities will crystallize into the final gamified system, namely: provisioning, implementation, testing and deployment. Although domain, gamification or business experts can help clarify or discuss the gamification concept, ICT experts are ultimately responsible for this workflow. There are two main options as regards implementing the gamification strategy: (1) using a general-purpose gamification platform, or (2) creating self-built solutions with which to support gamification.

General-purpose gamification platforms can simplify the implementation process at the cost of reduced flexibility and higher

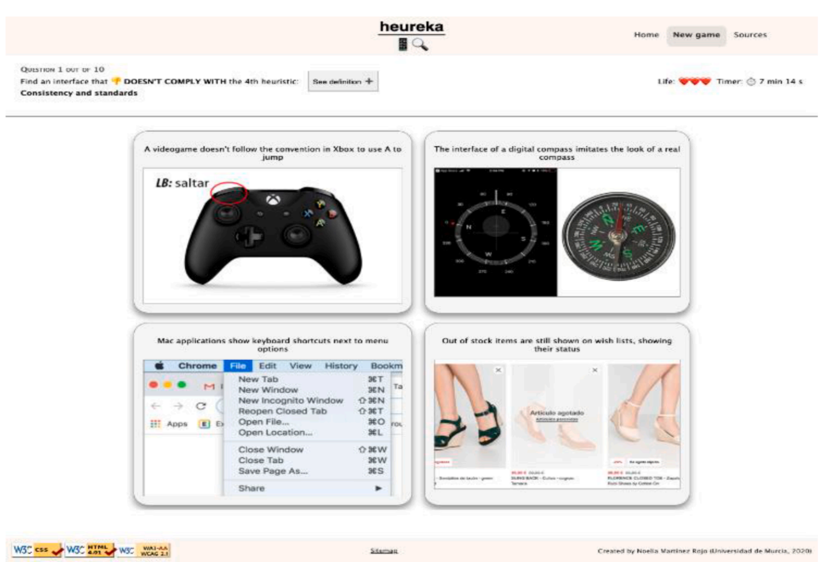


Fig. 2. Heureka game screen interface.

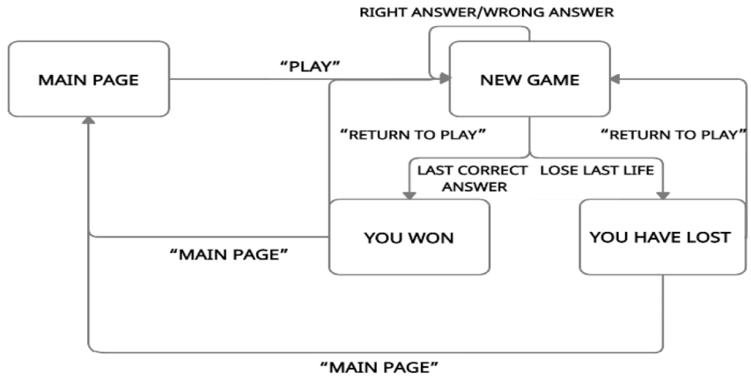


Fig. 3. Heureka interface navigation map, following the notation of Constantine & Lockwood (1999).

integration effort (Calderón et al., 2018). This could be the solution of choice when there is a lack of experience in gamification, gamification complexity is low, or knowledge or resources are insufficient to create a self-built solution. Heureka was conceived as a self-built gamified web application. This decision made it easier to design a tool that would be aligned with the business goals defined, ensure control over the gamification engine, and streamline the processing, controlling and monitoring of the user data generated (Herzig et al., 2012; Maican et al., 2016). This made it possible to modify, change and adjust all aspects of the system to our needs without being limited by the functionality of a generic gamification platform. Heureka has been developed with JavaScript, HTML and CSS. This choice was based on the criterion of design flexibility, and the learning curve is low. Git was used for version control, in addition to which Handlebars.js (an HTML template engine) and JSON were used as supporting technologies. The HTML code for the usability quizzes presented by the gamified tool was dynamically generated from a database in JSON format (see Fig. 4).

As the construction of the gamified system progressed, it was necessary to carry out the usual validation and verification (V&V) activities. In the tests performed on Heureka, it was possible to verify the requirements with the prototype in operation. This included the technical testing of functional correctness and non-functional attributes, along with verifying design constraints thanks to the cooperation of domain, gamification and business experts. With regard to non-functional attributes, particular attention was paid to accessibility, by testing WCAG 2.1 accessibility guidelines. Moreover, QUnit was used as a unit testing framework in JavaScript, Express.js (Node.js) for the server code, while Heroku was employed as a cloud computing service for temporary deployment during validation with end users. Once all the tests had been passed, the tool was deployed and its access granted to all end users (available only at https://docentis.inf.um.es:5050/).

4.4. Monitoring and adaptation

When the gamified system is running, the operational end user data is processed and analyzed in order to discover whether the system is successful and identify any possible modifications. In the case of Heureka, an empirical study was carried out as part of the monitoring tasks, given that the greatest challenge identified in gamification is the use of quantitative and qualitative data to obtain reliable information and adequately guide educators' decision making (Alhammad and Moreno, 2018). In this respect, a new functionality was added to Heureka that allowed the empirical study data of the experiment to be downloaded in JSON format once the users were interacting with the gamified tool. Examples of data that can be collected are the students' percentage of success in each of the heuristics and the average time they spent on each

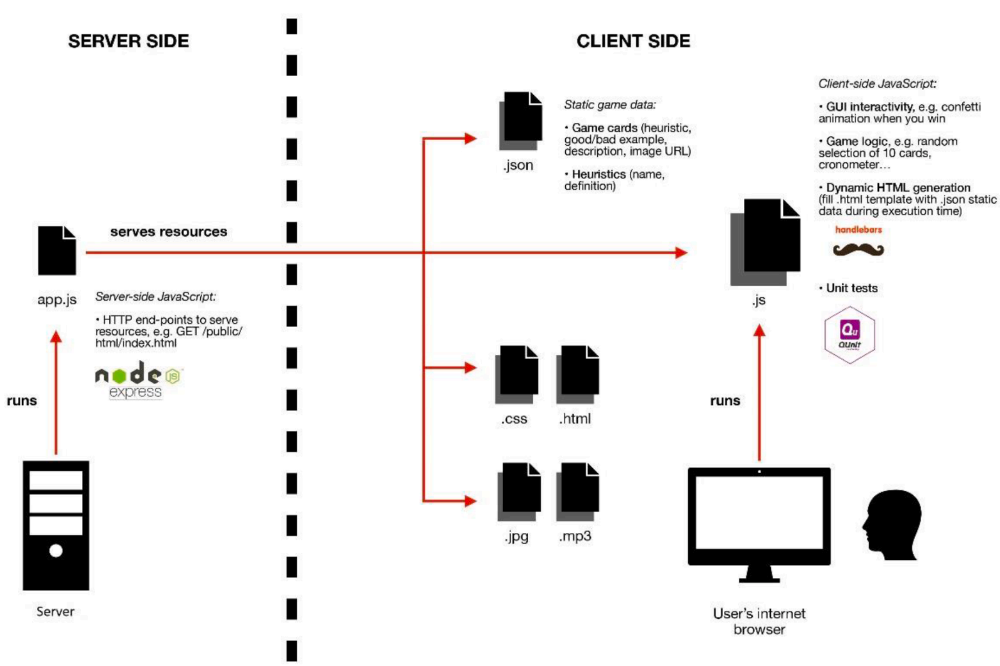


Fig. 4. Architecture diagram of Heureka.

heuristic. Detailed information on the procedure employed to design and conduct an experiment with which to evaluate our contribution to the use of gamification in usability learning is shown below.

5. Experimental methodology

5.1. Participants

In the academic year 2020/2021, a total of 66 students were enrolled on "User Interfaces", a 4th year, first-term course on the BSc in Computer Engineering at the University of OMITTED FOR REVIEW. Students have to complete a total of 6 ECTS credits, which are distributed into 2 h/week of lectures and 1 h 40 min/week of skills practice during a period of 15 weeks. The objective of this subject is to introduce students to HCI issues, including the development and auditing of usable and accessible user interfaces, paying special attention to the application of standards and style guides in web, mobile and desktop computer applications. The two key terms on this course are usability and accessibility.

The participants were recruited during the teaching of the subject, in which the teacher explained the purpose of the study, its duration and the activities to be carried out, and requested their verbal consent.

5.2. Design

One important aspect is the fact that, owing to the Covid-19 pandemic this year, all classes are being conducted via video conferencing (using Zoom, in our case), a software application that students have been familiar with from the beginning of the course. The experiment was, therefore, carried out remotely by means of video conference.

The students were given prior training in Jakob Nielsen's 10 rules of usability during the first 25 min of the experiment. They were then randomly assigned to two groups: (1) the control group (CG) and (2) the experimental group (EG). A total of 55 students attended the class and were divided into two groups by creating two Zoom rooms, i.e., each student was randomly assigned to one of two rooms. A total of 29 students participated in the CG experiment and 26 in the EG, thus representing a participation of 83.33% of the enrolled students.

The instructor provided the documentation on the Nielsen heuristics to the CG. The CG used a series of materials provided by the teacher that represented the traditional way of teaching Nielsen's principles. As Nielsen's heuristics are well-known and widespread, a lot of resources are, therefore, available on the Internet, and the materials provided to the students were basically a selection obtained from Internet sites. The CG students, therefore, spent 40 min using a set of links that contained information about the 10 Nielsen heuristics rules.

The EG received 10 minutes' training on the functioning of the Heureka tool, after which the group made use of Heureka by playing with the tool as much they could for 30 min. The EG interacted with Heureka through a link provided by a guidelines document. Once the link was opened, the students recorded a video showing their interaction with the game. At the end of the time given to them, and following the instructions given by the teachers, the students uploaded these videos to the Multimedia Gallery of the User Interfaces subject in the virtual classroom of the University of OMITTED FOR REVIEW.

Finally, students from both groups accessed the exam area of the virtual classroom in order to take a 20-minute test (Exam1). Both groups subsequently had the opportunity to continue practicing during the following week with either the traditional materials or with the Heureka tool, depending on whether they had participated in CG or EG. After one week, they repeated the exam in the virtual classroom (Exam2). On this occasion, the students in the experimentation group did not have to record videos of their interaction with Heureka.

5.3. Research goals

We shall present the objective of our empirical evaluation by following the recommendations of Basili and Dieter Rombach (1988), which are based on the application of the Goal/Question/Metric (GQM) method. Our goal is thus defined as follows:

- To analyze the Use of gamification in the learning of usability
- for the purposes of Evaluating and Improving the learning of usability
- in terms of the Adequacy of gamifying usability learning tools
- from the point of view of the Researcher
- in the context of Jakob Nielsen's usability heuristics.

The aforementioned goals were considered in order to pose the following research question:

- RQ. How does the use of a gamified tool affect the learning of usability heuristics?

5.4. Variable description

The next step was to decide on the variables to be used to carry out the experiment. Three independent variables were defined. TimePoint represents the points in time at which Exam1 and Exam2 were performed. EducationTool denotes the learning tool used: Heureka (EG) and UsabilityWebs (CG). Time_Inver_H represents the timeinterval spent on each heuristic: High_TimeInterval (time interval required to answer a heuristic, first tercile), Intermediate_TimeInterval (time interval required to answer a heuristic, second tercile) and Low_TimeInterval (time interval required to answer a heuristic, third tercile).

The dependent variables were the scores of those students who used Heureka, ScoreHeureka (M1); the scores of those students who used the usability webs, ScoreUsabilityWeb (M2); the scores attained by the students in the first exam, ScoreofExam1 (M3); the scores attained by the students in the second exam, ScoreofExam2 (M4); the success rate of the students who used Heureka for each heuristic, Success_Rate (M5); and the difference between the scores attained in the first and second exams, Score_Difference (M6).

5.5. Hypotheses

In order to answer the aforementioned research question, the following hypotheses were defined on the basis of the measures selected in the design of the empirical evaluation:

- H10, Null Hypothesis: Scores of students who used Heureka (M1–ScoreHeureka) are not affected by the timing of the students' evaluation (TimePoint).
- H11, Alternative Hypothesis: Scores of students who used Heureka (M1–ScoreHeureka) are affected by the timing of the students' evaluation (TimePoint).
- H20, Null Hypothesis: Scores of students who used usability websites (M2-ScoreUsabilityWeb) are not affected by the timing of the student's evaluation (TimePoint).
- H21, Alternative Hypothesis: Scores of students who used usability websites (M2-ScoreUsabilityWeb) are affected by the timing of the student's evaluation (TimePoint).
- H30, Null Hypothesis: Scores of the students in the first exam (M3-ScoreofExam1) are not affected by the type of learning tool used (EducationTool).
- H31, Alternative Hypothesis: Scores of the students in the first exam (M3-ScoreofExam1) are affected by the type of learning tool used (EducationTool).

- H40, Null Hypothesis: Scores of the students in the second exam (M4-ScoreofExam2) are not affected by the type of learning tool used (EducationTool).
- H41, Alternative Hypothesis: Scores of the students in the second exam (M4-ScoreofExam2) are affected by the type of learning tool used (EducationTool).
- H50, Null Hypothesis: The success rate of the students who used Heureka (M5-Success_Rate) is not affected by time they spent answering the heuristics (Time_Inver_H).
- H51, Alternative Hypothesis: The success rate of the students who used Heureka (M5-Success_Rate) is affected by time they spent answering the heuristics (Time_Inver_H).
- H60, Null Hypothesis: The difference between the scores attained in the first and second exams (M6- Score_Difference) is not affected by the type of learning tool used (EducationTool).
- H61, Alternative Hypothesis: The difference between the scores attained in the first and second exams (M6- Score_Difference) is affected by the type of learning tool used (EducationTool).
- The objective of each hypothesis is explained as follows:
- Hypothesis H1. Attempts to discover the extent to which Heureka enables students to retain, in the short term, the knowledge learned as regards Jakob Nielsen's 10 Usability Heuristics.
- Hypothesis H2. Attempts to discover to what extent traditional usability learning allows students to retain, in the short term, the knowledge learned as regards Jakob Nielsen's 10 Usability Heuristics.
- Hypothesis H3. Attempts to discover the extent to which the type of learning approach used enables students to achieve better knowledge on Jakob Nielsen's 10 Usability Heuristics immediately after the intervention.
- Hypothesis H4. Attempts to examine the extent to which the type of learning approach used enables students, in the short term, to achieve better knowledge on Jakob Nielsen's 10 Usability Heuristics.
- Hypothesis H5. Attempts to explore what heuristics are more difficult to learn with the gamified approach and to discover whether this is related to the time the students required in order to understand them.
- Hypothesis H6. Attempts to ascertain what learning approach allows students to achieve a greater increase in knowledge as regards Jakob Nielsen's 10 Usability Heuristics.

5.6. Statistical analysis

The data gathered was analyzed using the SPSS 24.0 statistical software package and Microsoft Office Excel 2016. The Shapiro-Wilk W statistical test was applied in order to verify whether the scores in the two tasks had a normal distribution. A standard level of significance (0.05) was selected so as to reject the null hypothesis. The Levene test was used to verify variance homogeneity.

Wilcoxon and Student-T tests for paired samples were used to search for any significant differences in the scores obtained by those students who made use of the usability websites and the Heureka tool in the first and second exams.

The Student-T test for unpaired samples was used to discover any significant differences between the scores attained by those students who used the usability websites and those who used the Heureka tool in the first exam. Furthermore, the Man-Whitney test was used to discover any significant differences between the scores attained by those students who used the usability websites and those who made use of the Heureka tool in the second exam. The Man-Whitney test was also used to discover whether the learning tool had an effect on the differences between the scores attained in the first and second exams.

Finally, the Kruskall-Wallis test for unpaired samples was used to discover whether the time spent answering questions (high, intermediate and low time intervals) had an impact on the success rates of the heuristics.

6. Results

6.1. Participants' characteristics

A demographic analysis of the 55 computer science students showed that 87.27% of the participants were male ($n = 48$) and 12.73% were female ($n = 7$). The population sample had a similar academic background, in the 20- to 21-year-old age range.

6.2. Descriptive statistics

Table 1 shows that CG students scored better in the first exam than in the second exam, while the scores attained by the EG students in the second exam were slightly higher than those attained in the first exam.

Table 2 shows the percentage of success of the students in the experimentation group for each of the heuristics. The study revealed that the heuristics for which the students obtained better results, considering the data obtained after giving an explanation by the teacher about the Nielsen heuristics, were: h8 (85.45%) and h3 (80.35%), while heuristics h6 (55.35%) and h1 (63.79%) were the worst performers.

Upon considering the results obtained by the students from the control and experimental groups in both the first and second examinations, it was determined that the heuristics with the highest percentage of success in the control group were h2 and h8, while those with the worst results were h5 and h7. In the experimental group, the heuristics with the highest percentage of success were h2, h3 and h8, while heuristics h5 and h4 obtained the worst results. Table 3 shows the results obtained by the students in the control and experimental groups after the first and second tests.

The study revealed that the heuristics on which the students spent most time were h5 ($A = 168.98$ s) and h9 ($A = 162.87$ s); while the heuristics on which the students spent least time were h2 ($A = 133.29$ s) and h8 ($A = 144.96$ s). Table 4 shows the average time the students spent on each heuristic.

In order to determine whether the heuristics on which most and least time was spent influenced the students' success rate, Table 5 was organized by considering the average time that the students took to respond to each heuristic.

Table 6 shows the descriptive statistics (number of heuristics, mean, median, and standard deviation) obtained for success rate.

Table 7 shows the descriptive statistics (number of students, mean, median, and standard deviation) obtained for mean difference. On average, the students in the experimental group (using Heureka) achieved better results than those in the control group.

6.3. Data analysis

6.3.1. H1: TimePoint – ScoreHeureka (M1)

The result obtained by the related sample *t*-test revealed that, owing to the means of the two exams, there was no statistically significant improvement as regards the scores obtained by the students in the period of time just after being taught the 10 Jakob Nielsen usability heuristics (Exam1) and those obtained after the experiment was repeated one week later (Exam2), with the statistical result of $t\ (25) = -0.75$, $p = 0.46$.

Table 1a
Descriptive statistics for student scores. "N": sample size; "M": mean; "Md": median; "SD": standard deviation.

Variables	N	M	Md	SD
ScoreUsabilityWeb				
Exam1	29	6,65	6,01	1,93
Exam2	29	6,44	7,34	2,32
ScoreHeureka				
Exam1	26	6,29	6,67	2,15
Exam2	26	6,57	6,34	1,69

Table 2
Results obtained by students from the experimentation group after interacting with Heureka.

Heuristics	Matches	Total	Success rate
h1-Visibility of system status	37	58	63.79
h2- Match between system and the real world	70	97	72.16
h3-User control and freedom	45	56	80.35
h4-Consistency and standards	79	102	77.45
h5-Error prevention	36	56	64.28
h6-Recognition rather than recall	62	112	55.35
h7-Flexibility and efficiency of use	35	51	68.62
h8- Aesthetic and minimalist design	47	55	85.45
h9-Help users recognize, diagnose, and recover from errors	81	101	80.19
h10-Help and documentation	65	90	72.22

6.3.2. H2: TimePoint – ScoreUsabilityweb (M2)

In order to discover out whether the scores of those students who learned by employing usability webs (M2ScoreUsabilityWeb) were affected by the timing of the students' evaluation (TimePoint), a Wilcoxon signed-rank test was performed for the variable ScoreUsabilityWeb. The results obtained did not show a statistically significant change between the scores obtained by the students just after being taught the 10 Jakob Nielsen usability heuristics (Exam1) and those attained after the experiment was repeated one week later (Exam2), with the statistical result of $Z = -0.09$ and $p = 0.99$.

6.3.3. H3: EducationTool – ScoreofExam1 (M3)

An independent sample test (Student T) for the variable ScoreofExam1 was used to compare the scores obtained by the CG students with those from the EG in Exam1. The results revealed that there was no statistically significant difference between the two groups. The statistic t (53) = 0.66, with $p = 0.52$.

6.3.4. H4: EducationTool – ScoreofExam2 (M4)

Although the EG students scored better in the second test (Exam2), the differences are not sufficiently statistically significant to be able to state that the students in one group achieved greater knowledge on Jakob Nielsen's heuristics than those in the other group. A Mann-Whitney U test for the ScoreofExam2 variable was applied to Exam2. The results revealed that there was no significant difference between CG and EG. In this case, the statistic obtained was $U = 365$ and $p = 0.84$.

6.3.5. H5: Time_Inver_H – Success_Rate (M5)

A Kruskal-Wallis H test showed that there was no statistically significant difference between the average time it took the students to respond to the heuristics and the students' percentage of success for each

Table 4
Heuristics on which the students from the experimentation group spent most and least time.

Heuristics	Total submissions	Seconds	Average
h1	58	8572	147.79
h2	97	12,930	133.29
h3	56	8781	156.80
h4	102	15,849	155.38
h5	56	9463	168.98
h6	112	17,566	156.83
h7	51	7770	152.35
h8	55	7973	144.96
h9	101	16,450	162.87
h10	90	13,148	146.08

Table 5
Heuristics organized according to the average time spent by students on each heuristic.

Heuristics	Average	Success Rate
Heuristics on which most time was spent		
h5	168.98	64.28
h9	162.87	80.19
h6	156.83	55.35
Heuristics on which an intermediate amount of time was spent		
h3	156.80	80.35
h4	155.38	77.45
h7	152.35	68.62
h1	147.79	63.79
Heuristics on which least time was spent		
h10	146.08	72.22
h8	144.96	85.45
h2	133.29	72.16

Table 6
Descriptive statistics for success rate. "N": Number of students; "M": Mean; "Md": Median; "SD": Standard deviation.

Success_Rate	N	M	Md	SD
Most_Time Spent	3	66.60	64.28	12.58
Intermediate_Time	4	72.55	73.03	7.68
Least_Time Spent	3	76.61	72.22	7.65

Table 7
Descriptive statistics for mean difference. "N": Number of students; "M": Mean; "Md": Median; "SD": Standard deviation.

Score_Difference	N	M	Md	SD
Experimentation (Heureka)	26	0.94	0.93	0.09
Control	29	0.90	0.93	0.19

Table 3
Results obtained by the students from the control and experimental groups in the first and second exams.

Heuristics	Matches				Success rate			
	First Exam (CG)	First Exam (EG)	Second Exam (CG)	Second Exam (EG)	First Exam (CG)	First Exam (EG)	Second Exam (CG)	Second Exam (EG)
h1	20	19	18	20	68.96	73.07	62.06	76.92
h2	29	26	28	25	100	100	96.55	96.15
h3	27	25	27	25	93.10	96.15	93.10	96.15
h4	18	11	21	15	62.06	42.30	72.41	57.69
h5	12	8	11	8	41.37	30.76	37.93	30.76
h6	19	16	16	19	65.51	61.53	55.17	73.07
h7	13	15	13	13	44.82	57.58	44.82	50.00
h8	29	25	28	25	100	96.15	96.55	96.15
h9	28	24	26	24	96.55	92.30	89.65	92.30
h10	22	18	22	19	75.86	69.23	75.86	73.07

heuristic, $\chi2(2) = 1.47$ and $p = 0.48$, with a mean percentage of success of 66.61 for High_TimeInterval, 72.55 for Intermediate_TimeInterval and 76.61 for Low_TimeInterval.

6.3.6. H6: EducationTool – Score_Difference (M6)

In order to determine whether the difference between the control group's scores in the first and second tests was greater than the difference between the scores attained by the students in the experimental group, an independent sample test (Mann-Whitney U) was used for the variable Score_Difference. The results revealed that there was no statistically significant difference. The statistic $U = 402{,}50$ and $p = 0.66$.

6.4. Survey

A questionnaire concerning the participation in the experiment was filled out by the students. The aim of this survey was to collect feedback regarding the students' perceptions with respect to their experience with Heureka. A total of 14 questions were administrated, employing a five-point Likert-type scale (1 = Completely disagree; 2 = Disagree; 3 = Neither agree nor disagree; 4 = Agree; 5 = Completely agree).

The questionnaire was designed by keeping in mind the Technology Acceptance Model (TAM)(Al-Qaysi et al., 2021), which helps to put the results into a perspective of technology adoption. The TAM highlights the need to be conscious of the socially constructed processes in which tools are deployed and used on a daily basis (Al-Qaysi et al., 2021). This is because, when users are presented with new technology, a number of factors—in particular perceived usefulness and perceived ease of use—influence their decision regarding how and when they will use this new technology. Table 8 shows the questions, the means, standard deviations and median for the students' answers.

Table 8
Means, standard deviations and medians of students' perceptions. "M": mean; "SD": standard deviations; "Md": median.

Id	Question	M	SD	Md
	Block 1: Attitude of use			
Q1	I would use the Heureka tool if I needed to study Jakob Nielsen's heuristics.	4	0.73	4
Q2	I would use the Heureka tool to improve my performance in the subject User Interfaces.	3.7	0.66	4
	Block 2: Intention of use			
Q3	I would recommend the Heureka tool to future UI students.	3.85	0.93	4
Q4	I would recommend the Heureka tool in order to learn Jakob Nielsen's heuristics.	4.2	0.83	4
	Block 3: Perceived ease of use			
Q5	In general, I found the application intuitive.	4.35	0.75	4.5
Q6	I easily realized that I could use the 'View Definition' button to get help.	3.95	1.00	4
Q7	I easily realized that I had to press the 'Confirm Answer' button to set my final answer.	4.4	0.82	5
Q8	I easily noticed that some questions asked for an example that did not comply with the heuristics.	3.8	1.06	4
Q9	I knew for sure how many mistakes I could make at any given time.	3.6	1.10	3.5
Q10	I easily noticed when I got a question right or wrong.	4.3	0.80	4.5
Q11	When I missed a question, I easily realized the right one.	4.1	0.72	4
	Block 4: Perceived utility			
Q12	Using Heureka would make it easier for me to study Jakob Nielsen's heuristics.	4.15	0.88	4
Q13	I would have preferred to use the Heureka tool in class in order to learn about Jakob Nielsen's heuristics, instead of listening to the teacher's verbal explanation.	3.5	1.15	4
Q14	I would have preferred to have used the Heureka tool to review Jakob Nielsen's heuristics before the exam, instead of studying the slides.	3.9	0.97	4

7. Discussion and lessons learned

7.1. Discussion

There is little consensus as to whether gamification has positive effects on performance. Previous studies have reported mixed findings. The gamification component is believed to be effective in enhancing students' motivation and improving their learning experience, engagement and performance (Legaki et al., 2020). Several studies (Manzano-le et al., 2021) have shown that the addition of game mechanics (such as badges, levels and leader boards) has a positive effect on learner engagement. However, Zainuddin et al. (2020) claim that critics have argued that these mechanics create only extrinsic motivation, not intrinsic motivation; that is, learners complete a task simply to earn a badge, not for the satisfaction of gaining new knowledge and skills. An increase in performance was observed in the Heureka group between Exam1 and Exam2, in the period of time of a week, although this improvement was not statistically significant (H1). The Heureka group students played during the week and thus probably remembered the heuristics better thanks to repetition. This is partly in line with the results obtained in an empirical study carried out by (Barreto et al., 2015), which compared results of the pretest and post-test for students using UsabilityGame (experimental group) and the Monopoly board game (control group). Significant statistical improvements were found as regards learning the concepts of usability. This means that those students who used UsabilityGame learned heuristic evaluation concepts better than those who used the Monopoly board game.

Nevertheless, the results shown in Section 5 indicate that the use of Heureka did not help the students gain better academic results in Exam1 and Exam 2 (hypotheses H3 and H4). We believe that the reason for these results lies in the fact that the students were exposed to the game for only a short amount of time. Another reason could be that some Heureka cards may recall several of Nielsen's heuristics at the same time, which could have created confusion in the learners. An exploration-based learning tool such as Heureka could be an intuitive and effective approach in domains less prone to subjective evaluation by the students. For instance, this is the case of teaching software process improvement, such as SPICE (Software Process Improvement Capability dEtermination) (Dorling and McCaffery, 2012). What is more, another factor to consider is that gamification may have less impact on usability heuristics than occurs in other disciplines, as the material and examples used in the traditional teaching of Nielsen heuristics can also be enjoyable.

These results show that further research is required in order to add new gamified elements and analyze which of them are most influential so as to achieve a better learning performance. Note that Heureka does not employ the game elements most frequently used in literature, such as badges and leaderboards, which will be included in future work. This is because evidence has been found concerning a positive effect on academic performance in teaching programming fundamentals by using the "Clara" framework gamified with star ratings, badges and challenges (Bogdanovych and Trescak, 2016). A positive impact on academic performance was also evidenced when using the UDPiler compiler to teach C programming in comparison to using a non-gamified platform (Marín et al., 2019). UDPiler is gamified with points, badges and leaderboards. De Marcos et al. (2014) ascertained that a gamification learning approach improved academic achievement in practical assignments. However, a traditional e-learning approach was better for students in terms of knowledge. The result obtained in De Marcos' study could have been owing to the fact that UDPiler compiler was not easy to use, as reported by the students in an attitudinal survey. Note that perceived usefulness and perceived ease of use are considered key variables that explain outcome measures such as performance (Marangunić and Granić, 2015). The positive effects of gamification may be blunted by a poorly designed tool.

Although the hypotheses test did not find statistically significant difference in the percentage of success between time interval groups, extreme behavior in the percentage of success was observed for three heuristics. First, two heuristics attained lower success rates and needed more time to be answered, namely h5 – "Error prevention" and h6 – "Recognition rather than recall". Extensive work can be found on tips to improve both the h5 and h6 heuristics, but these tips may be difficult to translate into Heureka cards: (1) on the one hand, in the case of h5, one of the main goals of a well-designed user interface is to prevent interaction problems, thus promoting error prevention by eliminating error-prone conditions or checking for them and presenting users with confirmation dialogs so as to avoid "unconscious slips" and "conscious mistakes**"** (Sherwin, 2019**)**. Norman identifies two categories of user errors (Laubheimer, 2015): slips occur "when a user is on autopilot, and takes the wrong actions in service of a reasonable goal"; mistakes occur "when a user has developed a mental model of the interface that is not correct, and forms a goal that does not suit the situation well"; (2) on the other hand, with regard to h6, some useful tips include providing easy access to the history and previously visited content as long as visible, and intuitive interfaces (Budiu, 2014). According to (Dix et al., 2004) one way in which to achieve the latter is by ensuring that the interface is synthesizable, i.e., users must be able to evaluate the effect of previous operations on the current state. As stated above, it would appear that some of these tips cannot be easily translated into static content, i.e., Heureka cards. If these heuristics are to be illustrated, dynamic visual resources are required. On the contrary, heuristic h8 – "Aesthetic and minimalist design" attained the best success rate and required the least amount of time to be solved: this heuristic is intuitively well-suited to representation on static cards.

With regard to subjective perceptions, the students pointed out that the Heureka tool is recommendable as regards learning Jakob Nielsen's heuristics ($M = 4.20$). They also perceived Heureka to be intuitive ($M = 4.35$). The utility perceived by the subjects in our experiment confirms findings obtained with UsabilityGame (Barreto et al., 2015). In our survey, the students affirmed that using Heureka would make it easier to study Jakob Nielsen's heuristics ($M = 4.15$). Positive perceptions were also found in the UsabilityGame experiment, in which around 65% of the students responded that they strongly agreed or partially agreed with the utility of the game as regards teaching the evaluation of heuristics. Moreover, more than 80% of the students believed that the idea of teaching usability through UsabilityGame was adequate. Motivation was similarly highlighted by more than 75% of the students who used UsabilityCity (Ferreira et al., 2014). These results are also similar to gamified experiences in other settings such as the UDPiler (Marín et al., 2019), in which more than 80% of the students stated that UDPiler helped them to obtain better results.

7.2. Lessons learned

Based on the experience reported in this paper, this section synthesizes some lessons learned in relation to the development of a gamified learning system in the domain of usability and user interface interaction.

On the software engineering methodology

- Adopt user-centered design to address the students' interests and needs. The system must engage students' attention for learning to be effective.
- Employ an agile software development based on user stories which reduces development time and allows to provide the trainer with a non-complex gamified system in a short period of time.
- Select the best implementation strategy between a general-purpose gamification platform to reduce cost or a self-built application to achieve a higher integration and flexibility. This decision should be based on previous knowledge, available resources and the pre-established objectives.
- Develop a robust tracking and data collection system in advance. Data on the use of the gamified system is crucial to provide feedback to the designer and guide educators' decision-making. A video recording tool, especially in the case of a subject on user interfaces, is an interesting instrument to analyze student interactions.

On the gamification process

- Follow a gamification process that gives the team a clear direction, helping to approach the problem in a systematic way in order to integrate the various elements involved (stakeholders, theories, methodologies and technologies) and to manage the activities necessary for the success of the project.
- Take into account all the key elements to design the gamification strategy such as game elements, mechanics, user type, context and motivation. A conceptual framework to help designer to cover these principles can be useful, such as Gamicards (c.f. Section 4.2).
- Determine the type of user targeted by the gamification effort in order to fine-tune the design of the gamification strategy and the design elements to be included in the gamified system.
- Sufficient time should be planned to validate and review the questions and answers used in the game. The usability domain is given to writing questions that may later be ambiguous, in the sense that these questions may refer to more usability issues than originally expected, so all the questions should be carefully reviewed by a team, especially if they are self-correcting multiple-choice questions, where students are not allowed to reason their answers.

8. Threats to the validity of the study

The effect of different threats to validity has been analyzed in the present research, focusing on threats to internal, external and the conclusion validity.

8.1. Internal validity

During the selection process, the students who participated in the experiment received prior preparation on Nielsen's 10 Usability Heuristics for User Interface Design, thus counteracting the internal threat to validity that may be caused by the effect of the confounding variable related to prior knowledge and experience in this topic. In addition, the effect of this, and other possible confounding variables such as: (1) students' motivation, (2) students' personality and (3) students' fatigue, was controlled by randomly forming the groups of students in the design of the experiment.

8.2. External validity

Students may feel overwhelmed when taking extensive exams, thus resulting in the fatigue effect. An exam comprising only 10 questions was, therefore, prepared, one question for each heuristic. However, the small size of this experimental object could have been a threat to the external validity of the results. The size and complexity of the exam had to be limited, as it took place within a time-constrained HCI undergraduate course.

Internal consistency, construct validity and reliability were not quantified in the questionnaire used in the experiment. In order to mitigate the threat of construct validity, several reviews of the content of the questionnaire were carried out by three professionals in the field of usability engineering, and the TAM (Al-Qaysi et al., 2021) was used to discover the acceptance of the technology. In this regard, one of the instructors has ten years of experience in teaching usability, which could add reliability to the instrument developed. Errors related to some questions were identified and corrected in these reviews. In addition, we avoided writing negatives or double negatives in the survey questions, as respondents tend to spend a lot of time figuring out whether they agree or disagree with the questions. However, there is a general tendency for assent rather than dissent (acquiescence) (J. M. Johnson et al., 2011), signifying that the mean of all responses may tend toward the side of agreement.

Finally, when experiments are conducted with students, the external validity may also be threatened; if usability engineering professionals are used, the representativeness of the participants may be improved. Nevertheless, controlled experiments provide insights into issues and problems that can later be considered in industrial case studies (Arisholm et al., 2006). As suggested by Carver et al. (2003), the results obtained through empirical studies conducted with students have relevance in the progress of Software Engineering (Salman et al., 2015).

8.3. Conclusion validity

The collection of statistical data, the reliability of the measurement and the validity of the statistical tests is related to the validity of the conclusions and may affect the ability to draw a correct conclusion. The results of the tests performed in the experiment were obtained from the virtual classroom of the University of OMITTED FOR REVIEW in order to avoid threats related to data collection procedures.

The statistical inference method was employed in order to mitigate threats to the validity of the conclusions. Parametric and nonparametric statistical tests were used to test the hypotheses, complying with all the necessary requirements, and thus ensuring that the validity of the results obtained would be acceptable.

9. Conclusion

To the best of our knowledge, no tool related to the learning of Jakob Nielsen's 10 usability heuristics has been built. This work presents a tool that makes use of gamification elements to support an HCI subject on the BSc in Computer Engineering. The Heureka gamified environment was designed in order to allow users to self-study concepts related to Jakob Nielsen's 10 usability heuristics.

A survey using questionnaires was conducted so as to examine the students' perception of the game by considering the TAM. Heureka obtained positive results when analyzing the overall results attained by the students who played the game. In this context, we believe that the game helps students learn some concepts about Jakob Nielsen's 10 heuristics, although we found no statistical evidence regarding the benefits of Heureka using a 95% confidence interval. According to Nielsen (1993), a less strong degree of confidence can be adopted in usability engineering as long as sometimes decisions have to be made on the basis of fairly unreliable data, and "one should certainly do so since *some* data is better than *no* data". In our case, there is some evidence that Heureka was effective as regards teaching Jakob Nielsen's 10 usability heuristics. The empirical results obtained may be associated with the limited amount of time that the students were exposed to Heureka. Moreover, the randomness of the selection of the cards resulted in the students being able to interact with some heuristics but not with others. This means that some heuristics may not have been addressed during a game.

In future work, we plan to incorporate other gamification elements into the Heureka tool in order to attain a comparative framework from which to select the most appropriate gamification elements for usability learning and, more specifically, with which to learn Jakob Nielsen's 10 usability heuristics. Although not specifically relevant to the user types initially targeted in this work, gamification elements of the socializer player types (such as social networks, social status, competition and guilds or teams) and philanthropists (i.e., gifs, collection, meaning, sharing knowledge and trading) should be considered in the new version of the tool, as they may have an impact on intrinsic motivation, enable students to re-engage in challenging and creative activities, and help extend the spectrum of potential participants and application scenarios. Moreover, the new elements could promote engagement, altruism and student performance, and create social connections between participants, thus allowing them to interact with one another (Nasirzadeh and Fathian, 2020) (Klock et al., 2020).

CRediT authorship contribution statement

Raimel Sobrino-Duque: Methodology, Formal analysis, Writing – original draft, Writing – review & editing. **Noelia Martínez-Rojo:** Conceptualization, Software. **Juan Manuel Carrillo-de-Gea:** Methodology, Writing – review & editing. **Juan José López-Jiménez:** Data curation. **Joaquín Nicolás:** Conceptualization, Methodology, Writing – review & editing, Funding acquisition. **José Luis Fernández-Alemán:** Methodology, Validation, Formal analysis, Writing – review & editing, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research is part of the OASSIS-UMU (PID2021-122554OB-C32) project (supported by the Spanish Ministry of Science and Innovation), the BIZDEVOPSGLOBAL-UMU (RTI2018-098309-B-C33) project (supported by the Spanish Ministry of Science, Innovation, and Universities), and the Network of Excellence in Software Quality and Sustainability (TIN2017-90689-REDT). These projects are also founded by the European Regional Development Fund (ERDF).

References

Al-Qaysi, N., Mohamad-Nordin, N., Al-Emran, M, 2021. Factors affecting the adoption of social media in higher education: a systematic review of the technology acceptance model BT - recent advances in intelligent systems and smart applications. In: Al-Emran, M., Shaalan, K., Hassanien, A.E. (Eds.), Recent Advances in Intelligent Systems and Smart Applications. Springer International Publishing, pp. 571–584. https://doi.org/10.1007/978-3-030-47411-9_31.

Alhammad, M.M., Moreno, A.M., 2018. Gamification in software engineering education: a systematic mapping. J. Syst. Softw. 141, 131–150. https://doi.org/10.1016/j.jss.2018.03.065.

Aparicio, A.F., Vela, F.L.G., Sánchez, J.L.G., Montes, J.L.I., 2012. Analysis and application of gamification. In: ACM International Conference Proceeding Series. https://doi.org/10.1145/2379636.2379653.

Arisholm, E., Briand, L.C., Hove, S.E., Labiche, Y., 2006. The impact of UML documentation on software maintenance: an experimental evaluation. IEEE Trans. Software Eng. 32 (6), 365–381. https://doi.org/10.1109/TSE.2006.59.

Bai, S., Hew, K.F., Huang, B., 2020. Does gamification improve student learning outcome? Evidence from a meta-analysis and synthesis of qualitative data in educational contexts. Educ. Res. Rev. 30 (December 2019), 1–20. https://doi.org/10.1016/j.edurev.2020.100322.

Barata, G., Gama, S., Jorge, J., Gonçalves, D., Goncalves, D., 2013. Engaging engineering students with gamification an empirical study. In: Games and Virtual Worlds for Serious Applications (VS-GAMES), 2013 5th International Conference On, pp. 1–8. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6624228.

Barber, C.S., 2021. When students are players : toward a theory of student-centric edu-gamification systems. J. Inf. Syst. Educ. 32, 53–64.

Barreto, F., Benitti, V., Campus, S.C., Sommariva, L., Uruguai, R., 2015. Evaluation of a game used to teach usability to undergraduate students in computer science. J Usability Stud 11 (1), 21–39.

Basili, V.R., Dieter Rombach, H., 1988. The TAME project: towards improvement-oriented software environments. IEEE Trans. Software Eng. 14 (6), 758–773. https://doi.org/10.1109/32.6156.

Blohm, I., Leimeister, J.M., 2013. Gamification: design of IT-based enhancing services for motivational support and behavioral change. Bus. Inf. Syst. Eng. 5 (4), 275–278. https://doi.org/10.1007/s12599-013-0273-5.

Bogdanovych, A., Trescak, T., 2016. Teaching programming fundamentals to modern university students. In: CSEDU 2016 - Proceedings of the 8th International Conference on Computer Supported Education, 2, pp. 308–317. https://doi.org/10.5220/0005809403080317.

Budiu, R., 2014. Memory Recognition and Recall in User Interfaces. World Leaders in Research-Based User Experience. https://www.nngroup.com/articles/recognition-and-recall/.

Calderón, A., Boubeta-Puig, J., Ruiz, M., 2018. MEdit4CEP-Gam: a model-driven approach for user-friendly gamification design, monitoring and code generation in CEP-based systems. Inf. Softw. Technol. 95 (November 2017), 238–264. https://doi.org/10.1016/j.infsof.2017.11.009.

Carrion, M., Santorum, M., Flores, H., Aguilar, J., Perez, M., 2019. Serious game, gamified applications, educational software: a comparative study. In: Proceedings - 2019 International Conference on Information Systems and Software Technologies, ICI2ST 2019, pp. 55–62. https://doi.org/10.1109/ICI2ST.2019.00015.

Carroll, J.M., Rosson, M.B., 2006. Case studies as minimalist information. IEEE Trans. Prof. Commun. 49 (4), 297–310. https://doi.org/10.1109/TPC.2006.885836.

Carver, J., Jaccheri, L., Morasca, S., Shull, F., 2003. Issues in using students in empirical studies in software engineering education. In: Proceedings - International Software Metrics Symposium, pp. 239–249. https://doi.org/10.1109/METRIC.2003.1232471.

Constantine, L., Lockwood, L., 1999. Software For use: A Practical Guide to the Models and Methods of Usage-Centered Design. Addison-Wesley.

Crittenden, W.F., Biel, I.K., Lovely, W.A., 2019. Embracing digitalization: student learning and new technologies. J. Market. Educ. 41 (1), 1–10. https://doi.org/10.1177/0273475318820895.

Dale, S., 2014. Gamification: making work fun, or making fun of work? Bus. Inf. Rev. 31 (2), 82–90. https://doi.org/10.1177/0266382114538350.

De-Marcos, L., Domínguez, A., Saenz-De-Navarrete, J., Pagés, C., 2014. An empirical study comparing gamification and social networking on e-learning. Comput. Educ. 75, 82–91. https://doi.org/10.1016/j.compedu.2014.01.012.

Deterding, S., 2011. Situated motivational affordances of game elements: a conceptual model. In: CHI 2011 Workshop "Gamification," January 2011, pp. 2–4. http://gamification-research.org/chi2011/papers.

Deterding, S., Dixon, D., Khaled, R., Nacke, L., 2011. From game design elements to gamefulness: defining "gamification. In: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments, MindTrek 2011, pp. 9–15.

Dicheva, D., Dichev, C., Agre, G., Angelova, G., 2015. Gamification in education: a systematic mapping study. Educ. Technol. Soc. 18 (3), 75–88.

Dix, A., Finlay, J., Abowd, G.D., Beale, R., 2004. Human–computer Interaction. Pearson Education.

Domínguez, A., Saenz-De-Navarrete, J., De-Marcos, L., Fernández-Sanz, L., Pagés, C., Martínez-Herráiz, J.J., 2013. Gamifying learning experiences: practical implications and outcomes. Comput. Educ. 63, 380–392. https://doi.org/10.1016/j.compedu.2012.12.020.

Dorling, A., McCaffery, F., 2012. The gamification of SPICE. Commun. Comput. Inf. Sci., 290 CCIS 295–301. https://doi.org/10.1007/978-3-642-30439-2_35.

Santos, Dos, L, A., Souza, M.R.D.A., Figueiredo, E., Dayrell, M, 2018. Game elements for learning programming: a mapping study. In: CSEDU 2018 - Proceedings of the 10th International Conference on Computer Supported Education, 2, pp. 89–101. https://doi.org/10.5220/0006682200890101.

Faulkner, X., Culwin, F., 2001. The internet as a resource for teaching non internet topics. In: Proceedings - 2001 Symposium on Applications and the Internet Workshops, SAINT 2001, pp. 51–55. https://doi.org/10.1109/SAINTW.2001.998209.

Ferreira, B.M., Rivero, L., Lopes, A., Marques, A.B., Conte, T., 2014. UsabiliCity: Um Jogo de Apoio ao Ensino de Propriedades de Usabilidade de Software Através de Analogias. In: Anais Do XXV Simpósio Brasileiro de Informática Na Educação (SBIE 2014), 1, pp. 1273–1282. https://doi.org/10.5753/cbie.sbie.2014.1273.

Ferro, L.S., 2021. The Game Element and Mechanic (GEM) framework: a structural approach for implementing game elements and mechanics into game experiences. Entertain. Comput. 36 (July 2020), 100375 https://doi.org/10.1016/j.entcom.2020.100375.

Ferro, L.S., Walz, S.P., Greuter, S., 2014. Gamicards - an alternative method for paper-prototyping the design of gamified systems. In: International Conference on Entertainment Computing, 8770, pp. 11–18. https://doi.org/10.1007/978-3-662-45212-7_2.

Gordon, N., Brayshaw, M., Grey, S., 2013. Maximising gain for minimal pain: utilising natural game mechanics. ITALICS Innov. Teach. Learn. Inf. Comput. Sci. 12 (1), 27–38. https://doi.org/10.11120/ital.2013.00004.

Gounaridou, A., Siamtanidou, E., Dimoulas, C., 2021. A serious game for mediated education on traffic behavior and safety awareness. Educ. Sci. 11 (3), 127. https://doi.org/10.3390/educsci11030127.

Guerino, G.C., Valentim, N.M.C., 2020. Usability and user experience evaluation of natural user interfaces: a systematic mapping study. IET Software 14 (5), 451–467. https://doi.org/10.1049/iet-sen.2020.0051.

Haaranen, L., Ihantola, P., Hakulinen, L., Korhonen, A., 2014. How (not) to introduce badges to online exercises. In: SIGCSE 2014 - Proceedings of the 45th ACM Technical Symposium on Computer Science Education, pp. 33–38. https://doi.org/10.1145/2538862.2538921.

Hakulinen, L., Auvinen, T., Korhonen, A., 2013. Empirical study on the effect of achievement badges in TRAKLA2 online learning environment. In: Proceedings - 2013 Learning and Teaching in Computing and Engineering, LaTiCE 2013, pp. 47–54. https://doi.org/10.1109/LaTiCE.2013.34.

Harley, A. (2019). Usability *heuristic 1: visibility of system* status (Video). https://www.nngroup.com/videos/usability-heuristic-system-status/.

Herbert, B., Charles, D., Moore, A., Charles, T., 2014. An investigation of gamification typologies for enhancing learner motivation. In: Proceedings - 2014 International Conference on Interactive Technologies and Games, ITAG 2014, pp. 71–78. https://doi.org/10.1109/iTAG.2014.17.

Herzig, P., Ameling, M., Schill, A., 2012. A generic platform for enterprise gamification. In: Proceedings of the 2012 Joint Working Conference on Software Architecture and 6th European Conference on Software Architecture, WICSA/ECSA 2012, pp. 219–223. https://doi.org/10.1109/WICSA-ECSA.212.33.

Herzing, P., Ameling, M., Wolf, B., Schill, A., 2015. Implementing gamification: requirements and gamification platforms. Gamification Educ. Bus. 431–450. https://doi.org/10.1007/978-3-319-10208-5.

Hill, J.M.D., Ray, C.K., Blair, J.R.S., Carver, C.A., 2003. Puzzles and games: addressing different learning styles in teaching operating systems concepts. In: SIGCSE Bulletin (Association for Computing Machinery, Special Interest Group on Computer Science Education), pp. 182–186. https://doi.org/10.1145/792548.611964.

Huotari, K., Hamari, J., 2012. Defining gamification - a service marketing perspective. In: Proceedings of the IADIS International Conference Interfaces and Human Computer Interaction 2012, IHCI 2012, Proceedings of the IADIS International Conference Game and Entertainment Technologies 2012, pp. 17–22. https://doi.org/10.1145/2393132.2393137.

ISO 9241. (2018). *Ergonomics of human-system interaction — Part 11: usability: definitions and concepts*. https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-2:v1:en.

Johnson, J.M., Bristow, D.N., Schneider, K.C., 2011. Did you not understand the question or not? An investigation of negatively worded questions in survey research. J. Appl. Bus. Res. (JABR) 20 (1), 75–86. https://doi.org/10.19030/jabr.v20i1.2197.

Johnson, L., Adams Becker, S., Cummins, M., Estrada, V., Freeman, A., Hall, C, 2016. Horizon report - 2016 higher education edition. NMC Horizon Rep. http://www.nmc.org/publications/2014-horizon-report-higher-ed.

Kapp, K.M., 2012. The Gamification of Learning and instruction: Game-Based Methods and Strategies for Training and Education, 148. John Wiley & Sons.

Khalil, M., Wong, J., De Koning, B., Ebner, M., Paas, F., 2018. Gamification in MOOCs: a review of the state of the art. In: *IEEE Global Engineering Education Conference, EDUCON, 2018-April*, pp. 1629–1638. https://doi.org/10.1109/EDUCON.2018.8363430.

Kim, S., 2021. How a company's gamification strategy influences corporate learning: a study based on gamified MSLP (Mobile social learning platform). Telematics Inform. 57 (September 2020), 1–19. https://doi.org/10.1016/j.tele.2020.101505.

Klock, A.C.T., Gasparini, I., Pimenta, M.S., Hamari, J., 2020. Tailored gamification: a review of literature. Int. J. Hum. Comput. Stud. 144 (September 2019) https://doi.org/10.1016/j.ijhcs.2020.102495.

Klock, A.C.T., Pimenta, M.S., Gasparini, I., 2018. A systematic mapping of the customization of game elements in gamified systems. In: XVII Brazilian Symposium on Computer Games and Digital Entertainment (SBGames 2018)N, pp. 1–8. In: http://www.sbgames.org/sbgames2018/proceedings-eng.

Labrador, E., Villegas, E., 2014. Sistema Fun Experience Design (FED) aplicado en el aula. ReVisión 7 (2). http://www.aenui.net/ojs/index.php?journal=revision&page=article&op=viewArticle&path%5B%5D=147&path%5B%5D=242.

Laubheimer, P., 2015. Preventing User Errors: Avoiding Conscious Mistakes. World Leaders in Research-Based User Experience. https://www.nngroup.com/articles/user-mistakes/.

Legaki, N.Z., Xi, N., Hamari, J., Karpouzis, K., Assimakopoulos, V., 2020. The effect of challenge-based gamification on learning: an experiment in the context of statistics education. Int. J. Hum. Comput. Stud. 144 (November 2019) https://doi.org/10.1016/j.ijhcs.2020.102496.

Ludi, S., 2005. Providing students with usability testing experience: bringing home the lesson "the user is not like you. In: *Proceedings - Frontiers in Education Conference, FIE, 2005*, pp. 6–11. https://doi.org/10.1109/fie.2005.1611949.

Maican, C., Lixandroiu, R., Constantin, C., 2016. Interactivia.ro - a study of a gamification framework using zero-cost tools. Comput Human Behav 61, 186–197. https://doi.org/10.1016/j.chb.2016.03.023.

Manzano-le, A., Camacho-lazarraga, P., Guerrero, M.A., Guerrero-puerta, L., Aguilar-parra, J.M., & Alias, A. (2021). Betw*een level up and game over : a systematic literature review of gamification in educati*on. 1–14.

Marangunić, N., Granić, A., 2015. Technology acceptance model: a literature review from 1986 to 2013. Univ. Access Inf. Soc. 14 (1), 81–95. https://doi.org/10.1007/s10209-014-0348-1.

Marczewski, A., 2015. Gamification, Game Thinking and Motivational design. *Even Ninja Monkeys Like to Play*. CreateSpace Independent Publishing Platform. https://www.gamified.uk/wp-content/uploads/2018/10/Narrative-Chapter.pdf.

Marín, B., Frez, J., Cruz-Lemus, J., Genero, M., 2019. An empirical investigation on the benefits of gamification in programming courses. ACM Trans. Comput. Educ. 19 (1), 1–22. https://doi.org/10.1145/3231709.

Meijer, H.A.W., Graafland, M., Obdeijn, M.C., Dieren, S.Van, Goslings, J.C., Schijven, M P, 2021. Serious game versus standard care for rehabilitation after distal radius fractures : a protocol for a multicentre randomised controlled trial. BMJ Open 11, 1–8. https://doi.org/10.1136/bmjopen-2020-042629.

Muratet, M., Torguet, P., Viallet, F., Jessel, J.P., 2011. Experimental feedback on Prog&Play: a serious game for programming practice. Comput. Graphics Forum 30 (1), 61–73. https://doi.org/10.1111/j.1467-8659.2010.01829.x.

Nasirzadeh, E., Fathian, M., 2020. Investigating the effect of gamification elements on bank customers to personalize gamified systems. Int. J. Hum. Comput. Stud. 143 (March 2019) https://doi.org/10.1016/j.ijhcs.2020.102469.

Nicholson, S., 2012. A user-centered theoretical framework for meaningful gamification. Games+Learning+Society 8.0 18, S66–S67. https://doi.org/10.1089/dia.2016.2506.

Nielsen, Jakob., 2020. 10 Usability Heuristics for User Interface Design. World Leaders in Research-Based User Experience. https://tfa.stanford.edu/download/TenUsabilityHeuristics.pdf.

Nielsen, Jakob., 1993. Usability Engineering. Morgan Kaufmann.

Norman, Don., 2013. The Design of Everyday things: Revised and Expanded Edition. Basic Books. https://doi.org/10.1145/1340961.1340979.

Norman, Donald., 1989. The Psychology of Everyday Things. Doubleday.

O'Donovan, S., Gain, J., Marais, P., 2013. A case study in the gamification of a university-level games development course. In: ACM International Conference Proceeding Series, pp. 242–251. https://doi.org/10.1145/2513456.2513469.

Quiñones, D., Rusu, C., 2017. How to develop usability heuristics: a systematic literature review. Comput. Standards Interfaces 53 (March), 89–122. https://doi.org/10.1016/j.csi.2017.03.009.

Rodrigues, E., Schwan-estrada, K.R.F., Fiori-tutida, A.C.G., Stangarlin, J.R., Eugênia, M., & Cruz, S. (2007). *Orgânico contra Sclerotinia sclerotiorum pelo extrato de gengibre*. 124–128.

Rodrigues, L., Oliveira, A., Rodrigues, H., 2019. Main gamification concepts: a systematic mapping study. Heliyon 5 (7), e01993. https://doi.org/10.1016/j.heliyon.2019.e01993.

Rose, D.H., Meyer, A., 2007. Teaching every student in the digital age: universal design for learning. Educ. Technol. Res. Dev. 55 (5), 521–525. https://doi.org/10.1007/s11423-007-9056-3.

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

Sakamoto, M., Nakajima, T., Alexandrova, T., 2012. Value-based design for gamifying daily activities. In: Entertainment Computing - ICEC 2012, pp. 421–424.

Salman, I., Misirli, A.T., Juristo, N., 2015. Are students representatives of professionals in software engineering experiments?. In: Proceedings - International Conference on Software Engineering, 1, pp. 666–676. https://doi.org/10.1109/ICSE.2015.82.

Sardi, L., Idri, A., Fernández-Alemán, J.L., 2017. A systematic review of gamification in e-Health. J. Biomed. Inform. 71, 31–48. https://doi.org/10.1016/j.jbi.2017.05.011.

Schell, J., 2019. The Art of Game Design: A book of Lenses. CRC Press. https://doi.org/10.1201/b22101.

Seaborn, K., Fels, D.I., 2015. Gamification in theory and action: a survey. Int. J. Hum. Comput. Stud. 74, 14–31. https://doi.org/10.1016/j.ijhcs.2014.09.006.

Sharma, M.K., Sharma, R.C., 2021. Innovation framework for excellence in higher education institutions. Glob. J. Flexible Syst. Manage. https://doi.org/10.1007/s40171-021-00265-x.

Sherwin, K. (2019). Usability Heuristic 5: Error prevention (Video). World Leaders in Research-Based User Experience. https://www.nngroup.com/videos/usability-heuristic-error-prevention/.

Sindre, G., Natvig, L., Jahre, M., 2009. Experimental validation of the learning effect for a pedagogical game on computer fundamentals. IEEE Trans. Educ. 52 (1), 10–18. https://doi.org/10.1109/TE.2007.914944.

Skinner, B., 1965. Science and Human Behavior. Macmillan, New York. https://doi.org/10.1016/B978-012370509-9.00087-5.

Tao, Y., 2005. Work in progress - Introducing usability concepts in early phases of software. In: *Proceedings - Frontiers in Education Conference, FIE, 2005*, pp. 19–20.

Tondello, G.F., Wehbe, R.R., Diamond, L., Busch, M., Marczewski, A., Nacke, L.E., 2016. The gamification user types hexad scale. In: Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play, pp. 229–243.

von Wangenheim, C.G., Shull, F., 2009. Voice of evidence To Game or Not to Game? IEEE Softw. 26 (2), 92–94.

Wahl, N.J., 2000. Student-run usability testing. In: Software Engineering Education Conference, Proceedings, pp. 123–131. https://doi.org/10.1109/CSEE.2000.827030.

Wilson, P., 1973. Situational relevance. Inf. Storage Retrieval 9 (8), 457–471. https://doi.org/10.1016/0020-0271(73)90096-X.

Zainuddin, Z., Chu, S.K.W., Shujahat, M., Perera, C.J., 2020. The impact of gamification on learning and instruction: a systematic review of empirical evidence. Educ. Res. Rev. 30 (March) https://doi.org/10.1016/j.edurev.2020.100326.

Zichermann, G., Cunningham, C., 2011. Gamification By design: Implementing game Mechanics in Web and Mobile Apps. O'Reilly Media, Inc. http://storage.libre.life/Gamification_by_Design.pdf.