Investigating the Effects of Tailored Gamification on Learners' Engagement over Time in a Learning Environment

AUDREY SERNA, University of Lyon, INSA Lyon, CNRS, UCBL, LIRIS, UMR5205, France
STUART HALLIFAX, University of Lyon, University Jean Moulin Lyon 3, iaelyon school of Management, CNRS, INSA Lyon, UCBL, LIRIS, UMR5205, France
ÉLISE LAVOUÉ, University of Lyon, University Jean Moulin Lyon 3, iaelyon school of Management, CNRS, INSA Lyon, UCBL, LIRIS, UMR5205, France

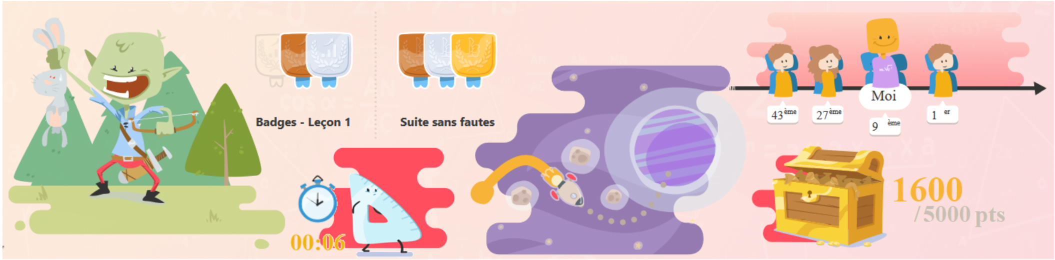

Fig. 1. Examples of the six game elements used in the experiment.

Gamification has been widely used to increase learners' motivation and engagement in digital learning environments. Various studies have highlighted the need to tailor gamification according to users' characteristics. However, little is known about how tailoring gamification affects learners' engagement when interacting with the environment. In this paper, we analyse learners' behaviours in a large-scale field study in real-world classroom conditions over a six-week period. We identify three behavioural patterns and show at a global level that two of these patterns are influenced by adaptation. When we look at how learners' engagement evolves over time, we see more differences in the adapted condition, specifically in the final lessons of the experiment. Globally learners' engaged behaviours gradually decreased over time but tailoring the game elements to learners seemed to reduce this decrease or make it more stable, depending on the behavioural patterns.

CCS Concepts: • **Human-centered computing** → *Empirical studies in HCI*; *User models*; • **Applied computing** → Computer games.

Additional Key Words and Phrases: Tailored Gamification, Digital learning environment, Behaviour analysis, Engagement, Behaviour patterns

Authors' addresses: Audrey Serna, audrey.serna@insa-lyon.fr, University of Lyon, INSA Lyon, CNRS, UCBL, LIRIS, UMR5205, 19 avenue Jean Capelle, Villeurbanne, -, France, F-69621; Stuart Hallifax, stuhallifax@live.co.uk, University of Lyon, University Jean Moulin Lyon 3, iaelyon school of Management, CNRS, INSA Lyon, UCBL, LIRIS, UMR5205, 1C avenue des Frères Lumière - CS 78242, Lyon, -, France, F-69008; Élise Lavoué, elise.lavoue@univ-lyon3.fr, University of Lyon, University Jean Moulin Lyon 3, iaelyon school of Management, CNRS, INSA Lyon, UCBL, LIRIS, UMR5205, 1C avenue des Frères Lumière - CS 78242, Lyon, -, France, F-69008.

2573-0142/2023/11-ART384 $15.00
https://doi.org/10.1145/3611030

ACM Reference Format:
Audrey Serna, Stuart Hallifax, and Élise Lavoué. 2023. Investigating the Effects of Tailored Gamification on Learners' Engagement over Time in a Learning Environment. *Proc. ACM Hum.-Comput. Interact.* 7, CHI PLAY, Article 384 (November 2023), 25 pages. https://doi.org/10.1145/3611030

1 INTRODUCTION

Gamification relies on the use of game design elements in non-game contexts [18]. Over the last decade, gamification has been widely integrated in learning environments as a way to stimulate learners' motivation, engagement and user experience as shown in recent systematic reviews on gamification within the educational domain [7, 19, 61, 65, 72]. These meta-analyses identify empirical studies that investigate the effects of gamification mostly on learners' motivation, performances and engagement. The results of these studies tend to show positive effects [7, 65, 72]. Zainuddin et al. [72] particularly underline the importance and the role of engagement in learning outcomes, as they observe that values of engagement and motivation were always positively correlated with academic performance. However, other reviews underline contradictory or mixed results. For instance, Sailer et. al [61] show that results on motivational and behavioural learning outcomes cannot be interpreted as stable due to different factors.

These mixed results of gamification have also been observed in other domains [36, 37] and researchers have started to investigate if adapting game elements to individual learners can have a greater impact. Several systematic literature reviews analyse the current results of such studies. Rodrigues et al. [57] identify primary studies testing personalisation approaches and report mixed results for motivational, behavioural and cognitive outcomes. Oliveira et al. [49] also highlight that most of the experiments do not provide sufficient statistical evidence. Several of these studies point out the fact that the impact depends on individual factors, such as the player profile [29, 42, 45, 51] or learners' motivation for the learning task [5, 26, 40]. In addition to these individual static factors, Klock et al. [36] underline the importance of considering the dynamic and cyclical nature of gamification to improve user experience in the long-term. They highly recommend investigating "how users may change from time to time and how interaction evolves" to periodically update these dynamic models. This need for a deeper understanding of users' experiences is also highlighted by other meta-analyses [49, 57].

To date, most of the existing studies are focused on the effects of tailored gamification after using the learning environment and do not analyse behavioural outcomes while performing the learning activity. Thus, to achieve a better understanding of these aspects, we propose to analyse the impact of tailoring game elements integrated in a learning environment on learners' engagement across learning sessions and its evolution over time. To investigate this, we conducted a large-scale field study in real-world classroom conditions, involving 145 students (aged between 13-15 years) from French secondary schools. The students completed 6 maths lessons of approximately 40 minutes in a gamified learning environment. We performed a factor analysis on the interaction logs collected during the experiment to identify behavioural patterns when using the gamified digital learning environment.

We also compared the levels of engagement between learners who used a tailored game element and those of learners who used randomly assigned ones. Our results reveal that (1) we can distinguish three behavioural patterns corresponding to different kinds of engaged behaviours (2) learners' engaged behaviours gradually decreased over time but (3) tailoring the game elements to learners seemed to reduce this decrease or make it more stable, depending on the behavioural patterns, as learners in the non-tailored condition showed a higher decrease than those in the tailored condition (specifically with regard to two of the three identified patterns).

We believe that these findings contribute to a better understanding of how tailoring gamification to learners can impact their engagement during a gamified learning activity. We derive design implications for tailored gamification in education. We also provide insights into how engagement should be tracked and evaluated using a more granular approach. This is a first step towards the dynamic adaptation of gamification and its automation, which is an important challenge for future research in the field of adaptive gamification [49].

2 RELATED WORK

2.1 Impact of Gamification on Learners

Gamification has gained interest in recent years in several domains. Different recent systematic reviews of literature intended to summarise the results obtained in various studies analysing the effects of gamification in general [30, 37, 63] and specifically in the educational domain [7, 19, 61, 65, 72].

The literature reviews conducted between 2014 and 2015 reported rather positive results on motivation, engagement and enjoyment [19, 30, 63], thus supporting its potential for beneficial effects. However, they also underlined some issues regarding the lack of theoretical foundations and the design of these empirical studies. In particular, most of the studies did not isolate the effects of gamification and specific game elements [63], and the effects were greatly dependent on the context and users [30]. In their meta-review in education, Dicheva et al. [19] reported a majority of positive results, including significantly higher engagement of students with increased participation and motivation. They also pointed out some major obstacles and needs, in particular the need for more substantial empirical research to demonstrate reliable results using specific game elements.

Since then, several studies have tried to analyse the effect of specific game elements on different aspects of user experience and performances. In education, these aspects are evaluated in terms of cognitive, motivational and behavioural outcomes, with rather stable results for cognitive learning outcomes and less stable results for motivational and behavioural outcomes [61]. The positive effects of gamification on learner performances is reported in several meta-analyses [7, 65, 72]. Regarding motivation and engagement, several studies report positive effects of gamification [11, 20, 21, 71], as underlined by the review conducted by Zainuddin et al. [72]. Other studies report mixed results. For instance, Landers et al. [38] demonstrated the effectiveness of leaderboards for simple tasks where they served as a goal setting tool for users. However, their effectiveness decreased as task difficulty increased. Also, Denny et al. [15] tested the effect of badges and scores on learner behaviour and found that only badges had an effect.

Even if these results are rather positive, some variability was underlined by Sailer et al. [61] and Koivisto and Hamari [37], who state that "while the results in general lean towards positive findings [..] the amount of mixed results is remarkable". The authors recommend investigating the role of the users, their goals and their individual attributes on the effectiveness of gamification. In fact, the majority of the meta-analyses cited above agree that the effects of gamification are highly dependent on specific contexts, which vary across individuals. Based on this assumption, a recent approach investigated the adaptation of gamification to users' characteristics.

2.2 Tailoring Gamification to Learners: Methods and Results

Tailored gamification is a current trend that consists in taking into account users' inter-individual differences when gamifying a system [6, 55]. Several systematic literature reviews [29, 36, 49] show that tailored gamification studies have mainly explored ways to model the user profile. These studies usually rely on statistical techniques to correlate and test the most suitable game elements for each user characteristic. The characteristics most commonly considered in the user profile are player preferences (45% according to [36]), gender and personality traits. Regarding player preferences, a recent comparative study of several typologies showed that the Gamification User Types Hexad framework [43] was the one most suited to express player preferences towards game elements [28]. Indeed, Hexad was specifically developed for gamification and is based on the Self Determination Theory [59]. The Hexad typology distinguishes six different user types: Philanthropists, Socialisers, Free Spirits, Achievers, Players, and Disruptors. Several recent experimental studies were conducted in order to identify the motivational impact of game elements regarding this typology [52] or propose personalisation approaches [46].

Another systematic literature review on personalised gamification [57] focuses on the comparison of the effects of adapted gamification with one size fits all gamification. The authors identify primary studies testing personalisation approaches and report mixed results for motivational, behavioural and cognitive outcomes alike.

Regarding more specifically the field of education, tailored gamification has also gained interest because learners have different motivations when using a learning environment [30, 45, 70]. Recent meta-analyses on tailored gamification in education [5, 29, 49] confirm that most of the studies adapt gamification using player preferences thanks to statistical approaches. However, few studies proposed to adapt gamification to learners' behaviours, such as [53] using students' interactional profile or to learners' motivation, such as Roosta et al. [58] and Hassan et al. [32], who use various forms of user task motivation as a basis for adaptation. In addition, in their literature review on adaptive gamification in e-learning, Bennani et al. [5] identify several studies arguing that future adaptation improvement should be based on motivation considerations.

Regardless of the adaptation process, the meta-analyses also identify mixed results and highlight that most of the experiments do not provide sufficient statistical evidence. Roosta et al. [58] showed significant differences in motivation, engagement (participation rates) and quiz results between learners who used a tailored (to their motivation) or a randomised game element. Mora et al. [46] reported an increase in students' behavioural and emotional engagement when adapting to their Hexad player type. More mixed results are reported. For example, Monterrat at al. [45] showed that the adaptation process had a negative impact on the perceived usefulness and fun of gaming features. However, when performing a similar study, Lavoué et al. [40] found that providing learners with tailored game elements made the most engaged learners spend significantly more time in the learning environment. In another study, Paiva et al. [53] analysed the usage data during the month after the introduction of tailored goals in their learning tool used for learning mathematics. Learners received personalised goals to encourage them to increase the number of specific learning actions they performed. The results showed that only specific goals were effective in increasing the number of related actions. Finally, Oliveria et al. [50] found no differences between tailored and non-tailored conditions on learners' flow experience. This finding is somewhat contradictory with a previous study [48] in which the same authors found that, for some player types, the tailored system induced better learner concentration than the counter-tailored one, while for other player types the counter-tailored game elements functioned better. More recently, [26] simulated adapting game elements to learners based both on their Hexad profile and their initial motivation for the learning content. They proposed an algorithm to combine the recommendations issued from each profile into a unique recommendation. This approach resulted in an adaptation that was more effective on learners' motivation and engagement than the one based on a unique profile.

To conclude, these studies report mitigated results when tailoring gamification to learners, thus highlighting the need for more empirical studies on this recent research issue [36, 49, 57]. Moreover, as reported by Hallifax et al. [29], studies are generally focused on the effects of tailored gamification on learners' motivation and performances after using the learning environment, but do not investigate if and how tailoring game elements affects their behaviours while performing the learning activity. This is also underlined by Klock et al. [36], who state that tailored gamification could be improved by considering how learners' interactions with the learning environment evolve from time to time. In this paper, we aim to fulfil the need for more understanding on this issue by analysing learners' engagement while performing the activity.

2.3 Analysing Learners' Engagement

Engagement is a complex and multidimensional process. O'Brien & Toms [47] define it as "a quality of user experience characterised by attributes of challenge, positive affect, endurability, aesthetic and sensory appeal, attention, feedback, variety/novelty, interactivity, and perceived user control". It is particularly a dynamic process, with engagement and disengagement, that therefore changes over time. It is generally accepted that engagement is composed of three complementary dimensions: *cognitive*, *motivational (or affective)*, and *behavioural* [25, 41]. Motivational engagement covers the interest, emotions, and values perceived by learners during learning activities. Cognitive engagement is related to the deployment of learning strategies: cognitive, self-regulated or resource management-related [54]. Behavioural engagement refers to the observable actions of the learner in completing a learning task [25]. Considering the recommendations of Klock et al. [36], we are particularly interested in this latter type of engagement.

Many studies have focused on analysing students' engagement in learning environments, whether gamified or not. Different approaches coexist, generating a large variety of indicators for measuring the level of engagement of students. For instance, da Rocha Seixas et al. [11] defined engagement indicators (such as autonomy or participation) with ethnographic methods, using observation and semi-structured interviews. Thanks to a survey completed by students at the end of the experiment, they also performed a cluster analysis to classify student engagement. Many other studies prefer to adopt the gamification analytics approach [34]. They rely on student activity (i.e. actions in the gamified application) to measure behavioural engagement, extracting indicators from data logs. Generally, these indicators correspond to the time spent in the learning environment, number of logins, content completion rate, completed test rate, number of questions answered correctly, etc. [66]. For instance, Ding et al. [20] measured student behavioural engagement through students' number of entries and frequency of logins. Rodrigues et al. [56] used the number of attempts that students made, the time spent on the environment and the number of system accesses to measure students' intentions to practice, consult learning materials, and interact with the game elements. Usually, these studies present a cumulative view of the indicators, without temporal analysis sessions after sessions. Recently, some longitudinal studies have examined this temporal aspect, such as Barata et al. [3] who used performance and participation measures to identify clusters of students and analyse their evolution over three years. Tacskin et al. [66] examined the role of gamification on the behavioural engagement of students, using the number of logins, time spent in the environment and content review rate over ten weeks, and observed the distribution of the indicators over time.

Finally, some studies started to investigate correlation between these indicators, defining behavioural patterns or engagement models. For instance, Codish et al. [9] defined gamification behavioural patterns as sequences of actions performed by a user that can be extracted from the data logs. Another interesting method is presented in Fincham et al. [24], where the authors use a factor analysis to establish an engagement model based on various simple indicators. The indicators are inspired by the review performed by Joksimović et al. [35] and are derived from data logs, referring to behavioural or academic engagement (e.g. days active, question accuracy). From all these metrics, the authors built a final engagement model based on three factors of engagement. More recently, Lavoué et al. [39] studied the data logs of a gamified learning environment to identify and analyse a model of engagement using the same approach. They identified two types of engaged behaviours (achievement-oriented and perfection-oriented) and showed that the learner profile influences these observed behaviours.

To sum up, to investigate behavioural engagement over time, in this study we follow approaches based on the analysis of learners' interactions with the learning environment. According to recent approaches based on factor analysis, we believe that indicator-based exploratory methods offer a promising approach for a comprehensive evaluation of learners' engagement.

3 RESEARCH QUESTIONS

In this paper, we aim to enrich our knowledge of the impact of tailored gamification on learner experience and behaviour in education, as suggested in the research agenda of different literature reviews on tailored gamification [36], specifically for the educational field [29, 57]. We propose to investigate how tailoring game elements to learners affects their engagement during the use of a gamified learning environment, relying on the analysis of learners' interactions and exploratory approaches based on factor analysis to identify behavioural patterns. Based on the literature, we hypothesise that tailoring game elements to both learners' player profile (relying on Hexad typology [43], widely used in the gamification literature) and to initial motivation for the learning task, can lead to an increase in engagement as compared to non-tailored game elements. Furthermore, considering the dynamic nature of engagement, we propose to investigate how learner engagement evolves over time, and we hypothesise that tailoring game elements will influence this evolution. Thus, we address the following research questions:

- **RQ1:** Are there some behavioural patterns emerging from learners' interactions with the gamified learning environment? We explore the use of data logs to build a latent variable model structure based on an exploratory factor analysis.
- **RQ2:** How does the adaptation of game elements affect learners' behaviours? We compare the average scores of each pattern identified in answer to RQ1 between two conditions: an experimental group of learners provided with tailored game elements and a control group with randomly assigned ones.
- **RQ3:** How do learners' behaviours evolve over time depending on the adaptation of game elements? We explore the differences between the scores of behavioural patterns identified in answer to RQ1 at each lesson between the two conditions, as well as the differences between lessons for each pattern.

4 GAMIFIED LEARNING ENVIRONMENT

Learners used the gamified learning platform, LudiMoodle a modified version of the Moodle Learning Management System (see Figure 2). This platform was developed within the scope of the LudiMoodle project, which brings together researchers in computer science and in educational sciences, pedagogical designers, four middle schools, a Moodle development company.

4.1 Learning Content

All of the learning content was created by the participating teachers so that it would be as close as possible to their teaching practices. The teachers designed six lessons, composed of several quizzes (4 to 10) that covered the topic of secondary school level basic algebra (in particular literal arithmetic). The quizzes were designed as *training exercises* since teachers had observed that learners generally found these exercises to be boring or too repetitive, and they wanted to make these exercises more engaging for learners. Within a lesson, to successfully complete a quiz and progress to the next one, learners had to answer at least 70% of all questions correctly. Otherwise they had to start the quiz again.

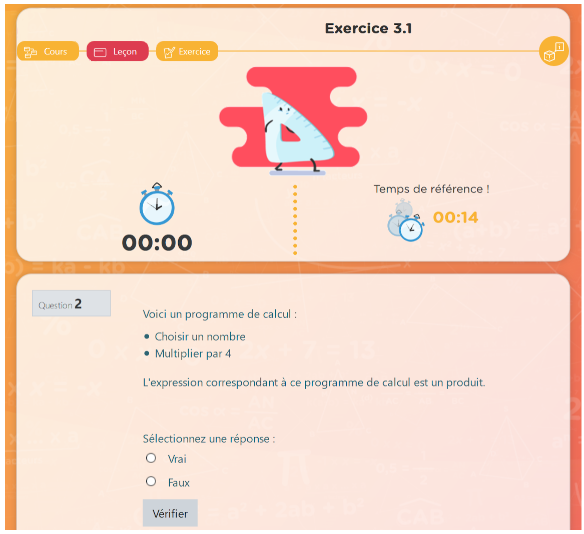

Fig. 2. A screenshot of the LudiMoodle learning platform showing the Timer game element and a question in a quiz.

4.2 Game Elements

To gamify the learning platform, we used an iterative design process with participatory design sessions with the different stakeholders of the project (teachers, game designers, educational engineers, the company in charge of development), using the design method and design space presented in [27].

First, following the same process as in [68], we ensured that we selected game elements that would appeal to different learners by covering the main game elements identified in the literature.

The most frequently used game elements in studies in education are points, badges and leaderboards [65, 72], followed by avatars and progress bars or levels [3, 72]. In tailored gamification, the most commonly used game elements, all fields considered, are, in order: customisation (i.e. avatar), badge, challenge, level, competition, leaderboard and points [36]. We compared these common elements with the classification proposed in the design space of [27] to select game elements that would cover the different game dynamics. We selected points and badges (for the Reward dynamic), progress bar and ranking (corresponding to a mix between leaderboards and competition) (for the Progress dynamic), and avatar (for the Self representation and customization dynamic). In addition, teachers selected timer (for the Time pressure dynamic) for its suitability for the quiz format, although this element has been rarely studied in the literature [8, 44]. We verified that each Hexad user type had several suited game elements, in line with both the recommendations of the Hexad framework and the positive influences observed in related studies. Finally, we ensured that the different game elements covered the different psychological needs as defined in the Self Determination Theory (SDT) [12, 13]. This well-established psychological theory is the one most frequently used in gamification research, as underlined by different literature reviews [63, 72]. It distinguishes several types of motivation (intrinsic or extrinsic) and argues that human beings are intrinsically motivated to engage in activities that satisfy three innate, universal psychological needs, which are competence (sense of efficacy), autonomy (volition and personal agency), and relatedness (social interaction). Most design approaches rely on this theory to implement game elements that intend to satisfy these needs [17].

The learners themselves were not directly involved in the design of the gamified platform. However, similar learners (same age and class level), who had interacted with previous prototypes of the platform, participated in focus groups to refine the design and functionalities of the game elements.

4.2.1 Points: Each correct answer given by the learners awarded them points (generally 100 points depending on the difficulty of the question). These points were displayed using a sack or chest of gold coins (the points corresponding to a quiz were shown using a sack, and those to a lesson using a larger chest). Learners were also shown the maximum number of points they could score for each lesson and quiz. Points generally matched the need for competence regarding the SDT [60, 72]. As this game element gives learners a clear representation of how well they are doing in the course and rewards them for performing better, it is generally given to Players [36, 43].

4.2.2 Badges: For each lesson, learners could earn two categories of badges, one based on how many questions they correctly answered in a row, and one based on how much of the lesson they completed. Each of these badges came in a bronze-silver-gold version based on how well learners achieved these goals. In general, these badges were awarded if the learners completed respectively 70-85-100% of the quizzes in a lesson. There was also a set of "medals" for each quiz in the lesson, so that learners could easily identify which quizzes they needed to try again if they wanted to earn gold in each lesson. Badges match the need for competence regarding the SDT [60, 72] and are generally shown to be motivating for all types of users [28] and explicitly for Players [36] and Achievers [43].

4.2.3 Avatar: The avatar game element showed a goblin-like character that learners could personalise with various clothes and equipment. As the learners progressed in a lesson, they could unlock a different set of objects to use (e.g. medieval, fairy tale, pirates). The avatar could be personalised via an inventory menu displayed near the top of the game element. In general, when learners achieved the required 70% in each quiz in a lesson, they could unlock 1 or 2 objects. Avatar matches the need for autonomy regarding the SDT [60, 72]. This kind of game element is generally recommended for Free Spirits, as it provides them with a personalised representation of themselves [36, 43], and for Disruptors [36].

4.2.4 Progress Bar: Learners were shown their progress in the quizzes by way of a rocket ship that travelled from earth to various planets. Each correct answer would charge a "boost" meter that, when filled, would propel the rocket further. When learners achieved a full 100% of correct answers for a lesson, they would reach the planet. This game element matches the need of competence regarding the SDT [72] and should be particularly interesting for Achievers as they have a clear goal. According to the Hexad framework, progress can support all user types [43]. It is explicitly attributed to Players in several studies [36] and can appeal to Disruptors [28].

4.2.5 Ranking: This game element is a combination between leaderboards and competition as defined in [36]. Learners were shown their position in a fictional "race" against other learners. This position was decided based on their answers (i.e. the more questions they answered correctly, the higher they were ranked in the race). Our initial idea was to show learners their ranking as compared to their classmates. However, teachers explained that they did not want a direct comparison with learners in the same class. We therefore made the compromise to compare the learners to a fictional class and told them that they were comparing themselves to previous years' attempts (thus still providing a sense of competition). This fictional class was set up so as a few learners scored 100% on all quizzes and all learners were at least in the top half of the ranking. As this game element allows learners to compare themselves to others (even if fictional), it should be motivating for Socialisers [36] (related to the relatedness need identified in the SDT [72]) and Players [36, 43]. Other studies report positive influences of ranking on the Disruptor player type [28, 36].

4.2.6 Timer: This game element showed a timer for each quiz. Each of the questions was timed and recorded. Learners were shown the average time taken to answer previous questions and, each time they beat this "reference time", a small maths related character ran faster and faster. Learners were thus encouraged to make their character as fast as possible by answering questions quickly. They were only rewarded for correct answers, as an incorrect answer would not affect the reference time or the animated character. With this functioning, learners are challenged to beat themselves in a race (matching the competence need related to the SDT). According to the Hexad framework, mechanics related to time pressure can support all user types [43].

5 STUDY DESIGN

5.1 Procedure

Before using the learning environment, we first asked learners to fill out two questionnaires to establish their Hexad player profile and their initial motivation for doing maths exercises (see section 5.3). Learners were then sorted into one of two experimental conditions. In the control group, learners were randomly assigned a game element. In the experimental group, learners were assigned a game element tailored to their profile. This assignment was based on an algorithm that determines the most relevant game element for each learner based both on their player type and motivation (see Section 5.4, the adaptation algorithm was previously tested in [26]). Only one condition was assigned to each class so that the users were homogeneous within the same class. Once a game element was assigned to learners, they used it for the entire duration of the experiment.

Then, learners followed the lessons either once or twice a week for four to six weeks. They accessed the quizzes individually using a tablet device. Each lesson was carried out in the same way: teachers gave a short introduction to the lesson's topic (10-15 minutes depending on the complexity). Learners then logged into the gamified learning platform to solve quizzes related to this lesson (25-30 minutes).

5.2 Participants

A total of 236 learners split over nine classes in five different secondary schools participated in our experiment. We filtered participants to only keep those who were present in at least four of the six lessons, leaving us with a total of 145 learners (72 self-reported as female, 73 as male, aged between 13-15 years). Regarding experimental conditions, 93 learners used a tailored game element (including those who were randomly assigned a tailored one), while 52 used a non-tailored game element.

5.3 Profile Questionnaires

Learners filled out both the Hexad [67] [1] and AMS [69] questionnaires to determine their profile (player and initial motivation for mathematics). The motivation scale proposed by Vallerand et al. [69] is inspired by the Self-Determination Theory (SDT) [14] and is especially designed for Education. It evaluates seven types of motivation: three for intrinsic motivation (IM), three for extrinsic motivation (EM) and one for amotivation. Each of these types identifies the reasons why

[1] we used the French version proposed here https://hcigames.com/wp-content/uploads/2019/11/Hexad-Survey-and-Instructions_FR.pdf

someone would perform an educational activity, in our case learning maths. We provide the details of the questionnaires in the supplementary materials.

5.4 Adaptation Algorithm

The platform we developed allowed us to assign a different game element to each learner during the initialisation step. Using the questionnaires filled out by learners during the pre-test, we were able to create a learner profile and therefore to propose the game element best suited to this profile. To determine which game elements to assign to learners, we used the adaptation algorithm proposed in [26], which considers both the game element that would be recommended for the player profile and the game element relevant to the initial motivation of learners. It takes into account the values of all dimensions of each profile and therefore learners' fine-grained preferences for game elements. This new approach is supported by the recent literature review on tailored gamification [36], which recommends that adaptation should not be based solely on one aspect of the users' characteristics. This method showed better results than adapting to either the Hexad profile or the initial motivation for mathematics separately on a simulated environment [26]. To establish relationships between the game elements and the dimensions of each profile, we used the data from 258 participants using the LudiMoodle platform during a study that we conducted in the same context (same learning content and game elements, same age learners). We chose to rely on field data rather than considering only pre-existing recommendations from the literature, because previous work insists on the fact that the context plays a major role in the impact of game elements on user motivation [28].

Figure 3 presents an overview and example of how this adaptation works (the algorithm is detailed in the supplementary materials). We used the results of previous statistical analyses (partial least square path modelling) that linked game elements with the different profile dimensions. These results were used to create an "influence matrix" for each of the profiles (Hexad and initial motivation). These matrices make it possible to code the positive or negative influences of each profile dimension according to each game element. We then multiplied the individual learner profiles with these influence matrices, which gave us two "affinity" vectors for each learner (one based on each profile). These affinity vectors showed how well suited each game element should be for a given learner. Finally, we used an algorithm that first checks for positive affinity overlaps between the two affinity vectors. If any game elements have positive affinities in both vectors, they combine the ranks of these game elements, selecting the lowest ranked game element. If there is no game element in this overlap, it then combines the rankings of the affinity vectors for all game elements (again selecting the lowest ranked). If at any point there is a tie for lowest ranked, it adds the affinities from both vectors and selects the game element with the highest affinity.

For example, a learner with the Hexad profile (Pl:0; Ac:-8; So:2; FS:0; Di:6; Ph:7), would have the following affinity vector ('Avatar': .385, 'Badges': .0364, 'Progress': -.241, 'Leaderboards': -.920, 'Points': -.577, 'Timer': .225) and would therefore be recommended the Avatar game element. A learner with the following initial motivation for mathematics (Mico:9; Miac:11; Mist:10; ExtReg:12; IdReg:7; IntReg:8; Amot:8) would have this affinity vector: ('Avatar':-6.188; 'Badges':-42.22; 'Progress':2.871; 'Leaderboards':-0.899; 'Points':-50.899; 'Timer':-23.807). As there is no positive overlap in these vectors, we combine the rankings for all game elements resulting in ('Avatar':4; 'Progress':5; 'Timer':6; 'Leaderboards':8; 'Badges':8, 'Points':11) and would therefore recommend the Avatar game element (see Figure 3).

5.5 Engagement Indicators

Learners' interactions with the learning environment were tracked using the Moodle data logging system. We followed the same process as previous works defining indicators extracted from data logs [9, 39, 56]. We defined indicators of how learners were interacting with the gamified system.

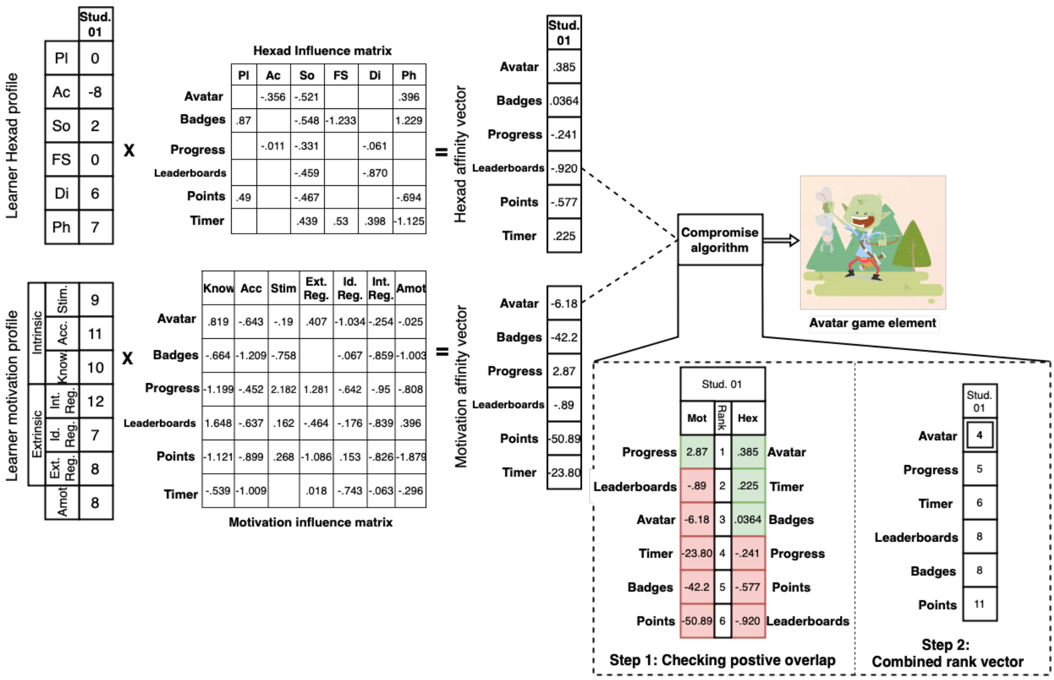

Fig. 3. Learner adaptation process. Both the learners' Hexad and initial motivation profiles are used to adapt. We use the compromise algorithm described in [26], fully described in the supplementary materials. In this example, we show how Student 01's profiles are used to recommend the **Avatar** game element: both the Hexad and Motivation influence matrices are obtained by adding the individual partial least squares result matrices for each game element. There is no positive overlap in the affinity vectors, so we combine the rankings for each game element.

These indicators represent behavioural outcomes related to quiz completion, feedback on answers and interaction with the game element. These types of indicators are generally used to measure accuracy, time on task, number of exercises completed [24, 39, 56, 66]. For some of the indicators, we decided to use ratios (obtained by dividing the count by the number of quizzes/questions attempted) instead of direct counts, since learners did not attempt the same number of quizzes. All these indicators are calculated either for all lessons (for answering RQ1 and RQ2) or for a given lesson (for RQ3):

- AvgQuestionTime: average time taken to answer a question (for all question attempts), calculated for the first attempt at each question (as questions did not change on successive attempts).
- NQuiz: total number of quizzes attempted for all lessons.
- NPassedQuiz: number of quizzes successfully completed (i.e. where a learner scored more than 70% on the quiz). Each quiz was counted only once when it was first successfully completed.
- NLoop: number of quizzes restarted after successfully completing them. Learners could gain more points, badges, etc. if they scored higher than the minimum required, and were therefore incentivised to restart a quiz that they completed only at 70%.

- QuestionRatio: correct question ratio (i.e. number of correct answers divided by the total number of questions in a quiz), averaged over the total number of quizzes attempted. It was calculated for the first attempt at each quiz.
- StreakRatio: average number of quizzes attempted in a session without restarting a successfully completed quiz, divided by the number of quizzes attempted in the session. For example, if Learner L1 completed Q1, Q2 and Q3, then restarted Q3 to get a higher score, and finally completed Q4 and Q5, they would have a "StreakRatio" of 3/5 = 0.6.
- LessonRatio: number of quizzes completed during a lesson divided by the total number of quizzes attempted during a lesson. For example, a learner who completed 5 different quizzes during a lesson and attempted 7 total quizzes, would have a *LessonRatio* of 71.42%.
- RestartedQuizzesRatio: ratio of successfully completed quizzes that were restarted. For example, if Learner L1 completed the following quizzes: Q1:70%, Q2:85%, Q3:50%, and then restarted Q1 2 times, Q2 3 times, and Q3 1 time, they would have a "RestartedQuizzesRatio": 2/2 = 1 (2 restarted quizzes out of 2 completed quizzes).
- FeedbackTime: time learners spent on the feedback page that was showed after submitting a quiz. On this page learners could see which questions they got wrong and prepare for restarting the quiz.
- GameElementInfo: number of times a learner clicked on the game element information popup. It was automatically shown at the start of a new lesson, and learners could reopen it at any time by clicking on an information button.

5.6 Analysis Procedure

To answer RQ1, the engagement model was obtained using the data from all the 145 participants who were filtered using the process defined in section 5.2, following a workflow inspired by [39]:

(1) We split our data into two same sized random samples (training and test datasets).
(2) We conducted a parallel analysis scree plot on the first half of the dataset (training set). This gives an indication of how many factors we should look for in the following steps. In brief, a parallel analysis involves the generation of a random dataset of the same dimensions as the data being analysed. Factor analysis is then performed on the random data to extract eigenvalues. To avoid bias, this process is repeated 20 times, and an average is taken for each eigenvalue. These random eigenvalues are then compared with the eigenvalues of the real data, and factors in the real data are only retained if their eigenvalues are greater than the eigenvalues from the random data [33].
(3) We ran an exploratory factor analysis (EFA) [23] using the recommended number of factors from step 1 on the training set. We conducted this process in an iterative manner, whereby variables that did not load or exhibited factor loadings greater than 1 were removed (as in [10]). We also defined a cut-off of 0.5 so that our factors were clearly and strongly defined [4].
(4) Finally, we assessed the fit of this model using a confirmatory factor analysis (CFA) on the second half of the dataset (test set).

These analyses were run using the Lavaan package in R. For the inital EFA, we used oblimin rotation to allow for correlations between factors, and given the relative normality of our data, standardised coefficients were estimated using maximum likelihood [10]. This permitted the computation of a wide range of goodness of fit indexes and allowed testing for the significance of factor loadings and correlations, as well as the computation of confidence intervals [22]. For the final CFA we used the DWLS (diagonally weighted least squares) estimator.

To answer RQ2 and RQ3, we then used a Wilcoxon rank sum test, since the data were not normally distributed and we calculated the effect sizes. We compared the scores obtained for each

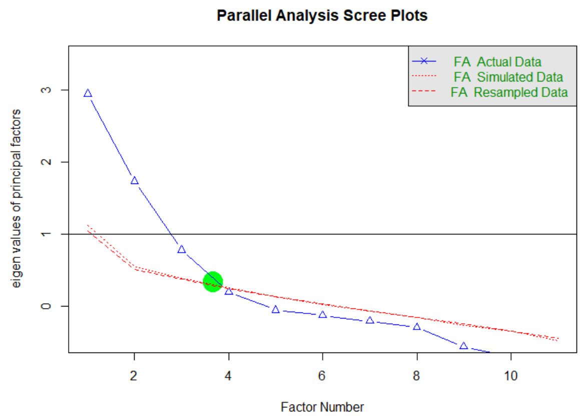

Fig. 4. Parallel analysis scree plots of exploratory factor analyses, where the blue line shows the scree plot of the first half of our dataset, and the red line shows the scree plot of random data of the same size. The "elbow" of the graph is highlighted in green.

factor defined in answer to RQ1 between the tailored and non tailored conditions for all lessons (RQ2) and for each lesson (RQ3). We then used a pairwise Wilcoxon signed rank test to compare the evolution of learners' engaged behaviours lesson after lesson in each condition. For this final comparison, we restricted our sample to the learners who were present in all of the lessons (99 learners in total) to be able to compare lessons in pairs.

6 RESULTS

6.1 RQ1: Behavioural Patterns

First, a parallel analysis using the training set suggested a three-factor structure. Figure 4 shows the scree plots that resulted from this analysis. The "elbow" of the graph where the eigenvalues seem to level off, and factors or components to the left of this point should be retained as significant.

We then ran an EFA on the same training set using the suggested three-factor model. We chose a cutoff value of 0.5 for factor loadings so as to be sure to retain only the most influential indicators. The EFA provided us with a three-factor model that uses 9 different indicators out of the 10 defined in Section 5.5 (TLI: 0.82, RMSEA: 0.148 (90% CI: 0.105 0.195), SRMR: 0.06). The standardised loadings for this model are presented in Table 1. We then performed a CFA using this model on the other half of the original data (test set). This also resulted in a fairly good fit to the dataset (χ^2(23, N = 73) = 53.137, p-value = 0.000, CFI = 0.903, TLI = 0.848, RMSEA = 0.135 (90% CI: 0.087-0.183), SRMR = 0.135). The estimated standardised solution (standardised loadings) and p-values (for testing the null hypothesis that the loading equals zero) may be found in Table 2.

In answer to RQ1, these results show that we can identify behavioural patterns from learners' interactions with the gamified learning environment. We identify three patterns that we believe correspond to three kinds of engaged behaviours [2] :

F1: composed positively of *NQuiz*, *GameElementInfo*, *NPassedQuiz* and negatively of *AvgQuestionTime*.

This behavioural pattern represents how fast learners answered questions (negative loading of average question time) and therefore how many quizzes they attempted on the platform

[2]The factors are colour-coded throughout the paper and figures to improve readability: these colours have no other meaning

Table 1. EFA Standardised Loadings

Indicator	Factor1	Factor2	Factor3
AvgQuestionTime	-0.681		
NQuiz	0.961		
NPassedQuiz	0.630		0.643
NLoop		0.923	
QuestionRatio			0.953
StreakRatio		-0.520	
LessonRatio			0.743
GameElementInfo	0.507		
RestartedQuizzesRatio		0.864	
FeedbackTime			

Table 2. CFA Standardised Loadings

		Load.	Std err	z	p
F1	Nquiz	0.123	0.014	8.669	0.000
	GameElementInfo	0.05	0.018	2.867	0.004
	AvgQuestionTime	-0.017	0.002	-8.658	0.000
	NPassedQuiz	0.043	0.008	5.155	0.000
F2	Nloop	0.45	0.057	7.933	0.000
	StreakRatio	-0.124	0.017	-7.445	0.000
	RestartedQuizzesRatio	0.092	0.014	6.408	0.000
F3	QuestionRatio	0.043	0.011	3.804	0.000
	NPassedQuiz	0.044	0.014	3.116	0.002
	LessonRatio	0.181	0.047	3.805	0.000

(positive loading of NQuiz). This naturally led to a higher number of completed quizzes (positive loading of NPassedQuiz) - as learners were told when they provided an incorrect answer. Finally, in this type of behaviour, learners checked their game element information often.

F2: composed positively of *NLoop*, *RestartedQuizzesRatio* and negatively of *StreakRatio*.
This behavioural pattern represents the amount of repeated quizzes learners performed (positive loading for RestartedQuizzesRatio), generally directly after attempts on the same quiz (positive loading for NLoop and negative loading for StreakRatio).

F3: composed positively of *QuestionRatio*, *NPassedQuiz* and *LessonRatio*.
This behavioural pattern corresponds to high scores at each question (positive loading for QuestionRatio), obtaining a high quiz accuracy during a lesson (positive loading for LessonRatio), and therefore a high number of completed quizzes (positive loading for NPassedQuiz). As we only used the first question attempt to calculate the QuestionRatio, learners who behaved in this way got more questions right directly.

6.2 RQ2: Comparison of Behaviours between Learners with Tailored and Non-Tailored Game Elements

To answer our second research question, we compare the behavioural patterns between experimental conditions and lessons, to analyse how tailoring game elements to learners has affected their engagement. In the next two subsections we provide the scores for each factor identified in answer to RQ1, for each experimental condition. The scores are calculated using the value of each engagement indicator composing the factor ponderated by the loadings identified by the EFA.

F1 score = $-0.681 * AvgQuestionTime + 0.961 * NQuiz + 0.630 * NPassedQuiz + 0.507 * GameElementInfo$
F2 score = $0.923 * Nloop - 0.520 * StreakRatio + 0.864 * RestartedQuizzesRatio$
F3 score= $0.643 * NPassedQuiz + 0.953 * QuestionRatio + 0.743 * LessonRatio$

We clarify that we are not interested in these individual scores themselves as they are just the linear combinations of the relevant indicators, but we use them to compare the two experimental conditions.

We used a wilcoxon rank sum test to compare the scores for each factor between the tailored and non-tailored conditions considering all lessons. We found no significant differences between both conditions for **F1** (p-value > 0.05), but we did observe a significant difference for **F2** (W = 798, p-value = 0.0310) and **F3** (W = 1490, p-value = 0.0029). A further comparison of the means of **F2** for both groups showed that learners in the tailored condition scored higher than those in the non-tailored condition (Tailored **F2** = 0.066 (SD = 0.090), Non-tailored **F2** = 0.030 (SD = 0.079)). We also observe that learners in the non-tailored condition scored higher for **F3** than those in the tailored condition (Tailored **F3** = 0.308 (SD = 0.0808), Non-tailored **F3** = 0.360 (SD = 0.080)).

Therefore, with regard to RQ2, the results show a difference in learners' behaviours between tailored and non-tailored conditions. Learners in the tailored condition scored higher than those in the non-tailored condition for the pattern corresponding to **F2**, whereas learners in the non-tailored condition scored higher than those in the tailored condition for the pattern corresponding to **F3**.

6.3 RQ3: Comparison of the Evolution of Learners' Behaviours with Tailored and Non-Tailored Game Elements over Lessons

Figure 5 shows the evolution of the scores of each behavioural pattern identified in RQ1 over the sessions in both the tailored and non-tailored conditions. In this figure, the Y-axis is split presenting each pattern separately as there is no interest in comparing their values between each other, only their evolution over the sessions and between the two conditions. Table 3 shows the average values for each pattern and condition in each lesson. We performed Wilcoxon rank sum tests between the conditions on each pattern for each lesson. The results of these comparisons are presented in Table 3 (column "W" for the value of the corresponding Wilcoxon rank sum test and column "d" for the effect sizes). From a general point of view, we can notice that learners' engagement decreases over time, for both conditions and for all behavioural patterns. Using a pairwise Wilcoxon signed-rank test we tested the significant differences in the values between lessons (i.e. if the observed variations in Figure 5 were statistically significant). The p-values and effect sizes of these analyses are presented in Table 4 (these tests were corrected using a Bonferroni correction).

Regarding **F1** we observe significant differences between the two conditions in the second, fourth, fifth and sixth lessons, with more pronounced behaviours in the non-tailored condition than in the tailored condition at the start of the experiment, and conversely in the second half. Regarding the evolution of these behaviours, **F1** is the pattern that has the largest decrease, with significant differences between almost every lesson for both conditions. This means that over time, learners attempted and completed less quizzes as they progressed, spent more time on questions, and checked their game element info less.

Regarding **F2**, we cannot see any significant differences between conditions except for the final lesson, with more pronounced behaviours in the tailored condition than in the non-tailored condition. This means that learners, who used an tailored game element, restarted quizzes more and had shorter streaks than those who used a non-tailored one at the end of the experiment. Over time, both non-tailored and tailored conditions were fairly stable throughout the experiment (no significant differences between lessons). We can only observe a difference between the first and the last lessons (lessons 1 and 6).

Finally, **F3** is the pattern that fluctuates the most over time. There is a significant difference in lesson 2, and then for lessons 5 and 6. This behaviour was more pronounced in the non-tailored condition in lessons 2 and 6, in opposition to lesson 5.

Regarding the evolution in detail, in the tailored condition this behaviour was fairly stable at the start of the experiment, only decreasing significantly at the end. In the non-tailored condition, however, this behaviour was a lot more erratic, with significant changes between almost all lessons. However, there is no significant difference between the first and the last lessons (lessons 1 and 6), meaning that even if there were a lot of changes throughout the experiment, there were no significant changes overall for this condition.

With regard to RQ3, we show that (1) in general all types of learners' behaviours decrease overtime, (2) the analysis lesson per lesson allows better qualification of the differences observed in RQ2 and reveals more detailed information: lots of significant differences for behaviours corresponding to **F1** and **F3** between lessons and no differences for **F2** except for the final lesson, and (3) behaviours corresponding to **F1** and **F2** decreased less in the tailored condition than in the non-tailored condition, especially when looking at the final lessons in the experiment.

7 DISCUSSION

7.1 Different Types of Engaged Behaviours

The analysis presented in Section 6.1 reveals three different behavioural patterns computed from students' interactions with the learning environment, corresponding to different kinds of engaged behaviours. We showed that these behaviours differ depending on whether learners have tailored or non-tailored game elements. As previously stated (in Section 6.2), we are not interested in the individual scores themselves for each engaged behaviour, but we use them to analyse their evolution along the lessons and the differences between experimental conditions.

7.1.1 F1: Trial and Error Pattern. In this behavioural pattern, learners were behaving following a trial and error method, favouring discovering the learning content, answering questions quickly, trying out lots of quizzes, and reading more about how their game element works. It is important to mention that, in education, trial and error can be a sign of learners who do not know what they are doing or who are operating without a useful heuristic or theory of action. However, in game spaces, players are encouraged to try and fail and retry to progress. In fact, freedom to fail has been documented as an important facet of gamification [64]. With regard to the SDT [59], this behaviour could refer to the basic motivational need of autonomy as it shows a certain curiosity for the environment. Autonomy is an indicator of engagement often used in related work. In their study, da Rocha Seixas et al. [11] underline that autonomy allows students to conduct activities making their own choices and thus being intrinsically motivated. This echoes [16] when explaining the links between SDT and video games: *"the voluntariness of play provides a strong experience of autonomy, which is intrinsically motivating".*

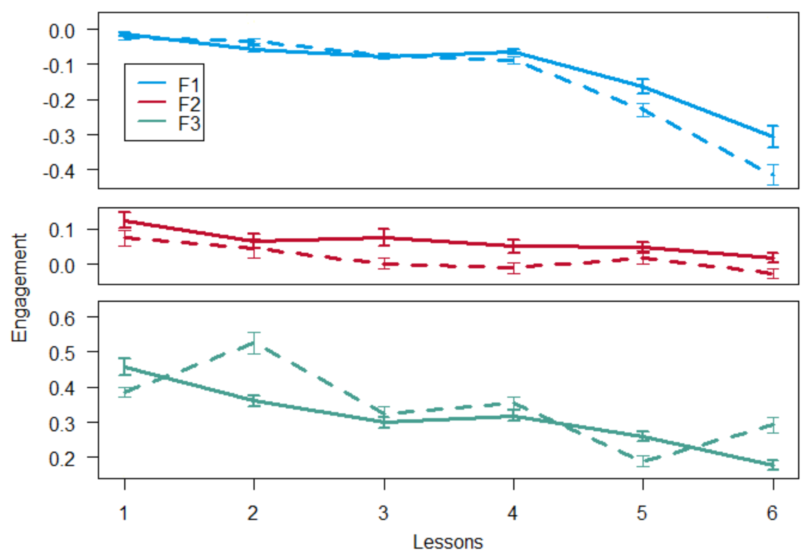

Fig. 5. Evolution graph for the engagement factors. Solid lines represent the learners in the tailored condition, and dashed lines those in the non-tailored condition

Table 3. Average engagement scores for each lesson, for each factor and condition. The "W" column shows the results of the Wilcoxon test comparing the tailored and non-tailored conditions and the "d" column shows the effect size. Values in grey are not significant p>.05.

	F1				F2				F3			
Lesson	Tailored	Non-T.	W	d	Tailored	Non-T.	W	d	Tailored	Non-T.	W	d
1	-0.0137	-0.0248	863	0.169	0.1272	0.0963	1002	0.066	0.4563	0.3961	983.5	0.079
2	-0.0572	-0.0264	1426	0.251	0.0669	0.0718	1040	0.040	0.3605	0.5855	1847.5	0.079
3	-0.0770	-0.0690	1168	0.059	0.0773	0.0040	878	0.179	0.2993	0.3473	1260	0.128
4	-0.0644	-0.0983	752	0.251	0.0519	-0.0074	874.5	0.187	0.3185	0.3494	1278.5	0.141
5	-0.1630	-0.2261	696	0.293	0.0484	0.0408	1020	0.056	0.2599	0.1866	668.5	0.314
6	-0.3064	-0.4232	620	0.350	0.0176	-0.0199	861	0.214	0.1777	0.3069	1691	0.449

Table 4. p-values of the wilcoxon tests between average engagement values of lessons for each situation, corrected with the Bonferroni correction and effect sizes in column "d". In grey p>.05, in black p<.05, highlighted in grey p<0.01, highlighted in black p<.001

	F1				F2				F3			
ΔL.	T.	d	Non-T.	d	T.	d	Non-T.	d	T.	d	Non-T.	d
1-2	3.67E-07	0.418	1	0.023	0.517626	0.178	1	0.192	0.787899	0.259	3.46E-05	0.594
2-3	0.007659	0.229	8.44E-05	0.476	1	0.001	1	0.174	0.159175	0.228	1.74E-05	0.619
3-4	0.224082	0.124	0.017807	0197	1	0.063	1	0.071	1	0.045	1	0.049
4-5	3.67E-06	0.328	0.0003	0.520	1	0.064	0.873318	0.203	0.019689	0.203	1.05E-08	0.631
5-6	1.30E-06	0.374	4.40E-06	0.547	1	0.123	0.643233	0.283	0.000608	0.381	0.020728	0.444
1-6	2.52E-11	0.850	3.49E-09	0.860	9.12E-05	0.334	0.008331	0.502	3.69E-11	0.749	0.092813	0.399

7.1.2 F2: Improvement by Repetition Pattern. Learners who behaved with this pattern were concerned with improving their performances and "perfecting" quizzes. They would complete the minimum required grade and restart said quizzes to try and obtain 100% over and over again until they achieved perfection. This type of pattern seems recurrent in learning situations, since it is similar to the "Perfection-oriented engagement" as identified in [39] and the repetition behavioural pattern identified in [9]. We believe that this behaviour is linked to the basic need for competence in the SDT theory [59] and that it is influenced by the motivational affordances related to the artefact, more specifically by how game elements were designed. Indeed, learners were free to restart any of the quizzes even if they had achieved the minimum required grade of 70%, and could gain extra rewards (points, badges, objects for their avatar, etc.) if they scored higher. This is in line with findings from other similar studies, where badges, ranking and scores may have increased the feeling of learners' competence [60, 71] or pushed them to increase their performance [38].

7.1.3 F3: Perfection Pattern. In this pattern, learners aimed to complete quizzes with the best possible performance on the first try. This behaviour is similar to the "Achievement-oriented engagement" as proposed in [39], as it contains similar indicators with similar rates and could be related to the need for competence in the SDT [59]. Several studies report correlations between users' activity (assimilated to their engagement) and their performance [15, 62], showing that students with gamified quizzes had significantly better scores on the first attempt.

7.2 Evolution of Learners' Behaviours

Overall, learners' engagement gradually decreased over time, with some nuances depending on each engaged behaviour type. The decrease in *Trial and error* pattern was more or less constant throughout the experiment. The *improvement by repetition* pattern was somewhat stable throughout the experiment (if we look on a lesson by lesson basis). However, when looking at the difference between the first and the last lesson, we observe a significant decrease. Finally, for the *perfection* pattern, learners in the tailored condition saw a more stable decrease, while those in the non-tailored condition were more erratic, with decreases and increases throughout the experiment. Several longitudinal studies also reported a decrease in engagement or performances in the long run [31, 56]. According to O'Brien and Toms [47], the process of engagement may consist of various stages (including points of engagement, disengagement or reengagement). In this study, the general decrease in engagement could be attributed to two main reasons: (1) the increasing complexity of the learning content, and (2) the lack of novelty or the weariness effect.

The learning content was designed so that each new lesson would introduce a new learning concept and make learners apply it in the quizzes. This was a decision made by the participating teachers and naturally meant that the last lessons are harder than the first ones. This could lead learners to have an increasingly harder time with the questions, potentially causing a loss of engagement. In their study on engagement, O'Brien and Toms identify the task difficulty as a potential factor of disengagement [47]. Landers et al. [38] showed that their gamification had a positive impact on learners until task difficulty became too great. This seems to suggest that gamification, even when tailored, cannot overcome the effects caused by the underlying learning content.

Regarding the second point, several studies discuss how a novelty effect could lead to positive effects of gamification that would not necessarily continue over time [30, 37, 63]. It is possible that, in this study, learners were engaged at first due to the introduction of a new tool in the classroom, and that their engagement progressively decreased over time. Several longitudinal studies corroborate that gamification suffers from the novelty effect [56, 62]. It is also possible that the students felt a certain weariness towards the platform content. Indeed, there was no variety in the format of the quizzes, and all the lessons were on the same part of the mathematics programme (introduction to basic algebra). This is in line with the study by O'Brien and Toms, who identified the lack of novelty as a potential factor of disengagement [47].

7.3 Effect of Tailoring Game Elements on Learners' Behaviours

An interesting finding in our results is the fact that tailoring game elements to learners' profile seems to have mitigated their loss of engagement. Differences between both conditions can be observed at a global level for two behavioural patterns: Improvement by repetition and Perfection. A detailed analysis, lesson by lesson, shows that these differences occur mainly for the latest lessons. It also highlights significant differences for Trial and error in the second half of the experiment. These results point to the need to observe engagement during its evolution and not only at the end of the experiment like most of the studies do. This also shows that tailoring gamification can make a difference in the long-term and argues for long-term studies when analysing the impact of tailored gamification on learners' engagement. This is in line with Altmeyer et al. [2], who showed more significant differences at the end of their study than at the beginning between adapted and non-adapted conditions. Oliveira et al. [49] also argue for more empirical and longitudinal studies.

Looking at each engaged behaviour in detail, the Trial and error pattern decreases less for learners with an tailored game element than for those who had non-tailored ones at the end of the experiment. This means that over time, learners who followed this behaviour with randomly assigned game elements were significantly less curious and experimental in their approach than those who had game elements tailored to their profile. They were probably more sensitive to the underlying mechanics of their game element, making them interact with various quizzes and their game element more than the others.

Regarding the improvement by repetition pattern, learners in the tailored condition were generally more *stubborn* in their approach than learners in the non-tailored condition (only significantly for the last lesson). Tailoring the game elements to their preferences for game mechanics and initial motivations for mathematics made them be more focused on each quiz and determined to achieve at 100%. This is in line with previous studies in education, such as the one conducted by Roosta et al. [58], who observed significant differences for engagement and quiz results between adapted and randomised conditions.

Finally, for the Perfection pattern, the results are not as conclusive. On the one hand, learners who followed this pattern with a tailored game element were progressively less engaged. We can suppose that tailoring to learners did not help improve perfection, but rather made it more stable (i.e. removed the erratic nature that we observe for learners in the non-tailored condition). On average, learners in the non-tailored condition showed higher engagement scores than learners in the tailored condition. This finding echoes the study conducted by Oliveira et al. [48], who found that concentration was improved in the counter-tailored condition for some player types. This is possibly due to the game elements being too distracting for some learners. We believe that the lack of variety in the learning content contributed to the decrease in engagement for this behavioural pattern, meaning that gamification (whether tailored or not) might have had less of an effect than for the other two patterns.

8 IMPLICATIONS FOR THE DESIGN OF ADAPTIVE GAMIFIED LEARNING ENVIRONMENTS

As stated previously, this paper aims at enriching our knowledge in the emerging field of adaptive gamification by investigating how tailoring game elements to learners affects their behaviours during the use of a gamified learning environment. Dynamic modelling has been identified as an important challenge for future research in adaptive gamification [36].

Even if our behavioural patterns are highly dependent on the engagement indicators chosen and thus on the learning environment studied, we believe that considering the dynamic nature of engagement thanks to interaction patterns is a step towards dynamic modelling issues that can be applied in other contexts. In particular, we showed that the analysis of learners' interactions, combined with an exploratory approach based on factor analysis, allows identification of objective models of engagement. Since our engagement indicators were inspired by indicators most commonly used to measure accuracy (time on task and number of exercises), they can be easily adapted according to the specificity of the learning environment and possible interactions. From these indicators, a factor analysis can be performed to obtain behavioural patterns.

Even if it cannot prevent a general decrease in learners' engagement after several learning sessions, we showed some positive effects of adaptive gamification. However, we also highlighted its complexity and pointed out that adaptive gamification should be designed with care in learning environments.

First, when designing the gamified environment, meaningful game elements should make sense to learners in the context of their learning activity and thus increase their feeling of competence and autonomy [39]. Designers should offer a wide variety of content and activities to avoid a weariness effect. Regarding the adaptation process, in this study, while we were only interested in the adaptation of the game elements, we could suppose that the adaptation of the content itself could prevent a decrease in learners' engagement, especially for learners who adopt a perfection oriented behaviour. It would be especially interesting to provide them with an adapted level of complexity so that they feel competent but also challenged (referring to the SDT) and stay involved in the task.

Second, we showed that learners' behaviours were more or less pronounced in the learning environment whether the gamification approach is tailored or not. It is important to note that we identified three behavioural patterns, but learners may exhibit one, multiple, or none of them. These engaged behaviours are rather complex and evolve over time. This has implications for the adaptation process of the game elements. While almost all approaches are based on a static model of learners to adapt the game elements to their profile before they use the environment, we believe that it would be important to also take into account how they behave while using the environment. This would lead to a richer learner model, based on both personal characteristics (player type, personality traits) and situational characteristics (e.g. motivation for the learning task, engagement for the task). As learners' engagement fluctuates over time, their profile could be updated according to the evolution of their behaviour (type and level) throughout the learning sessions, allowing a dynamic adaptation of both game elements and learning content. However, such an adaptive approach raises important well-known issues related to intelligibility and controllability of adaptation [1]. For instance, while the adaptation could be made automatically by the system based on the learners' profile, this may disrupt their activity if they do not understand this change. In a more flexible way, the system could suggest a game element to the learners and let them decide if they agree with this change. Further studies should be conducted to test and refine these approaches to dynamic adaptation of gamification.

9 LIMITATIONS

We identified a few limitations of our study. First, our experiment was conducted at a secondary school level, involving learners of the same age carrying out specific learning activities (quizzes), and was solely focused on mathematics. It is now well-known that the motivational impact of certain game elements varies according to the user activity or the domain of gamified systems [28]. Thus, other studies would be necessary to validate our model of engagement, i.e. the engaged behaviours identified in this context. However, the approach we propose for defining the different factors of engagement is context-independent. We believe that the method can be generalised to other contexts and systems. Furthermore, a previous study in the same context, but with different learners and game elements [39], did show similar patterns, leading us to believe that this model is robust. Second, in our study we decided to exclude all grades and questionnaires on learners' knowledge to prevent this from influencing their motivation. We believe that this makes our findings easier to use by researchers or pedagogical engineers in their practices. Further studies would be necessary to investigate the relationships between engaged behaviours induced by the use of the gamified learning environment and learning outcomes. Third, we specifically made the choice not to include an ungamified control situation. This is because we were interested in evaluating how effective the adaptation was on engaging learners and not the gamification itself, since the question of how effective gamification can be is something that has been widely investigated in the related literature (see Section 2.1). Fourth, our study was conducted over 6 learning sessions. Whilst we did start to see some differences in learners' engagement towards the end of the experiment, we would expect to see these differences become more marked over time. Finally, the engaged behaviours observed may depend on the adaptation process and on the matching between game elements proposed to learners and their preferences. We used previous field data from a very similar context to ensure the best match. However, we could suppose that learners' engagement would have been different depending on the adaptation algorithm. Future work should be conducted to investigate if combining each profile differently could lead to other recommendations and thus maybe to different behavioural outcomes.

10 CONCLUSION

In this paper, we presented the results of a large-scale study on the impact of a gamified learning environment on learners' engaged behaviours and how tailoring to learners' player profile and initial motivation for mathematics influences these behaviours.

Our main contributions are, with regard to RQ1, three different behavioural patterns that can be observed through learners' interactions: one related to the extent to which learners adopted a trial and error approach, one related to improvement by repetition, and the last one related to perfection. In answer to RQ2, we showed that, when looking at each engaged behaviour averaged over the six lessons, we observe global differences between the tailored and non-tailored condition for the improvement by repetition and perfection patterns. In addition, regarding RQ3 and when looking at a lesson by lesson basis, we can see more significant differences in learners' engagement. More importantly, these differences only emerge during the final lessons. It is important to note that, in general, learners' engagement decreased over time, but that tailoring the game elements to learners seemed to reduce this loss or make it more stable depending on the behavioural pattern.

These results contribute to a better understanding of how tailoring gamification to learners can affect their behaviours when using a gamified learning environment, enriching our knowledge on a question still under-explored in the field of gamification and, more particularly, in education, that could have important implications for the design of such environments. We also highlight the importance of running long-term studies on the effects of adaptive gamification, as the differences involved in our adaptation situations only really emerge during the final lessons.

Future research should investigate if specific game elements influenced these effects differently and how they can be replicated and extended to other situations. Another possible venue of future research could be to investigate how the learner profile influences these behaviours using clustering approaches as suggested in [5].

ACKNOWLEDGMENTS

This work is part of the LudiMoodle and LudiMoodle+ (ANR-22-FRAN-0005) projects financed by the e-FRAN Programme d'investissement d'avenir and the Agence nationale de la recherche for France 2030. We would also like to thank the teachers and learners who participated in this study.

REFERENCES

[1] Ashraf Abdul, Jo Vermeulen, Danding Wang, Brian Y. Lim, and Mohan Kankanhalli. 2018. *Trends and Trajectories for Explainable, Accountable and Intelligible Systems: An HCI Research Agenda.* Association for Computing Machinery, New York, NY, USA, 1–18.

[2] Maximilian Altmeyer, Marc Schubhan, Antonio Krüger, and Pascal Lessel. 2021. *A Long-Term Investigation on the Effects of (Personalized) Gamification on Course Participation in a Gym.* arXiv:2107.12597 [cs] http://arxiv.org/abs/2107.12597

[3] Gabriel Barata, Sandra Gama, Joaquim Jorge, and Daniel Gonçalves. 2017. Studying student differentiation in gamified education: A long-term study. *Computers in Human Behavior* 71 (2017), 550–585.

[4] Amy S Beavers, John W Lounsbury, Jennifer K Richards, Schuyler W Huck, Gary J Skolits, and Shelley L Esquivel. 2013. Practical considerations for using exploratory factor analysis in educational research. *Practical Assessment, Research, and Evaluation* 18, 1 (2013), 6.

[5] Souha Bennani, Ahmed Maalel, and Henda Ben Ghezala. 2021. Adaptive gamification in E-learning: A literature review and future challenges. *Computer Applications in Engineering Education* 30 (Dec. 2021). https://doi.org/10.1002/cae.22477

[6] Martin Böckle, Jasminko Novak, and Markus Bick. 2017. Towards adaptive gamification: a synthesis of current developments. (2017).

[7] Michelle Briffa, Nevena Jaftha, Giuseppe Loreto, Flavia Cristina Morone Pinto, Tatjana Chircop, and Corradino Hill. 2020. Improved students' performance within gamified learning environment: A meta-analysis study. *International Journal of Education and Research* 8, 1 (2020), 223–244.

[8] Charles Butler. 2014. A framework for evaluating the effectiveness of gamification techniques by personality type. In *HCI in Business: First International Conference, HCIB 2014, Held as Part of HCI International 2014, Heraklion, Crete, Greece, June 22-27, 2014. Proceedings 1.* Springer, 381–389.

[9] David Codish and Gilad Ravid. 2015. Detecting playfulness in educational gamification through behavior patterns. *IBM Journal of Research and Development* 59, 6 (2015), 6–1.

[10] Anna B Costello and Jason Osborne. 2005. Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical assessment, research, and evaluation* 10, 1 (2005), 7.

[11] Luma da Rocha Seixas, Alex Sandro Gomes, and Ivanildo José de Melo Filho. 2016. Effectiveness of Gamification in the Engagement of Students. *Computers in Human Behavior* 58 (2016), 48–63.

[12] Edward Deci and Richard Ryan. 2000. The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological inquiry* 11, 4 (2000), 227–268.

[13] Edward L. Deci and Richard M. Ryan. 1985. The General Causality Orientations Scale: Self-Determination in Personality. *Journal of research in personality* 19, 2 (1985), 109–134.

[14] Edward L. Deci and Richard M. Ryan. 2000. The "What" and "Why" of Goal Pursuits: Human Needs and the Self-Determination of Behavior. 11, 4 (2000), 227–268. https://doi.org/10.1207/S15327965PLI1104_01

[15] Paul Denny, Fiona McDonald, Ruth Empson, Philip Kelly, and Andrew Petersen. 2018. Empirical Support for a Causal Relationship Between Gamification and Learning Outcomes. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, CHI 2018, Montreal, QC, Canada, April 21-26, 2018*, Regan L. Mandryk, Mark Hancock, Mark Perry, and Anna L. Cox (Eds.). ACM, 311. https://doi.org/10.1145/3173574.3173885

[16] Sebastian Deterding. 2011. Situated Motivational Affordances of Game Elements: A Conceptual Model. In *Gamification: Using Game Design Elements in Non-Gaming Contexts, a Workshop at CHI*, Vol. 10.

[17] Sebastian Deterding. 2015. The Lens of Intrinsic Skill Atoms: A Method for Gameful Design. 30, 3-4 (2015), 294–335.

[18] Sebastian Deterding, Dan Dixon, Rilla Khaled, and Lennart Nacke. 2011. From Game Design Elements to Gamefulness: Defining Gamification. In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments.* ACM, 9–15. http://dl.acm.org/citation.cfm?id=2181040

[19] Darina Dicheva, Christo Dichev, Gennady Agre, and Galia Angelova. 2015. Gamification in Education: A Systematic Mapping Study. *Educational Technology & Society* 18 (July 2015), 75–88.

[20] Lu Ding, Erkan Er, and Michael Orey. 2018. An Exploratory Study of Student Engagement in Gamified Online Discussions. *Computers & Education* 120 (2018), 213–226.

[21] Lu Ding, ChanMin Kim, and Michael Orey. 2017. Studies of Student Engagement in Gamified Online Discussions. *Computers & Education* 115 (2017), 126–142.

[22] Leandre R Fabrigar, Duane T Wegener, Robert C MacCallum, and Erin J Strahan. 1999. Evaluating the use of exploratory factor analysis in psychological research. *Psychological methods* 4, 3 (1999), 272.

[23] Andy Field, Jeremy Miles, and Zoë Field. 2012. *Discovering statistics using R.* Sage publications.
[24] Ed Fincham, Alexander Whitelock-Wainwright, Vitomir Kovanović, Srećko Joksimović, Jan-Paul van Staalduinen, and Dragan Gašević. 2019. Counting Clicks Is Not Enough: Validating a Theorized Model of Engagement in Learning Analytics. In *Proceedings of the 9th International Conference on Learning Analytics & Knowledge.* 501–510.
[25] Jennifer A Fredricks, Phyllis C Blumenfeld, and Alison H Paris. 2004. School Engagement: Potential of the Concept, State of the Evidence. *Review of Educational Research* 74, 1 (2004), 59–109.
[26] Stuart Hallifax, Elise Lavoué, and Audrey Serna. 2020. To Tailor or Not to Tailor Gamification? An Analysis of the Impact of Tailored Game Elements on Learners' Behaviours and Motivation. In *Artificial Intelligence in Education* (Cham), Ig Ibert Bittencourt, Mutlu Cukurova, Kasia Muldner, Rose Luckin, and Eva Millán (Eds.). Springer International Publishing, 216–227.
[27] Stuart Hallifax, Audrey Serna, Jean-Charles Marty, and Elise Lavoué. 2018. A Design Space For Meaningful Structural Gamification. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems.* ACM, LBW073.
[28] Stuart Hallifax, Audrey Serna, Jean-Charles Marty, Guillaume Lavoué, and Elise Lavoué. 2019. Factors to Consider for Tailored Gamification. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play* (New York, NY, USA) *(CHI PLAY '19).* ACM, 559–572. https://doi.org/10.1145/3311350.3347167
[29] Stuart Hallifax, Audrey Serna, Jean-Charles Marty, and Élise Lavoué. 2019. Adaptive Gamification in Education: A Literature Review of Current Trends and Developments. In *Transforming Learning with Meaningful Technologies* (Cham), Maren Scheffel, Julien Broisin, Viktoria Pammer-Schindler, Andri Ioannou, and Jan Schneider (Eds.). Springer International Publishing, 294–307.
[30] Juho Hamari, Jonna Koivisto, and Harri Sarsa. 2014. Does Gamification Work?–A Literature Review of Empirical Studies on Gamification. In *System Sciences (HICSS), 2014 47th Hawaii International Conference On.* IEEE, 3025–3034. http://ieeexplore.ieee.org/abstract/document/6758978/
[31] Michael D. Hanus and Jesse Fox. 2015. Assessing the Effects of Gamification in the Classroom: A Longitudinal Study on Intrinsic Motivation, Social Comparison, Satisfaction, Effort, and Academic Performance. *Computers & Education* 80 (2015), 152–161.
[32] Muhammad Awais Hassan, Ume Habiba, Fiaz Majeed, and Muhammad Shoaib. 2019. Adaptive gamification in e-learning based on students' learning styles. *Interactive Learning Environments* (2019), 1–21.
[33] James C Hayton, David G Allen, and Vida Scarpello. 2004. Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis. *Organizational research methods* 7, 2 (2004), 191–205.
[34] Benjamin Heilbrunn, Philipp Herzig, and Alexander Schill. 2017. Gamification analytics—methods and tools for monitoring and adapting gamification designs. *Gamification: Using game elements in serious contexts* (2017), 31–47.
[35] Srećko Joksimović, Oleksandra Poquet, Vitomir Kovanović, Nia Dowell, Caitlin Mills, Dragan Gašević, Shane Dawson, Arthur C. Graesser, and Christopher Brooks. 2018. How Do We Model Learning at Scale? A Systematic Review of Research on MOOCs. 88, 1 (2018), 43–86.
[36] Ana Carolina Tomé Klock, Isabela Gasparini, Marcelo Soares Pimenta, and Juho Hamari. 2020. Tailored Gamification: A Review of Literature. 144 (2020), 102495. https://doi.org/10.1016/j.ijhcs.2020.102495
[37] Jonna Koivisto and Juho Hamari. 2019. The Rise of Motivational Information Systems: A Review of Gamification Research. 45 (2019), 191–210. https://doi.org/10.1016/j.ijinfomgt.2018.10.013
[38] Richard N. Landers, Kristina N. Bauer, and Rachel C. Callan. 2017. Gamification of Task Performance with Leaderboards: A Goal Setting Experiment. *Computers in Human Behavior* 71 (2017), 508–515.
[39] Elise Lavoué, Qinjie Ju, Stuart HALLIFAX, and Audrey Serna. 2021. Analyzing the Relationships between Learners' Motivation and Observable Engaged Behaviors in a Gamified Learning Environment. (2021). https://hal.archives-ouvertes.fr/hal-03221038
[40] Élise Lavoué, Baptiste Monterrat, Michel Desmarais, and Sébastien George. 2019. Adaptive Gamification for Learning Environments. 12, 1 (2019), 16–28. https://doi.org/10.1109/TLT.2018.2823710
[41] Elizabeth A. Linnenbrink and Paul R. Pintrich. 2003. The Role of Self-Efficacy Beliefs in Student Engagement and Learning in the Classroom. *Reading & Writing Quarterly: Overcoming Learning Difficulties* 19, 2 (2003), 119–137.
[42] C. Lopez and C. Tucker. 2018. Towards Personalized Adaptive Gamification: A Machine Learning Model for Predicting Performance. (2018), 1–1. https://doi.org/10.1109/TG.2018.2883661
[43] A. C. Marczewski. 2015. Even Ninja Monkeys like to Play. (2015).
[44] Baptiste Monterrat, Michel Desmarais, Elise Lavoué, and Sébastien George. 2015. A Player Model for Adaptive Gamification in Learning Environments. In *International Conference on Artificial Intelligence in Education.* Springer, 297–306. http://link.springer.com/chapter/10.1007/978-3-319-19773-9_30
[45] Baptiste Monterrat, Élise Lavoué, and Sébastien George. 2017. Adaptation of Gaming Features for Motivating Learners. 48, 5 (2017), 625–656. https://doi.org/10.1177/1046878117712632
[46] A. Mora, G. F. Tondello, L. E. Nacke, and J. Arnedo-Moreno. 2018. Effect of Personalized Gameful Design on Student Engagement. In *2018 IEEE Global Engineering Education Conference (EDUCON).* 1925–1933. https://doi.org/10.1109/

EDUCON.2018.8363471

[47] Heather L. O'Brien and Elaine G. Toms. 2008. What Is User Engagement? A Conceptual Framework for Defining User Engagement with Technology. 59, 6 (2008), 938–955. https://doi.org/10.1002/asi.20801

[48] Wilk Oliveira and Ig Ibert Bittencourt. 2019. *Tailored Gamification to Educational Technologies.* Springer Singapore. https://doi.org/10.1007/978-981-32-9812-5

[49] Wilk Oliveira, Juho Hamari, Lei Shi, Armando M Toda, Luiz Rodrigues, Paula T Palomino, and Seiji Isotani. 2022. Tailored gamification in education: A literature review and future agenda. *Education and Information Technologies* (2022), 1–34.

[50] Wilk Oliveira, Armando Toda, Paula Toledo, Lei Shi, Julita Vassileva, Ig Ibert Bittencourt, and Seiji Isotani. 2020. Does Tailoring Gamified Educational Systems Matter? The Impact on Students' Flow Experience. In *Proceedings of the 53rd Hawaii International Conference on System Sciences.*

[51] Rita Orji, Lennart E. Nacke, and Chrysanne DiMarco. 2017. Towards Personality-Driven Persuasive Health Games and Gamified Systems. In *Proceedings of SIGCHI Conference on Human Factors Computing System.* https://www.researchgate.net/profile/Rita_Orji/publication/314435151_Towards_Personality-driven_Persuasive_Health_Games_and_Gamified_Systems/links/58c2275aaca272e36dcc8d91/Towards-Personality-driven-Persuasive-Health-Games-and-Gamified-Systems.pdf

[52] Rita Orji, Gustavo F. Tondello, and Lennart E. Nacke. 2018. Personalizing Persuasive Strategies in Gameful Systems to Gamification User Types. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, Vol. 61. 62.

[53] Ranilson Paiva, Ig Ibert Bittencourt, Thyago Tenório, Patricia Jaques, and Seiji Isotani. 2016. What do students do on-line? Modeling students' interactions to improve their learning experience. *Computers in Human Behavior* 64 (2016), 769–781.

[54] Paul R. Pintrich. 2004. A Conceptual Framework for Assessing Motivation and Self-Regulated Learning in College Students. *Educational Psychology Review* 16, 4 (Dec. 2004), 385–407.

[55] Amon Rapp, Frank Hopfgartner, Juho Hamari, Conor Linehan, and Federica Cena. 2019. Strengthening Gamification Studies: Current Trends and Future Opportunities of Gamification Research. 127 (2019), 1–6. https://doi.org/10.1016/j.ijhcs.2018.11.007

[56] Luiz Rodrigues, Filipe D. Pereira, Armando M. Toda, Paula T. Palomino, Marcela Pessoa, Leandro Silva Galvão Carvalho, David Fernandes, Elaine H. T. Oliveira, Alexandra I. Cristea, and Seiji Isotani. 2022. Gamification suffers from the novelty effect but benefits from the familiarization effect: Findings from a longitudinal study. *International Journal of Educational Technology in Higher Education* 19, 1 (Feb. 2022), 13. https://doi.org/10.1186/s41239-021-00314-6

[57] Luiz Rodrigues, Armando M. Toda, Paula T. Palomino, Wilk Oliveira, and Seiji Isotani. 2021. Personalized Gamification: A Literature Review of Outcomes, Experiments, and Approaches. (2021), 699–706. https://doi.org/10.1145/3434780.3436665

[58] F. Roosta, F. Taghiyareh, and M. Mosharraf. 2016. Personalization of Gamification-Elements in an e-Learning Environment Based on Learners' Motivation. In *2016 8th International Symposium on Telecommunications (IST).* 637–642. https://doi.org/10.1109/ISTEL.2016.7881899

[59] R. M. Ryan and E. L. Deci. 2000. Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. 55, 1 (2000), 68–78. pmid:11392867

[60] Michael Sailer, Jan Ulrich Hense, Sarah Katharina Mayr, and Heinz Mandl. 2017. How Gamification Motivates: An Experimental Study of the Effects of Specific Game Design Elements on Psychological Need Satisfaction. *Computers in Human Behavior* 69 (2017), 371–380.

[61] Michael Sailer and Lisa Homner. 2020. The Gamification of Learning: a Meta-analysis. *Educational Psychology Review* 32 (March 2020), 77–112. https://doi.org/10.1007/s10648-019-09498-w

[62] Diana R Sanchez, Markus Langer, and Rupinder Kaur. 2020. Gamification in the classroom: Examining the impact of gamified quizzes on student learning. *Computers & Education* 144 (2020), 103666.

[63] Katie Seaborn and Deborah I Fels. 2015. Gamification in theory and action: A survey. *International Journal of human-computer studies* 74 (2015), 14–31.

[64] Daniel Strmecki, Andrija Bernik, and Danijel Radosevic. 2015. Gamification in E-Learning: Introducing Gamified Design Elements into E-Learning Systems. *J. Comput. Sci.* 11, 12 (2015), 1108–1117.

[65] Sujit Subhash and Elizabeth A. Cudney. 2018. Gamified learning in higher education: A systematic review of the literature. *Computers in Human Behavior* 87 (Oct. 2018), 192–206. https://doi.org/10.1016/j.chb.2018.05.028

[66] Necati Taşkın and Ebru Kılıç Çakmak. 2022. Effects of gamification on behavioral and cognitive engagement of students in the online learning environment. *International Journal of Human–Computer Interaction* (2022), 1–12.

[67] Gustavo F. Tondello, Alberto Mora, Andrzej Marczewski, and Lennart E. Nacke. [n.d.]. Empirical Validation of the Gamification User Types Hexad Scale in English and Spanish. 127 ([n. d.]), 95–111.

[68] Gustavo F. Tondello and Lennart E. Nacke. 2020. Validation of User Preferences and Effects of Personalized Gamification on Task Performance. 2 (2020). https://doi.org/10.3389/fcomp.2020.00029

[69] Robert J. Vallerand, Luc G. Pelletier, Marc R. Blais, Nathalie M. Briere, Caroline Senecal, and Evelyne F. Vallieres. 1992. The Academic Motivation Scale: A Measure of Intrinsic, Extrinsic, and Amotivation in Education. 52, 4 (1992), 1003–1017.

[70] Julita Vassileva. 2012. Motivating Participation in Social Computing Applications: A User Modeling Perspective. 22, 1 (2012), 177–201. http://www.springerlink.com/index/e40qq8928p3318h6.pdf

[71] Zamzami Zainuddin. 2018. Students' learning performance and perceived motivation in gamified flipped-class instruction. *Computers & Education* 126 (Nov. 2018), 75–88.

[72] Zamzami Zainuddin, Samuel Kai Wah Chu, Muhammad Shujahat, and Corinne Jacqueline Perera. 2020. The impact of gamification on learning and instruction: A systematic review of empirical evidence. *Educational Research Review* 30 (June 2020), 100326. https://doi.org/10.1016/j.edurev.2020.100326

Received 2023-02-21; accepted 2023-07-07