HAL
open science

Factors to Consider for Tailored Gamification

Stuart Hallifax, Audrey Serna, Jean-Charles Marty, Guillaume Lavoué, Elise Lavoué

To cite this version:

Stuart Hallifax, Audrey Serna, Jean-Charles Marty, Guillaume Lavoué, Elise Lavoué. Factors to Consider for Tailored Gamification. CHI Play, Oct 2019, Barcelona, Spain. pp.559-572, 10.1145/3311350.3347167 . hal-02185647

HAL Id: hal-02185647

https://hal.science/hal-02185647v1

Submitted on 16 Jul 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Factors to Consider for Tailored Gamification

Stuart Hallifax
Univ. Jean Moulin Lyon 3,
iaelyon school of Management
CNRS, LIRIS, France
stuart.hallifax@liris.cnrs.fr

Audrey Serna
INSA de Lyon
CNRS, LIRIS, France
audrey.serna@liris.cnrs.fr

Jean-Charles Marty
Univ. de Savoie Mont Blanc
CNRS, LIRIS, France
jean-charles.marty@liris.cnrs.fr

Guillaume Lavoué
INSA de Lyon
CNRS, LIRIS, France
guillaume.lavoue@liris.cnrs.fr

Elise Lavoué
Univ. Jean Moulin Lyon 3,
iaelyon school of Management
CNRS, LIRIS, France
elise.lavoue@liris.cnrs.fr

ABSTRACT
Gamification is widely used to foster user motivation. Recent studies show that users can be more or less receptive to different game elements, based on their personality or player profile. Consequently, recent work on tailored gamification tries to identify links between user types and motivating game elements. However findings are very heterogeneous due to different contexts, different typologies to characterize users, and different implementations of game elements. Our work seeks to obtain more generalizable findings in order to identify the main factors that will support design choices when tailoring gamification to users' profiles and provide designers with concrete recommendations for designing tailored gamification systems. For this purpose, we ran a crowdsourced study with 300 participants to identify the motivational impact of game elements. Our study differs from previous work in three ways: first, it is independent from a specific user activity and domain; second, it considers three user typologies; and third, it clearly distinguishes motivational strategies and their implementation using multiple different game elements. Our results reveal that (1) different implementations of a same motivational strategy have different impacts on motivation, (2) dominant user type is not sufficient to differentiate users according to their preferences for game elements, (3) Hexad is the most appropriate user typology for tailored gamification and (4) the motivational impact of certain game elements varies with the user activity or the domain of gamified systems.

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

ACM ISBN .
DOI:

Author Keywords
Gamification; User motivation; User Types; Tailoring; Crowdsourcing

INTRODUCTION
Since the early 2000s, gamification has been increasingly used to enhance user motivation [21, 49, 54, 67, 77]. Gamification, defined as *"the use of design elements characteristic of games in non game contexts"* [14], is used nowadays in many domains, from education [12, 44, 46], to health [39, 47], and the workplace [17, 22, 59]. These "design elements" are named differently, often using the terms "pervasive strategies", "motivational strategies", "game mechanics", or "game elements".

Several studies in this field show that users can be more or less receptive to different game elements [6, 41, 53, 55, 61]. These studies suggest that user personality and preferences have a great influence on the effect that game elements have on user motivation. Appropriate game elements can lead to higher levels of user motivation, whereas inappropriate game elements can demotivate users. Recent work on tailored gamification [21, 39, 56, 60] has provided valuable results that identify links between user types and relevant or motivating game elements. However, these results are very heterogeneous due to three reasons. First, the studies are generally carried out in different and particular domains (usually in health or education). Second, they rely on specific and different user typologies or personality models, mainly BrainHex [57], Hexad [74], or Big Five [25]. Third, they do not consider the same game elements or study different levels of abstraction of motivational strategies.

Thus, the motivational impact of game elements considered in these studies is difficult to isolate and the results are hard to reuse when designing tailored gamification. The goal of our work is to obtain more generalizable findings on factors to consider to support gameful design choices for tailored gamification. For this purpose, we conducted a new study on the motivational impact of game elements according to user types. Our study differs from existing works in three ways: first, our study is context-independent meaning that the

scenarios that illustrate the game elements are not related to a specific user activity or domain; second, we characterize users using two different user typologies (BrainHex, Hexad) and a personality trait model (Big Five); third, we make a clear distinction between the high-level motivational strategies and their implementation in the form of game elements.

Our study was conducted using a crowdsourcing platform and reached 616 participants. They were then filtered down to a set of 300 high-quality and consistent participants. The motivational impact of game elements was evaluated using a pairwise comparison protocol involving 66 comparisons between pairs of game elements illustrated by storyboards. Our results show that: (1) game elements from a same motivational strategy have significantly different impacts on user motivation, (2) the dominant user type is not sufficient to differentiate users according to their preferences for game elements (3) out of the user models tested, Hexad is most appropriate for tailored gamification, and (4) the motivational impact of certain game elements varies with the user activity and the domain of gamified systems. Finally, we provide designers with concrete recommendations for tailored gamification.

RELATED WORK

Gamification design and user motivation

Meaningful design advocates that game elements should make sense to users, creating explicit connections to the activity [13, 29, 58]. To do so, global design processes generally recommend to take the following steps: define the main objective, understand users' motivation for the activity, identify the game mechanics and analyze the effect of gamification [43, 50, 76].

To broaden these guidelines, several studies have been conducted to measure the impact of game elements on user motivation in different domains, mainly health [39, 62, 63, 64, 74] and education [12, 46, 55, 56]. They are usually based on a small set of game elements in specific contexts and do not reach a consensus on the impact on user motivation [69]. For instance, in the context of online communities users who used badges were significantly more engaged [30]. Another study showed that the integration of a leaderboard and badges induced a lower level of motivation and lower final exam scores [34]. However, leaderboards have also been shown to be interesting for supporting goal setting [45]. These contradictory results suggest that the observed impact of game elements may be influenced by the user activity (e.g. exchanging with peers, taking exams or goal setting) and/or the domain (e.g. online community, education). Hence these specific results are hard to reuse when designing gamification. Only one study [73] was conducted independently from a context, but considering only component elements and not specific game elements.

Theoretical foundations of gamification research mainly rely on the Self-Determination Theory (SDT [68]) that argues that human beings are intrinsically motivated to engage in activities that satisfy three basic psychological needs: competence (sense of efficacy), autonomy (volition and personal agency), and relatedness (social interaction). SDT also argues that Humans strive to fulfill these three needs in order to enhance well-being. Meaningful gamification should spontaneously provoke the satisfaction of these three user needs [66], especially the sense of competence [13]. This theoretical framework suggests that user motivation when using gamified environments may be enhanced by game elements independently from any context. For instance, we could suppose that the need for relatedness can be fulfilled with any kind of social interaction. However, existing studies do not allow us to identify general recommendations for gamification design as they were all conducted in specific contexts.

No ground studies have been conducted on the influence of the user activity and domain on the motivational impact of game elements. We address this question by considering context-independent game elements, and comparing our results with previous studies in specific contexts. Our study aims to provide general recommendations for game elements that can be used in any context, and identify those whose motivational impact varies with the context.

User typologies for tailored gamification

Several studies focus on player types to differentiate the motivational impact of game elements. They rely on different typologies, such as BrainHex [57] and Hexad [51].

The **BrainHex player typology** [57] was originally developed specifically as a "gamer" typology. However, research has applied it indiscriminately in both game and gamification studies [7, 46, 55, 60, 61, 64]. It proposes seven player types : *Seeker*, *Survivor*, *Daredevil*, *Mastermind*, *Conqueror*, *Socialiser*, and *Achiever*. BrainHex is based on a top-down approach, taking inspiration from neurobiological research [3]. For instance, Lavoué et al. [46] use BrainHex to adapt five game elements to users' dominant player type (i.e. the highest scoring type).

Marczewski more recently developed the **Gamification User Types Hexad framework** [51] specifically for gamification. This typology is based on SDT [68]. The Hexad typology designates six different user types: *Philanthropists*, *Socialisers*, *Free Spirits*, *Achievers*, *Players*, and *Disruptors*. Several recent experimental studies have been conducted in order to identify the motivational impact of game elements regarding the Hexad profile [56, 63].

Finally, a few recent studies [39, 62] on tailored gamification consider "personality traits" to identify user preferences, mainly the **Big Five Factors** [25]. The five user types are: *Openness to experience*, *Conscientiousness*, *Extraversion*, *Agreeableness*, and *Emotional stability*.

In this context, we first question the influence of the use of one particular user model for tailored gamification. Studies on tailored gamification were conducted using a particular user model, making it difficult to exploit the results obtained. For instance, in the field of persuasive technologies for health, three different studies contribute to identify strategies the most suited to the user profile using BrainHex [64], Big Five [62] and Hexad [63]. But the results obtained are not compared and designers are not advised on a particular typology or model to use. Furthermore, even if some user types have the same name, they do not have the same definitions, or represent the same user motivations (e.g. Achievers in BrainHex like collecting

everything, whereas achievers in Hexad are motivated by mastery). In our study, we propose to test all three commonly used user models, and to compare the influences found for each user model dimension in relation to their definitions.

We also question the relevance of considering only the dominant type for player typologies when tailoring gamification. Most of the studies only consider the dominant user type, i.e. the type that has the highest score [24, 8, 46, 21, 71]. Studies considering personality traits take all dimensions into account [62] and few very recent studies do the same for player typologies [63] or establish clusters with hybrid profiles for more fine-grained analysis [56]. These considerations could have an impact on tailoring gamification.

Motivational strategies and game elements

Werbach et al. [76] propose the DMC (Dynamics, Mechanics and Components) framework to classify game elements through three levels, from the most abstract to the most concrete elements. Dynamics are “the big-picture aspects of the gamified system” which are not directly used (constraints, emotions, narrative, etc.). Mechanics are “the basic processes that drive the action forward and generate player engagement” (challenges, chance, competition, etc.). Finally, components are “the specific instantiations of mechanics and dynamics” (achievements, avatars, badges, etc.).

In the persuasive technology field, studies are mainly focused on the game mechanics level, often named "persuasive strategies" or "motivational strategies" [64]. In field studies on the impact of motivational strategies, authors use commonly accepted categories of strategies such as competition, simulation, rewards, or comparison [33, 47, 62]. These studies usually rely on the evaluation of only one implementation of a given motivational strategy, and we can question the influence of the implementation choices on the effect of tailored gamification. Furthermore, the categories of persuasive strategies on which they rely do not clearly distinguish the motivational strategies from their implementation. For example, points and rewards are considered at the same level in [32], whereas rewards can be considered as a higher level concept that could be implemented using badges or points.

In the gamification field, studies generally focus on the "components" level of the DMC framework to evaluate the motivational impact of one or several commonly used game elements. For example, [2] examines how points affect test performance in mathematics. [12] compares the effectiveness of badges and points systems in an educational setting. [60] uses badges and leaderboards in an app designed to remedy unhealthy eating habits. [52] compares the effects of points, leaderboards, and levels on user motivation, competence and performance in an image annotation task. However, these studies generally do not consider the higher level of the motivational strategies they implement (mechanics and dynamics of the DMC framework).

More theoretical research provides lists of game elements as a first step to a more structured and systematic approach in the domain. For instance, [21] presents an extensive review of related literature to identify the commonly used game elements: achievements, combos, bars, leaderboards, points, status, badges, bonuses, levels, reward schedule, customization, quests, inventory (see [36, 40, 76] for other lists). Some theoretical studies identify direct links between game elements and the game mechanism they support [75], but they have not been tested on real ground yet.

Strategy	Game element	Equivalent to
Rewards	Badges	Badges [21, 39] Rewards [39] Badges or Achievements [51, 74]
	Points	Points [21, 39], Reward [62, 63, 64], Points [74, 51]
	Useful	Reward [62, 63, 64]
Goals	External	Goal setting & Goal suggestion [62, 63]
	Self	Clear goals [39]
Time	Schedule	Reward Schedule [21]
	Timer	Timer [46]
Social Interaction	Trading	Collection & trading [51, 74]
	Teams	Cooperation [62, 63, 64] Teamwork [38] Guilds or Team [51, 74]
	Discussion	Social network [51, 74]
Progress	Compared	Leaderboards [21, 39, 46] Competition and Comparison [64] Comparison [62, 63] Social Comparison [51, 74]
	Task	Bars [21] Progress [39] Self-monitoring and suggestion [64] Self-monitoring and feedback [62, 63] Advancement [38] Levels or progression [51, 74]

Table 1: The 12 game elements we used grouped by motivational strategy, and their equivalents in previous studies.

In conclusion, previous field studies on the impact of game elements on user motivation either consider low-level game elements (Components) without considering the motivational strategies (Mechanics and Dynamics) they support, or high-level strategies each represented by one single implementation. In our study, we structure game elements according to two different levels: an abstract level connected to motivational strategies and a more concrete level allowing to compare different implementations of these strategies. allowing us to study how user motivation varies with the implementation. We present the game elements used in our study in Section 3.

STUDY DESIGN

Our study seeks to identify the factors that will support design choices when tailoring gamification to user profiles. These choices underline a general question which is how we can identify users’ preferences for game elements. From the related work analysis, we address several research questions:

- *Considering the game elements,* "Do game elements implementing a same motivational strategy have different impacts on user motivation ?"
- *Considering the user typologies*, "Is the dominant user type sufficient to discriminate users’ preferences? And which typology should be chosen for tailored gamification?"
- *Considering the user activity and domain,* "To which extent does the context influence the motivational impact of game elements?"

To explore these questions, we consider 5 motivational strategies, implemented by 12 game elements. This list was created from the references listed in table 1, then filtered to only keep elements that are not directly linked to the content (such as storytelling) as they did not make sense for a domain independent task. We then grouped these game elements into an overarching motivational strategy they implement. The game elements are illustrated using storyboards that show implementations of motivational strategies independently from any specific user activity and domain. The perceived motivational impact of each element is evaluated using a paired comparison protocol which has been shown to be more reliable than direct rating [65, 70]. We compare variations based on users' profiles according to the three most commonly used user models: BrainHex, Big Five Factor and Hexad.

Materials

We designed a storyboard for each game element independent of any activity and domain (see an example in figure 1). This methodology was inspired by [62, 63, 64]. Each of the storyboards depicts three panels where a user completes a generic "task" with the game element changing accordingly. The full storyboards are presented in the supplementary materials.

Rewards game elements

In the *Points* storyboard, the user receives points each time s/he completes a task, his/her total score is shown in the game element area. For *Badges* the user gains a badge for completing the task, s/he can see a list of the badges s/he has obtained, as well as the badges s/he can still obtain in the game element area. For *Useful rewards*, the user completes a task and gains a "Give example" power. The area on the right shows the users inventory, with items such as "Skip task" and "Help" that suggest their usage.

Goals game elements

Both of the Goals storyboards have the same general structure: the game element area on the right shows a list of goals that the user has to complete, each one has a checkbox that shows whether a given goal has been completed or not. For *External Goals*, the storyboard shows the user completing a task, after which a popup informs him/her that the system has given them a new goal. A new goal and checkbox appear in their goal list. For *Self Goals* there is a button in the game element area that opens a window where the user can add a new goal himself/herself instead.

Time game elements

The *Schedule* storyboard shows a five day calendar. The user has to complete at least one task per day. Each day that a user completes a task, a check mark is placed in the corresponding day. After completing five days, the user receives a bonus. For the *Timer* storyboard, the user is timed the see how quickly s/he can complete a task. In the game element area, s/he is shown a table of his/her previous times. After completing a task the table updates, and the user is informed if s/he performed a new record (faster time than before).

Social Interaction game elements

In the *Trading* storyboard, a user selects a task to complete in the task area. A popup informs him/her that s/he needs a "key" to start that particular task, which s/he acquires by using the "chat" on the right to trade another user for it. For the *Teams* storyboard, the user acquires points each time s/he completes a task, the user can see his/her team score. A set of notifications on the right show when another user in his/her team completes a task and increases the team score. The *Discussion* storyboard shows a user stuck on a task. S/he uses the chat on the right to ask other users for help. Another user provides a solution to their problem and s/he can complete the task.

Progress game elements

Both of the Progress game element storyboards have a similar design. They both use a progress bar in the game element area that fills up when a user completes a task. In the *Compared Progress* storyboard, the bar shows the class average on the side and opens a popup when the user advances into the top 50% of their class. In the *Task Progress* storyboard the class average is replaced by a simple marking to show how much of the task the user has completed (50%, 60% etc.).

Storyboard validation

To ensure that these context-independent storyboards would be understood by all participants, we ran a pre-study comparing them to similar storyboards for a math learning activity. We asked participants to describe in their own words each storyboard. The descriptions were then reviewed by two different evaluators and graded out of 3. After an initial set of evaluations using 8 participants, we iterated on our storyboard design, and validated the final design using a further 2 participants. We found that the descriptions given for the context-free storyboards matched those given for the task-specific ones. Furthermore, to ensure that the comprehension of the storyboards was not influenced by age, or familiarity with video games (or game mechanics in general), we calculated the correlation between these factors. We found a very low correlation (0.10) between participant age and understanding, and a low correlation (0.28) between video game familiarity and understanding. We therefore judged that our storyboards could be understood as well as context-specific storyboards by all participants.

Procedure

As stated above, we used a paired comparison technique to evaluate the perceived motivational impact of each game element. Participants were shown pairs of storyboards, and were asked to choose which one they estimate "*would motivate them more to use the system*" (forced-choice methodology). Such paired comparison protocol offers 3 advantages over Likert-type rating [65]: (1) the experimental task is less cognitively demanding [10] (choosing a preference between two items is easier than providing an ordinal rating); (2) it avoids normalization issues which occur, for instance, when some users avoid extreme response categories; (3) it has been shown to provide higher sensitivity and lower measurement error [70].

We opted for a *full* paired-comparison design, meaning that each participant evaluates *all possible* pairs of storyboards, i.e. $\binom{n}{2} = n(n-1)/2$ pairs with n the total number of storyboards. In our experiment $n = 12$ leading to 66 comparisons. This full design, as opposed to incomplete ones, is more time consuming for individual participants but allows a complete

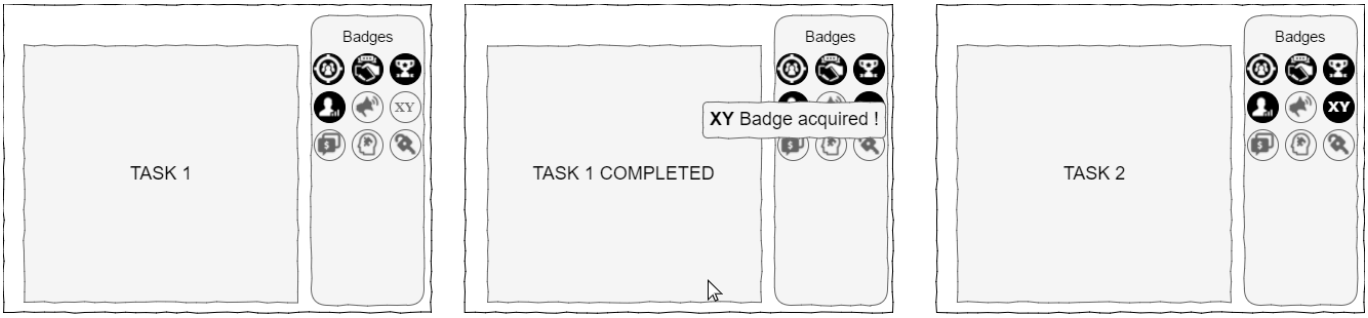

Figure 1: An example of one of the 12 storyboards created to illustrate the game elements that implement the 5 motivational strategies. This particular one depicts the Badges game element related to the Rewards strategy.

evaluation of participant agreement and consistency. Note that the order of pairs is randomized for each participant.

Data collection

As commonly used in HCI research (e.g., [35, 60, 63]), we leveraged the power of crowdsourcing to recruit a large number of participants. Our study used the Figure Eight platform [1] and proposed a task divided into two parts: firstly, participants were asked to complete the paired-comparison experiment described above; secondly, they were asked to fill questionnaires allowing us to determine their BrainHex, Hexad and Big Five profile. For the first two, we used the official questionnaires [74, 37], for the last one, we used a simplified version of traditional big five questionnaires called TIPI (Ten Item Personality Measure) [26]. Using the tools provided by the crowdsourcing platform we ensured that our participants came from a wide variety of different countries and that they could respond to our task only once. Participants were paid a total of US$1.25, and spent between 15 and 25 minutes to complete the surveys.

Data filtering

As with all crowdsourced studies, certain measures are required to ensure that the responses given by participants are genuine. We employed two mechanisms to filter careless participants (we did not filter any of the answers to the personality or player type questionnaires):

(1) In the pairwise comparison task, we inserted four "test" questions where participants were expected to answer a certain way. For example, one "test" question presented a situation where a user would gain 20 points, and another where the user would lose 10 points for performing the same actions. Participants were therefore expected to chose the first one as the more motivational. Participants with less than three correct answers to these four test questions were rejected.
(2) To evaluate the reliability of each participant, we checked their individual consistency, by calculating the number of cyclic triads occurring in their choices. A cyclic triad occurs when a pair comparison is intransitive, (e.g., A is preferred to B, B is preferred to C and C is preferred to A). The coefficient of consistency [42] is then computed as follows for each observer: $\zeta = 1 - \frac{24c}{n^3-n}$ where n is the number of stimuli (12 in our experiment) and c the number of cyclic triads. $\zeta = 1$ when there is no circular triads (i.e. perfect consistency) and will decrease to zero as the number of circular triads, and thus the inconsistency, increases. Participants' results were rejected if their coefficient of consistency was inferior to 0.75, as in [48]. This limit was decided to allow for some degree of input error (i.e. clicking on the wrong button) whilst still removing the most inconsistent participants.

A total of 616 participants performed the whole task; 180 were rejected according to (1) and 136 were rejected according to (2) giving a final set of 300 valid and consistent participants. This strict filtering insures a high reliability of our results. Participants came from a wide variety of backgrounds, a summary of the demographic information and distribution of the different user types is presented in the supplementary materials.

Data Analysis

We describe in this section the different tools used for data analysis. Note that we used Bonferroni's correction to compensate for the multiple comparisons in our statistical tests.

Perceived motivation score

As explained in the Procedure, each participant provides a "vote" for a storyboard for each of the $\binom{n}{2}$ possible pairs ($n = 12$ in our case). Results per participant can be recorded in a $n \times n$ preference matrix. These per-participant preference matrices are then summed into a single one. In this summed matrix P, each element $P_{i,j}$ represents the number of times the storyboard i was judged to be more motivating than storyboard j. An example of such a matrix is given in the supp. mat.

As classically done with pairwise comparison experiments [65], we can then consider the number of votes received by each storyboard as its score of perceived motivational impact, which may then be divided by the number of comparisons per storyboard for normalization purposes. This score computation can be done either for each individual participant, or for groups of participants (e.g., for calculating the preference scores for each dominant user type).

Note that more sophisticated statistical methods exist for inferring scale values from a preference matrix [4, 72]. However they were not shown to give a better representation of perceived motivational impact score than the vote counts. For example, we observed an average correlation of 0.999 (SD:

[1] https://www.figure-eight.com/

0.0002) between scores obtained by vote counts and Thurstone's Law of Comparative Judgements, Case V [72].

Participant agreement

Beyond motivational scores, it is also interesting to analyze the agreement of participants in their choices, i.e. the similarity of their votes. The coefficient of agreement u was defined by Kendall and Smith [42] as: $u = \frac{2\Sigma}{\binom{s}{2}\binom{n}{2}} - 1$ where s is the number of participants and Σ is the sum of the number of agreements between all $\binom{s}{2}$ possible pairs of participants and $\binom{n}{2}$ possible pairs of stimuli. It ranges from 1 (perfect agreement) to $-1/(s-1)$, if s is even, and $-1/s$, if s is odd.

PLS-PM

To calculate how well each user type affects the scores for each implementation we used a method called *partial least squares path modelling* (*PLS PM*) [28]. PLS PM is a method of structural equation modelling which allows estimating complex cause-effect relationship models with latent variables, that has been previously used in studies on the effects of gamification on user motivation [62, 63, 64, 74]. Essentially we use it to see how the values for each user type influence the scores for each game element. The influence values vary between -1 and 1 depending on how strong the effect is. As this is a statistical evaluation we use the calculated p-value to determine the validity of the given influences.

RESULTS

R1: Perceived motivation for different implementations of motivational strategies

As explained in Section 2, a given motivational strategy can be implemented in the form of different game elements. To investigate if different implementations of a same motivational strategy lead to different levels of perceived motivation, we analyzed the motivation scores obtained for the game elements on the entire set of participants. Scores are computed from the preference matrix as explained in Section 3. Instead of a single score per game element, we compute a score distribution using a bootstrap technique [16]: scores are computed 200 times, each time on a random set of participants of the same size as the original set, generated by sampling with replacement. The bootstrap distributions allows for statistical testing and their percentiles provide the 95% confidence intervals.

Figure 2 illustrates the perceived motivation scores. For each motivational strategy, we performed pairwise paired t-tests over the score distributions to assess if significant differences exist between game elements.

The Rewards strategy shows highly significant differences among the motivational impact of its implementations. Badges is the best perceived, followed by Useful Rewards and finally Points (Badges-Points t:137.56, Points-Useful t:-60.25, Useful-Badges t:-114.23, p<.001). Similar results have been found by Denny et al. [12] who showed that in educational settings their badge system was more effective than their points system. They attribute the differences to the fact that the points setting "lacked clear targets". Several other studies have also reported the efficacy of badge systems [1, 11, 27]. Studies that show an effectiveness of Points [9, 20] integrate this game element

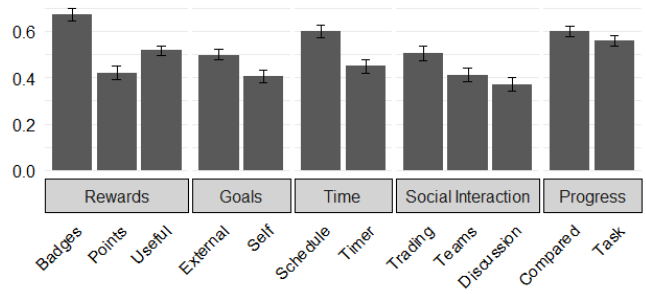

Figure 2: Perceived motivational impact scores of all game elements, for the whole set of participants. The error bars show the 95% confidence intervals. Full values are provided in the supp. mat.

with others like leaderboards or badges, thus preventing an isolated impact of this particular game element.

Regarding the Goals strategy, the difference between the two implementations is also significant (t:70.34, p<.001). Externally set Goals are perceived as more motivating than Self set Goals. This could be explained by the fact that users may find it more difficult to set their goals by themselves, especially without a specific (and meaningful) task to carry out. The studies in [15, 45] also show that External Goals are effective means for user performance and motivation.

Regarding the Time strategy, Timers score less than Schedule (t:-84.22, p<.001). We believe this to be due to the fact that Timers are usually seen as stressful for most users. In our storyboard, the Schedule game element shows the tasks accomplished on a week, users have more time to carry out their tasks and could perceive it as less stressful.

The three implementations of Social Interaction scored differently, with Trading scoring the highest, and Discussion scoring the lowest (Discussion-Teams t:-24,34, Teams-Trading t:-45.33, Trading-Discussion t:75.67, p<.001). Teams and Discussion are in the three least motivating game elements. Young [78] showed that discussion based interventions can be effective in situations where the users are intrinsically motivated by the task.

Finally the Progress implementations rank closely but still show a significant difference in scores (t:46.42, p<.001), with Progress Compared scoring higher. It is noteworthy that both game elements are ranked in the top four. Some previous studies also show the effectiveness of social comparison [31] and of progress bars [19].

As a conclusion, our results demonstrate that **user motivation varies significantly with the different implementations of a same motivational strategy**.

R2 a: Reliability of dominant user type

Many studies in tailored gamification or adaptive games consider only the dominant user type [8, 21, 24, 46, 71] defined as the type that scores the highest for a given user profile.

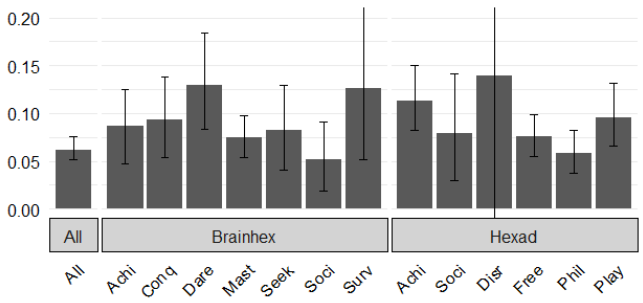

Figure 3: Coefficients of agreement for each dominant type. The error bars show the 95% confidence intervals. Full values are provided in the supp. mat.

To evaluate the reliability of the dominant user type for tailored gamification, we looked at how they affect the perceived motivational impact of the game elements. To do so, we clustered participants according to their dominant user type (for both of the player typologies, as Big Five cannot be considered in this manner) and calculated the Kendall Coefficients of Agreement u within each group and compared to the global value (obtained on the whole set of participants). The idea is to study if users sharing the same dominant type perceive the same motivational impact of game elements (and thus show a higher agreement than whole set). Figure 3 shows the results.

We find the coefficients of agreement within each dominant user type cluster to be low (no group scored an agreement greater than 0.15). While most of these values are nevertheless higher than the global one (0.062), this still shows that **dominant user types cannot be considered sufficient to differentiate users according to their game element preferences.** A recent study by Loria et al. also confirms this finding [18].

R2 b: Comparing user models

In this section we investigate more precisely the relationships between the user models (both player types and personality model) and game elements to identify which user typology is the most relevant to identify user preferences for game elements. Table 2 shows the PLS path coefficients that reflect the influence that each user model dimension has on the motivation score of the different game elements. We present our results grouped by user typology, and we discuss the extent to which our results are in line with the definition of each user type (the definitions are provided in the supp. mat.).

BrainHex

Daredevil affects *positively* two game elements of the Rewards motivational strategy: **Badges** and **Useful Rewards**. According to the definition, Daredevils appreciate "rushing around at high speed whilst still in control" [57]. Badges may reinforce the feeling of control and knowledge on the system, and Useful Rewards may help them to speed up progression.
The **Socialiser** type influences two game elements: **Discussion** *positively*, and **Progress Task** *negatively*. Nacke et al. [57] define Socialisers as "liking hanging around with, and helping people". This definition therefore tends to confirm that Socialisers are motivated by discussion. However, it does not explain the influence of this type on Progress Task.
The **Conqueror** user type has a significant *positive* influence on **Teams**, as well as a significant *negative* influence on **External Goals**. These influences are unrelated to the definition as conquerors are people who "like defeating impossibly difficult foes, struggling until they eventually achieve victory" [57].
The **Achiever** user type has two *negative* influences: **Timer** and **Progress Task**. Achievers "like collecting anything they can collect, and doing anything possible" [57]. The presence of a **timer** could hinder their abilities to achieve this. However the definition of Achiever does not provide any explanation for the negative influence on Progress Task.
The **Survivor** user type shows a significant *negative* influence on the **Progress Compared** game element which is unrelated to the definition stating that survivors enjoy "escaping from terrifying situations" [57].
Seekers, defined as people who have interest in "finding strange and wonderful things" [57], have no significant influences on any of the game elements tested.
Finally, **Mastermind** have also no significant influences. Nacke et al. [57] define them as people who "like solving puzzles and devising strategies", meaning that they might be more motivated by the task itself than the game elements used.

As a conclusion, for BrainHex, five user types have significant influences on the different game elements. Some of our results can be explained using the definition of BrainHex typology [57], especially for **Daredevil**, and partially for **Socialiser** and **Achiever**. However most of our results cannot be backed up by the definitions given in the typology. In addition, Mastermind and Seeker user types definitions seem not well suited for gamification. This result is in line with recent empirical investigation on the psychometric properties of BrainHex that has shown low reliability scores [5, 23]. This typology was built for games and there is no evidence of the generalizability of game motivation models to gameful design [73].

Big Five

The **Openness to experiences** trait has a significant *positive* influence on **Teams**. The appreciation for new ideas and curiosity [25] are two characteristics of this personality trait that can explain this result.
Regarding the **Agreeableness** trait, we observe a *positive* influence on **Discussion** and a *negative* influence on **Progress Compared**. The positive influence on social interactions is consistent with the definition stating that people with high agreeableness are generally generous, cooperative and helpful [25]. However the definition of this trait cannot explain why these people are demotivated by game elements related to the progression in the task.
Emotional stability shows no significant influences on any of the game elements. People with high emotional stability have a tendency to resist negative emotions such as anger or anxiety [25]. We can assume that people with this personality trait are not responsive to the game elements in terms of motivation.
The **Conscientiousness** trait influences significantly only one game element, **Trading** *negatively*. This result is not related to the definition of this trait since conscientious people are defined as self-disciplined and well organised.
Finally, **Extraversion** shows no significant influences. **Extraversion** is defined as a "pronounced engagement with the

		BrainHex							Hexad						Big Five				
Strategy	Game element	Seek	Surv	Dare	Mast	Conq	Soci	Achi	Phil	Soci	Free	Achi	Disr	Play	Extr	Agre	Cons	Emot	Open
Rewards	Badges	.01	-.01	.13	.11	.07	-.07	.02	-.08	.07	.04	.05	-.06	-.07	.00	-.01	.04	-.02	.00
	Points	-.12	.02	.01	-.02	-.08	-.05	-.06	-.06	-.14	.06	-.16	.00	.10	.05	-.07	.04	.04	-.06
	Useful	.11	-.04	.13	-.03	-.08	.02	.10	.03	-.13	.09	-.07	-.02	-.08	-.02	.01	-.12	.03	-.04
Goals	External	-.07	.04	-.04	.09	-.16	.02	.02	.13	-.30	.05	.10	-.05	-.01	.00	.03	.07	-.03	-.04
	Self	.06	.10	-.09	.01	-.10	.10	.01	.12	-.08	.05	.09	.00	-.08	.06	.05	.03	-.03	-.02
Time	Schedule	-.02	.03	-.05	.02	.04	-.07	-.00	.07	-.10	-.07	.03	-.08	-.04	-.05	.04	.06	.03	-.07
	Timer	-.08	.01	.02	-.05	.04	-.07	-.14	-.03	-.04	-.15	.13	.06	.01	.03	-.05	.08	-.03	.06
Social Interaction	Trading	.01	.06	.02	-.05	.10	.11	.04	-.06	.26	.05	.03	-.01	-.06	.03	.12	-.15	.10	.04
	Teams	.02	-.09	.01	-.05	.16	.04	.09	-.05	.34	.01	-.10	-.07	.03	-.05	-.04	-.06	-.06	.19
	Discussion	.08	.07	-.10	.01	.01	.18	.06	.02	.24	.13	-.11	-.07	-.01	-.03	.14	-.06	-.02	.04
Progress	Compared	.01	-.13	.03	.04	-.01	-.08	.01	.02	-.14	-.14	.08	.16	.08	-.02	-.14	.02	.02	-.05
	Task	.02	-.09	-.06	-.06	-.09	-.15	-.12	-.07	-.16	-.11	-.06	.17	.07	-.02	-.12	.07	-.06	-.10

Table 2: PLS Path coefficients for each user type of each typology. Values in grey are not significant (p>.05), highlighted in dark grey are significant (p<.05), and highlighted in black are highly significant (p<.001).

external world and enjoyment from interacting with people" [25]. To fit the definition, we would have expected a positive influence for game elements that implement the Social Interaction strategy.

As a conclusion regarding Big Five personality model, three traits have significant influences. As with BrainHex some of our results can be partially explained by the definition of the personality traits (**openness to experiences** and partially **agreeableness**), but most of them are not directly in line with the definitions. This result was predictable since Big Five is a general personality trait model and not specifically developed for games or gamification.

Hexad

The **Socialiser** player type affects *positively* the three Social Interaction game elements (**Trading**, **Discussion** and strongly **Teams**) and *negatively* **External** Goals and **Progress Task**. Regarding social interactions, our results are consistent with the definition that states that Socialisers "like to interact with others and create social connections" [51].

Disruptor has significant *positive* influence on both **Progress Compared** and **Progress Task**. This could be explained as disruptors seek to change a system [51]. Perceiving the boundaries of the system thanks to progress elements could help them to expand beyond these limits.

The **Achiever** user type has only a significant *negative* influence on **Points**. Some authors point out that reward systems (specifically points) can be perceived as useless if their implementation is not linked to the context [39]. As Achievers are motivated by competence [51], we can assume that they do not appreciate points as illustrated in our scenarios.

Free Spirit has only a significant *negative* influence on **Timer**. According to the definition, Free spirits are motivated by autonomy. They like to explore within a system and act without external control [51]. In this case, Timer can be perceived as constraining their freedom by time.

The **Philanthropist** user type has no effect on the motivation scores given for the different game elements. According to the definition [51], philanthropists are motivated by purpose. Thus we can think that this user type does not influence any preferences for our elements that are not connected to a specific user activity and domain.

Finally, we did not find any significant influences for the **Player** user type. Players are defined as being motivated by extrinsic rewards. They will do anything to earn a reward within a system, independently of the type of the activity [51]. Players seem to be able to appreciate anything, and therefore will react positively to almost any game element used, explaining that we do not find any particular influence.

As a conclusion regarding Hexad, four user types have significant influences, among which one is highly significant. Moreover, most of our results are consistent with the definitions of the Hexad typology [51]. This result reinforces the fact that this typology was designed especially for gamification and most of its player types are based on SDT [68], the major theoretical foundation for gamification research.

Seeing as our results with Hexad are the most consistent with the definitions of its user types, and that its types have more influence on the perceived user motivation than those from BrainHex and Big Five, we can state that **Hexad is the most relevant typology to identify user preferences for game elements and thus should be used to tailor gamification.**

R3: Activity and domain influence on the motivational impact of game elements

We finally compare our results to the findings of previous studies when their game elements or persuasive strategies are similar to ours (see table 1 for correspondences). Previous studies were conducted in specific contexts (specific domain and user activity), with possible influence of users' intrinsic motivation for the activity and/or domain on the observed user motivation. It is noteworthy that whilst the study in [64] focuses on a serious game, it uses game elements in a similar way as a gamified system. In our study, we use storyboards that show implementations of motivational strategies independent from any specific context. Our analysis in this section aims to identify the extent to which the context has an influence on the motivational impact of game elements according to user types.

BrainHex

Our results are *consistent* with other studies for the **Socialiser** type, for which Orji et al. [64] also found a negative influence

on self-monitoring (vs. Progress Task).
Three of our results on the **Achiever** and **Survivor** user types *contradict* the previous studies. Regarding **Achievers**, Orji et al. [64] found a positive influence on self-monitoring, whereas we find a negative influence on Progress Task and Lavoué et al. [46] predict a positive link with timer, whereas we find a negative influence. Concerning **Survivors**, Orji et al. [64] found a positive influence on competition & comparison, whereas we find a negative influence on Progress Compared.
We also find influences for game elements that are *not identified* in previous work for the **Daredevil** type. Finally, other studies also found *other influences* for the 7 user types.

We can conclude that our results obtained with the BrainHex user typology are quite different from the other studies conducted in specific contexts (gamified health system [64] and experts' recommendations in education [46]).

Big Five
Regarding the Big Five personality model, the positive influence we find of the **Agreeableness** trait on Discussion *is consistent* with the results of Orji et al. [62] on cooperation. Regarding **Emotional stability**, Orji et al. [62] studied people with a low emotional stability and also did not find any influences for those people. They stated that persuasion may not be effective for people who are emotionally unstable. In the same domain (health) but for a different activity, Jia et al. [39] found negative influence on points, badges, progress and rewards and they argued that for people with high emotional stability gamification may not be an effective approach.
Our results *contradict* previous studies only for **Openness to experiences**, Orji et al. [62] found a negative relation with cooperation, whereas we find a positive influence on Teams.
We also find a negative influence of **Conscientiousness** on Trading that is *not identified* in previous studies.
The three comparable studies also found *other influences* for **Extraversion**, **Agreeableness** and **Conscientiousness**, and the two studies held in the health domain do not find similar results. For instance for **Agreeableness**, Orji et al. [62] found influences on four game elements (self-monitoring & feedback, comparison, competition and reward) whereas no influences were found for these in Jia et al. [39].

Regarding the Big Five personality model, the comparisons with other studies highlight the differences in the results obtained both (1) between our study and studies conducted in a specific context, and (2) between the studies conducted in different contexts, even in the same domain like health.

Hexad
Our results are *consistent* with other studies for the **Philanthropist** user type, which do not find any influence on the motivational impact of game elements.
Our results are *partially consistent* both for the **Socialiser** and **Disruptor** user types. Regarding **Socialisers** and positive influences on social game elements, [74] also suggest positive influences on Teams, Social Network and Social Comparison, while Orji et al. [63] also found positive influences for Cooperation. Regarding **Disruptors**, all studies including ours observed positive influences for the **Progress Compared** element (Competition in [63, 74]). We observe *contradictions* only with [63] for these two player types regarding the **External Goals** and **Progress Task** elements. Our results are different from previous studies regarding three player types. We find influences for game elements that are *not identified* in previous works and other studies found *other influences* for **Achiever** and **Free Spirit** (previous work do not show any influence for elements comparable to the timer). Finally, regarding **Players**, we do not find any influences whereas [74] found positive influences for every game elements and [63] for most of game elements (except for Self-Monitoring & Feedback and Goal Setting & Suggestion).

As with Big Five, the results with the Hexad types are quite different from those found in studies conducted in specific different contexts (education [74] and health [63]), except for three user types (Philanthropist, Socialiser and Disruptor).

In conclusion, for all user typologies (BrainHex, Hexad) and the personality model (big five) we can see that **the motivational impact of certain game elements varies according to the activities or the domain of gamified systems**. In fact, our results are quite different (especially for BrainHex) from those found in contextualised studies. We also observe differences between the studies conducted in different contexts, even between two studies conducted in the same general domain (Health) with the Big Five user typology.

LIMITATIONS

About the experimental protocol
Regarding the experimental protocol, one major difference with previous studies deals with how we collected our data. We used forced-choice paired comparison instead of declarative statements to identify the perceived impact of game elements on user motivation. As stated in Section 3, this protocol has been shown to be less cognitively demanding [10] and to provide higher accuracy [70] as compared with Likert-type rating. We notice that it also impacts our results as it forces the users to choose which game elements they prefer. We therefore obtain a ranking of game elements (meaning that a user could not vote all game elements as equally motivating). We believe this has a direct consequence on two profile dimensions (Hexad-Player, and Big Five-Extraversion) for which people will tend to appreciate most of game elements, leading to no specific significant influences when applying the PLS PM method. In addition, using PLS PM allows us to observe negative influences on the perceived motivational impact of game elements whereas some of the other studies only measure positive influences [46, 74]. This could be a limitation to the comparisons we make in Section 4.4.

About the context-independent scenarios
We use context-independent scenarios to evaluate the motivational impact of game elements independently from a specific context. However certain game elements may be less motivating for users shown without a concrete task to carry out. We suppose this to be especially true for External Goals, Self Goals and Useful Rewards that are not perceived as the most motivating elements in our study and Progress Task that is negatively influenced for three user types. Furthermore, for some user types, no significant preferences can be shown without a

specific task to carry out. We think this is especially true for the philanthropist who is motivated by purpose.

About the implementation of motivational strategies

Finally, the results we obtain may differ from other studies due to the fact that the implementations of our motivational strategies may be quite different from those used in other studies. For instance, the Rewards strategy in [63] is implemented in the form of points that can be used to unlock new customisation options.This implementation can be considered as a combination of our Points and Useful Rewards (see Table 1). We believe that our approach allows us to study the isolated motivational impact for each game element more precisely.

IMPLICATIONS FOR DESIGN: TOWARDS TAILORED GAMIFICATION

According to our findings, we recommend to consider two main factors when designing tailored gamification: the choice of the user typology and the implementation of the motivational strategies. In this section we discuss both factors, considering also the influence of the context of the gamified system.

User typology recommendations

First, we recommend using **Hexad user types** when designing tailored gamification. Our results reinforce the fact that Hexad was created especially to address gamification (compared to BrainHex which focuses on player types in games and Big Five which focuses on personality traits).

Second, our results show that considering only the dominant player type is not sufficient to discriminate users' preferences. We thus advise designers to consider users' profiles as a combination of several player types, especially the four that identify the most significant influences: **Socialiser, Disruptor, Achiever** and **Free spirit**. The **Philanthropist** type does not appear to identify motivating game elements in any of the studies in a discriminant way. The **Player** type is similar as it either shows no significant influences (in our study) or significant influences for all [74] or most game elements [63].

Finally, the comparison of our results with contextualized studies reveals that the motivational impact of game elements for the Hexad user types is more or less influenced by the context. For Socialisers, designers should preferably implement social interactions (recommendation also found in the other context-independent study [73]). For Disruptors, progress compared is recommended. The design of these game elements can be made independently from the context. Whereas Achievers and Free-spirits have contradictory preferences according to the different studies. This can be explained by the fact that these two user types are highly dependent on the activity or on the environment of the gamified system, meaning that the design of game elements should take into account the context of the gamified system for these types of users.

Game element recommendations

First, care needs to be taken when implementing a motivational strategy. As shown in Section 4, different implementations of a same motivational strategy have different impacts on user motivation. **Rewards** is a good illustration of this: **Badges** is the highest rated of any of the implementations, whereas **Points** is one of the lowest rated ones (the same can be said for **Timer** and **Schedule**).

Whilst we cannot precisely recommend game elements for each user profile, we can make recommendations based how user motivation should vary with a game element using Hexad types. **Badges** and **Schedules** can be used as motivating game elements for all users. These were two of the highest scoring game elements and had no negative influences from any user types in both our study and related studies. Designers can therefore feel confident that these game elements will have no adverse effects on user motivation.
Progress game elements (**Compared** and **Task**) and **External Goals** are generally considered as motivating and could be used for various user types. In particular, **Progress compared** game element is recommended for users who have a high score in the *Disruptor* type. At the contrary, we recommend to carefully use **Progress Task** for high *Socialiser* users, as a negative influence was observed and also for high *Disruptor* users since we found contradictory results with previous studies. All three of the social interaction game elements (**Trading**, **Teams**, and **Discussion**) are generally perceived as less motivating, except for the *Socialiser* type which shows positive influences in all studies. We thus recommend to attribute these game elements only to high *Socialiser* users. **Points** and **Timer** show low motivation scores and the only influences they have from the various profile types are negative ones. We can therefore advise against using these game elements to motivate users. Finally, no influences were found for **Self goals** and **Useful items** which means that these elements are probably highly dependent from the context. They should be designed closely with the activity to be perceived as motivating game elements.

CONCLUSION

In this paper we showed that two major factors influence user motivation in tailored gamification: the implementation of a given motivational strategy, and the choice of the user typology. We also studied the influence of the user activity and domain of the gamified system to be able to identify general recommendations for designers. We found valuable insights on how to tailor gamification, notably by showing that the Hexad user typology seems the most relevant to identify user preferences for game elements. We recommend designers to consider several user types, not only the dominant one, when assigning game elements to users. We also provide designers with recommendations on game elements to assign to specific user types, independently or dependently from the context of the gamified system. Our future work will focus on three axes 1) the identification of clusters of hybrid user profiles to consider the impact of a combination of user type scores , 2) the development of an adaptive gamification system based on the recommendations from this study, and 3) the inclusion of user behaviour analysis to support a dynamic adaptation of game elements.

ACKNOWLEDGEMENTS

This work is a part of the LudiMoodle project financed by the e-FRAN Programme d'investissement d'avenir, operated by the Caisse des Dépots.

REFERENCES

[1] Ashton Anderson, Daniel Huttenlocher, Jon Kleinberg, and Jure Leskovec. 2013. Steering user behavior with badges. In *Proceedings of the 22nd international conference on World Wide Web*. ACM, 95–106.

[2] Yigal Attali and Meirav Arieli-Attali. 2015. Gamification in assessment: Do points affect test performance? *Computers & Education* 83 (April 2015), 57–63. DOI: http://dx.doi.org/10.1016/j.compedu.2014.12.012

[3] Chris Bateman, Rebecca Lowenhaupt, and Lennart Nacke. 2011. Player Typology in Theory and Practice.. In *DiGRA Conference*.

[4] Ralph Allan Bradley and Milton E. Terry. 1952. Rank Analysis of Incomplete Block Designs: I. The Method of Paired Comparisons. *Biometrika* 39, 3 (1952), 324–345.

[5] Marc Busch, Elke Mattheiss, Rita Orji, Peter Fröhlich, Michael Lankes, and Manfred Tscheligi. 2016b. Player Type Models: Towards Empirical Validation. In *CHI Extended Abstracts '16*. ACM, 1835–1841. DOI: http://dx.doi.org/10.1145/2851581.2892399

[6] Marc Busch, Elke Mattheiss, Rita Orji, Andrzej Marczewski, Wolfgang Hochleitner, Michael Lankes, Lennart E. Nacke, and Manfred Tscheligi. 2015. Personalization in serious and persuasive games and gamified interactions. In *CHIPLAY '15*. ACM, 811–816.

[7] Marc Busch, Elke E Mattheiss, Wolfgang Hochleitner, Christina Hochleitner, Michael Lankes, Peter Fröhlich, Rita Orji, and Manfred Tscheligi. 2016a. Using Player Type Models for Personalized Game Design-An Empirical Investigation. *IxD&A* 28 (2016), 145–163.

[8] Geiser Chalco Challco, Dilvan A Moreira, Riichiro Mizoguchi, and Seiji Isotani. 2014. An ontology engineering approach to gamify collaborative learning scenarios. In *CYTED-RITOS International Workshop on Groupware*. Springer, 185–198.

[9] Christopher Cheong, France Cheong, and Justin Filippou. 2013. Quick Quiz: A Gamified Approach for Enhancing Learning.. In *PACIS*. 206.

[10] Andrew P Clark, Kate L Howard, Andy T Woods, Ian S Penton-voak, and Christof Neumann. 2018. Why rate when you could compare ? Using the " EloChoice " package to assess pairwise comparisons of perceived physical strength. *PLoS ONE* 13, 1 (2018), 1–16. DOI: http://dx.doi.org/10.1371/journal.pone.0190393

[11] Paul Denny. 2013. The effect of virtual achievements on student engagement. In *CHI '13*. ACM, 763–772.

[12] Paul Denny, Fiona McDonald, Ruth Empson, Philip Kelly, and Andrew Petersen. 2018. Empirical Support for a Causal Relationship Between Gamification and Learning Outcomes. In *CHI '18*. ACM, 311. DOI: http://dx.doi.org/10.1145/3173574.3173885

[13] Sebastian Deterding. 2015. The lens of intrinsic skill atoms: A method for gameful design. *Human–Computer Interaction* 30, 3-4 (2015), 294–335.

[14] 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.

[15] Tao Dong, Mira Dontcheva, Diana Joseph, Karrie Karahalios, Mark Newman, and Mark Ackerman. 2012. Discovery-based games for learning software. In *CHI '12*. ACM, 2083–2086.

[16] Bradley Efron and R.J. Tibshirani. 1993. *An Introduction to the Bootstrap*. Chapman & Hall/CRC Monographs on Statistics & Applied Probability.

[17] Carsten Eickhoff, Christopher G. Harris, Arjen P. de Vries, and Padmini Srinivasan. 2012. Quality through flow and immersion: gamifying crowdsourced relevance assessments. In *SIGIR '12*. ACM, 871–880.

[18] Loria Enrica and Marconi Annapaola. Player Types and Player Behaviors: Analyzing Correlations in an On-the-field Gamified System. In *Extended abstracts CHI Play* (2018). 531–538. DOI: http://dx.doi.org/10.1145/3270316.3271526

[19] Rosta Farzan and Peter Brusilovsky. 2011. Encouraging user participation in a course recommender system: An impact on user behavior. *Computers in Human Behavior* 27, 1 (2011), 276–284.

[20] Rosta Farzan, Joan M. DiMicco, David R. Millen, Beth Brownholtz, Werner Geyer, and Casey Dugan. 2008. When the experiment is over: Deploying an incentive system to all the users. In *symposium on persuasive technology*.

[21] Lauren S. Ferro, Steffen P. Walz, and Stefan Greuter. 2013. Towards personalised, gamified systems: an investigation into game design, personality and player typologies. In *Proceedings of The 9th Australasian Conference on Interactive Entertainment: Matters of Life and Death*. ACM.

[22] David R. Flatla, Carl Gutwin, Lennart E. Nacke, Scott Bateman, and Regan L. Mandryk. 2011. Calibration games: making calibration tasks enjoyable by adding motivating game elements. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*. ACM, 403–412.

[23] Gustavo Fortes Tondello, Deltcho Valtchanov, Adrian Reetz, Rina R. Wehbe, Rita Orji, and Lennart E. Nacke. 2018. Towards a Trait Model of Video Game Preferences. (2018), 1–17.

[24] Borja Gil, Iván Cantador, and Andrzej Marczewski. 2015. Validating gamification mechanics and player types in an E-learning environment. In *Design for Teaching and Learning in a Networked World*. Springer, 568–572.

[25] Lewis R. Goldberg. 1992. The development of markers for the Big-Five factor structure. *Psychological assessment* 4, 1 (1992), 26.

[26] Samuel D. Gosling, Peter J. Rentfrow, and William B. Swann Jr. 2003. A very brief measure of the Big-Five personality domains. *Journal of Research in personality* 37, 6 (2003), 504–528.

[27] Scott Grant and Buddy Betts. 2013. Encouraging user behaviour with achievements: an empirical study. In *Proceedings of the 10th Working Conference on Mining Software Repositories*. IEEE Press, 65–68.

[28] Joseph F Hair Jr, G Tomas M Hult, Christian Ringle, and Marko Sarstedt. 2016. *A primer on partial least squares structural equation modeling (PLS-SEM)*. Sage Publications.

[29] Stuart Hallifax, Audrey Serna, Jean-Charles Marty, and Élise Lavoué. 2018. A Design Space For Meaningful Structural Gamification. In *CHI '18 Extended Abstracts*. ACM, LBW073.

[30] Juho Hamari. 2015. Do badges increase user activity? A field experiment on the effects of gamification. *Computers in Human Behavior* (2015). DOI: http://dx.doi.org/10.1016/j.chb.2015.03.036

[31] Juho Hamari and Jonna Koivisto. 2013. Social Motivations To Use Gamification: An Empirical Study Of Gamifying Exercise. In *ECIS*, Vol. 105.

[32] Juho Hamari, Jonna Koivisto, and Tuomas Pakkanen. 2014a. Do persuasive technologies persuade?-a review of empirical studies. In *International Conference on Persuasive Technology*. Springer, 118–136.

[33] Juho Hamari, Jonna Koivisto, and Harri Sarsa. 2014b. Does gamification work?–a literature review of empirical studies on gamification. In *Hawaii International Conference on System Sciences (HICSS)*. IEEE, 3025–3034.

[34] 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 (Jan. 2015), 152–161. DOI: http://dx.doi.org/10.1016/j.compedu.2014.08.019

[35] Jeffrey Heer and Michael Bostock. 2010. Crowdsourcing Graphical Perception: Using Mechanical Turk to Assess Visualization Design. In *CHI '10*. ACM, 203–212. DOI: http://dx.doi.org/10.1145/1753326.1753357

[36] Kai Huotari and Juho Hamari. 2012. Defining Gamification: A Service Marketing Perspective *(MindTrek '12)*. ACM, 17–22. DOI: http://dx.doi.org/10.1145/2393132.2393137

[37] international hobo. 2010. BrainHex questionnaire. http://survey.ihobo.com/BrainHex/. (2010). Accessed: 2019-04.

[38] Shih-Ping Jeng and Ching-I. Teng. 2008. Personality and motivations for playing online games. *Social Behavior and Personality: an international journal* 36, 8 (2008), 1053–1060.

[39] Yuan Jia, Bin Xu, Yamini Karanam, and Stephen Voida. 2016. Personality-targeted Gamification: A Survey Study on Personality Traits and Motivational Affordances. In *CHI '16*. ACM, 2001–2013. DOI: http://dx.doi.org/10.1145/2858036.2858515

[40] Karl M. Kapp. 2012. *The gamification of learning and instruction: game-based methods and strategies for training and education*. John Wiley & Sons.

[41] Maurits Kaptein, Boris De Ruyter, Panos Markopoulos, and Emile Aarts. 2012. Adaptive persuasive systems: a study of tailored persuasive text messages to reduce snacking. *ACM Transactions on Interactive Intelligent Systems (TiiS)* 2, 2 (2012), 10.

[42] Maurice G Kendall and B Babington Smith. 1940. On the method of paired comparisons. *Biometrika* 31, 3/4 (1940), 324–345.

[43] Janaki Kumar. 2013. Gamification at work: Designing engaging business software. In *International conference of design, user experience, and usability*. Springer, 528–537.

[44] Richard N. Landers and Michael B. Armstrong. 2015. Enhancing instructional outcomes with gamification: An empirical test of the Technology-Enhanced Training Effectiveness Model. *Computers in Human Behavior* (Sept. 2015). DOI: http://dx.doi.org/10.1016/j.chb.2015.07.031

[45] 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 (June 2017), 508–515. DOI: http://dx.doi.org/10.1016/j.chb.2015.08.008

[46] Élise Lavoué, Baptiste Monterrat, Michel Desmarais, and Sébastien George. 2018. Adaptive Gamification for Learning Environments. *IEEE Transactions on Learning Technologies* (2018).

[47] Tuomas Lehto and Harri Oinas-Kukkonen. 2011. Persuasive features in web-based alcohol and smoking interventions: a systematic review of the literature. *Journal of medical Internet research* 13, 3 (2011), e46.

[48] FB Leloup, MR Pointer, Philip Dutré, and Peter Hanselaer. 2010. Geometry of illumination, luminance contrast, and gloss perception. *Journal of the Optical Society of America* 27, 9 (2010), 2046–2054.

[49] Jemma Looyestyn, Jocelyn Kernot, Kobie Boshoff, Jillian Ryan, Sarah Edney, and Carol Maher. 2017. Does gamification increase engagement with online programs? A systematic review. *PLOS ONE* 12, 3 (2017), 1–19.

[50] Cathie Marache-Francisco and Eric Brangier. 2013. Process of gamification. *Proceedings of the 6th Centric* (2013), 126–131.

[51] A. C. Marczewski. 2015. *Even Ninja Monkeys like to play*. CreateSpace Indep. Publish Platform, Charleston, Chapter User Types, 69–84.

[52] Elisa D. Mekler, Florian Brühlmann, Alexandre N. Tuch, and Klaus Opwis. 2015. Towards understanding the effects of individual gamification elements on intrinsic motivation and performance. *Computers in Human Behavior* (Sept. 2015). DOI: http://dx.doi.org/10.1016/j.chb.2015.08.048

[53] Baptiste Monterrat, Michel Desmarais, Élise 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.

[54] Baptiste Monterrat, Élise Lavoué, and Sébastien George. 2014. Motivation for learning: Adaptive gamification for web-based learning environments. In *CSEDU 2014*. 117–125.

[55] Baptiste Monterrat, élise Lavoué, and Sébastien George. 2017. Adaptation of Gaming Features for Motivating Learners. *Simulation & Gaming* 48, 5 (Oct. 2017), 625–656. DOI: http://dx.doi.org/10.1177/1046878117712632

[56] Alberto Mora, Gustavo F. Tondello, Lennart E. Nacke, and Joan Arnedo-Moreno. 2018. Effect of personalized gameful design on student engagement. In *Global Engineering Education Conference (EDUCON), 2018 IEEE*. IEEE, 1925–1933.

[57] Lennart E. Nacke, Chris Bateman, and Regan L. Mandryk. 2011. BrainHex: preliminary results from a neurobiological gamer typology survey. In *International Conference on Entertainment Computing*. Springer, 288–293.

[58] Scott Nicholson. 2012. A user-centered theoretical framework for meaningful gamification. *Games+ Learning+ Society* 8, 1 (2012), 223–230.

[59] Florin Oprescu, Christian Jones, and Mary Katsikitis. 2014. I PLAY AT WORK—ten principles for transforming work processes through gamification. *Frontiers in psychology* 5 (2014), 14.

[60] Rita Orji, Regan L. Mandryk, and Julita Vassileva. 2017. Improving the Efficacy of Games for Change Using Personalization Models. *ACM Trans. Comput.-Hum. Interact.* 24, 5 (Oct. 2017). DOI: http://dx.doi.org/10.1145/3119929

[61] Rita Orji, Regan L. Mandryk, Julita Vassileva, and Kathrin M. Gerling. 2013. Tailoring persuasive health games to gamer type. In *CHI'13*. ACM, 2467–2476.

[62] Rita Orji, Lennart E. Nacke, and Chrysanne Di Marco. 2017. Towards Personality-driven Persuasive Health Games and Gamified Systems. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, 1015–1027. DOI: http://dx.doi.org/10.1145/3025453.3025577

[63] Rita Orji, Gustavo F. Tondello, and Lennart E. Nacke. 2018. Personalizing Persuasive Strategies in Gameful Systems to Gamification User Types. In *CHI ' 18*, Vol. 61. 62.

[64] Rita Orji, Julita Vassileva, and Regan L. Mandryk. 2014. Modeling the efficacy of persuasive strategies for different gamer types in serious games for health. *User Modeling and User-Adapted Interaction* 24, 5 (2014), 453–498.

[65] Maria Perez-Ortiz and Rafal K. Mantiuk. 2017. *A practical guide and software for analysing pairwise comparison experiments*. Technical Report.

[66] Scott Rigby and Richard M Ryan. 2011. *Glued to games: How video games draw us in and hold us spellbound (New directions in media)*. Praeger Santa Barbara, CA.

[67] David Robinson and Victoria Bellotti. 2013. A preliminary taxonomy of gamification elements for varying anticipated commitment. In *CHI'13 Workshop on Designing Gamification: Creating Gameful and Playful Experiences*.

[68] R. M. Ryan and E. L. Deci. 2000. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *The American Psychologist* 55, 1 (Jan. 2000), 68–78.

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

[70] Nihar B. Shah, Sivaraman Balakrishnan, Joseph Bradley, Abhay Parekh, Kannan Ramchandran, and Martin J. Wainwright. 2016. Estimation from Pairwise Comparisons: Sharp Minimax Bounds with Topology Dependence. *Journal of Machine Learning Research* 17 (2016), 1–47.

[71] David Thue, Vadim Bulitko, Marcia Spetch, and Eric Wasylishen. 2007. Interactive Storytelling: A Player Modelling Approach.. In *AIIDE*. 43–48.

[72] L. L. Thurstone. 1927. A law of comparative judgments. *Psychological review* 34 (1927), 273–286.

[73] Gustavo F. Tondello, Alberto Mora, and Lennart E. Nacke. 2017. Elements of Gameful Design Emerging from User Preferences. In *CHIPLAY'17*. ACM.

[74] Gustavo F. Tondello, Rina R. Wehbe, Lisa Diamond, Marc Busch, Andrzej Marczewski, and Lennart E. Nacke. 2016. The Gamification User Types Hexad Scale. In *CHIPLAY'16*. ACM, 229–243.

[75] Julita Vassileva. 2012. Motivating participation in social computing applications: a user modeling perspective. *User Modeling and User-Adapted Interaction* 22, 1 (2012), 177–201.

[76] Kevin Werbach and Dan Hunter. 2012. *For the win: How game thinking can revolutionize your business*. Wharton Digital Press.

[77] Nick Yee. 2006. Motivations for play in online games. *CyberPsychology & behavior* 9, 6 (2006), 772–775.

[78] Mark Mingyi Young. 2010. Twitter Me: Using Micro-blogging to Motivate Teenagers to Exercise. In *Global Perspectives on Design Science Research*, Robert Winter, J. Leon Zhao, and Stephan Aier (Eds.). Springer Berlin Heidelberg, 439–448.