Gameful Design Heuristics: A Gamification Inspection Tool

Gustavo F. Tondello[1,2(✉)] , Dennis L. Kappen[5] , Marim Ganaba[1],
and Lennart E. Nacke[1,2,3,4]

[1] HCI Games Group, Games Institute, University of Waterloo,
Waterloo, ON, Canada
`gustavo@tondello.com`, `mariganaba@gmail.com`,
`lennart.nacke@acm.org`
[2] Cheriton School of Computer Science, University of Waterloo,
Waterloo, ON, Canada
[3] Department of Communication Arts, University of Waterloo,
Waterloo, ON, Canada
[4] Stratford School of Interaction and Business, University of Waterloo,
Waterloo, ON, Canada
[5] Humber College, Toronto, ON, Canada
`dennis.kappen@humber.ca`

Abstract. Despite the emergence of many gameful design methodologies in the literature, there is a lack of methods to evaluate the resulting designs. Gameful design techniques aim to increase the user's motivation to interact with a software, but there are presently no accepted guidelines on how to find out if this goal was achieved during the design phase of a project. This paper presents the *Gameful Design Heuristics*, a novel set of guidelines that facilitate a heuristic evaluation of gameful software, with a focus on the software's potential to afford intrinsic and extrinsic motivation for the user. First, we reviewed several gameful design methods to identify the most frequently employed dimensions of motivational affordances. Then, we devised a set of 28 gamification heuristics that can be used to rapidly evaluate a gameful system. Finally, we conducted a summative empirical evaluation study with five user experience professionals, which demonstrated that our heuristics can help the evaluators find more motivational issues in interactive systems than they would without the heuristics. The suggested method fulfills the need for evaluation tools specific to gameful design, which could help evaluators assess the potential user experience of a gameful application in the early phases of a project.

Keywords: Gameful design heuristics · Heuristic evaluation ·
User experience · Gamification · Gameful design

1 Introduction

Many gameful design methods have recently emerged as part of the user experience (UX) design toolkit. They aim to augment and improve the UX of interactive systems with gamification—defined as using game design elements in non-game contexts [1].

Even though these tools have been increasingly adopted during the design phase of software projects, designers still lack standard evaluation methods. There are no guidelines for experts (i.e., people with background knowledge in UX) to evaluate a gameful implementation early on in a project.

For usability evaluation, two standard approaches exist. First, the gold standard is a usability test. UX researchers can either run a formative usability test (where they usually sit close to the participant and observe their behaviour) or a summative one (where they are often present locally or virtually, but the participant is working through an assigned task or scenario while some outcome measures are recorded). However, the second type (the heuristic evaluation or usability inspection) is cheaper and easier to set up—and can be conducted before planning an expensive usability test. Heuristic evaluation or usability inspections allow experts to evaluate a design based on a set of principles or guidelines (i.e., heuristics). These are fast and inexpensive methods that can be used to identify and address design issues.

These expert guidelines date back to the early days of software design (e.g., Smith and Mosier [2]) and have over the past decades improved how we develop software and interactive applications. In the established areas of UX, heuristic evaluation or inspection methods [3, 4] are commonly used as evaluation tools during the project design and implementation phases. These are not meant to replace user testing, but rather complement the set of evaluation tools. While it has become more common to conduct user tests with gamified applications (just as games user researchers have done in the video game industry), the domain is still lacking robust methodologies for evaluating gameful designs.

The benefit of using a gamification inspection method is that it allows rapid and early evaluation of a gameful design. While several studies have investigated the effectiveness of gameful applications by studying their users [5], user tests are conducted after a prototype has already been implemented. Although concerns have been voiced that heuristic evaluation can be influenced by subjective interpretations [6], it remains a valuable tool for practitioners, who operate under tighter time constraints than researchers. Heuristic evaluation affords researchers a finer focus in the user tests that are usually done subsequently to this initial validation, since the most basic issues will have already been discovered at that point.

While UX tests focus on identifying issues related to usability, ergonomics, cognitive load, and affective experiences, gamification is concerned with understanding and fostering the user's motivation to use a product, system, or service. Thus, gamification methods rely on motivational psychology research, such as self-determination theory (SDT) [7–10], to understand human motivation. Our heuristics were informed by this theoretical framework.

Several gameful design frameworks and methods have been suggested [11, 12] with prescriptive guidelines for augmenting an application with motivational affordances (note that we refer to gamification and gameful design interchangeably because both frame the same set of phenomena from different points of view [1]). Motivational affordances are properties added to an object, which allow its users to experience the satisfaction of their psychological needs [13, 14]. In gameful design, motivational affordances are used to facilitate intrinsic and extrinsic motivation. Thus, motivational affordances supporting a user's feelings of competence, autonomy, and relatedness can facilitate intrinsic motivation, whereas external incentives or rewards facilitate extrinsic motivation.

Our work contributes to the human-computer interaction (HCI) and gamification communities by presenting a new set of guidelines for heuristic evaluation of gameful design in interactive systems. We began our research by reviewing several gameful design frameworks and methods to identify which dimensions of motivational affordances were common among them. Next, we created a set of heuristics focused on each of the identified dimensions. The resulting set of heuristics provides a new way of evaluating gameful user experiences. It is the first inspection tool focused specifically on evaluating gameful design through the lens of intrinsic and extrinsic motivational affordances. The aim of our inspection tool is to enable any UX expert to conduct a heuristic evaluation of a gameful application more easily, even if they have no background expertise in gameful design or motivational psychology.

To evaluate the proposed heuristics, we conducted a study with five UX or HCI professionals who evaluated two online gameful applications. Three participants used our gameful design heuristics, while the remaining two used a two-page description of gamification and motivational affordances. Results showed that usage of our heuristics led to more motivational issues being identified in the evaluated applications, as well as a broader range of identified issues, comprising a larger number of different dimensions.

2   Related Work and Model Development

2.1   Heuristic Evaluation for Games

In usability engineering, heuristics are broad usability guidelines that have been used to design and evaluate interactive systems [15]. Heuristic evaluation is the use of these principles by experts in a usability inspection process to identify usability problems in an existing design as part of an iterative design process [3, 4]. These inspections are usually done early in the design process to identify application errors before scheduling user tests.

Several authors have suggested heuristic evaluation models for games. These models vary both in their goals and in the dimensions they address: while some are more general, aimed at evaluating any game genre or type, others are more focused for example on networked or mobile games. Some of the most relevant heuristic evaluation models for game design are shown in Table 1.

Some heuristics for evaluating games or playability may also be applied to gameful applications. Some of the dimensions addressed by most game design heuristics are of relevance to gameful design, such as goals, challenge, feedback, and social interaction. However, heuristics for games include several dimensions that are not applicable to most gameful applications, such as control and concentration.

Additionally, some of the game heuristics cover issues that can be addressed in gameful applications using general UX principles, such as screen layout or navigation. These heuristics might be necessary when evaluating games because game design often uses its own user interface principles, which can be different from traditional application interfaces. However, most gameful applications follow current design standards for user interfaces; thus, general UX evaluation methods can be easily applied to gameful applications to address issues such as usability or ergonomics.

Table 1. Existing heuristic evaluation models for games.

Model	Description
Heuristic Evaluation for Playability (HEP) [31]	A set of heuristics for playability comprising four categories: gameplay, game story, game mechanics, and game usability
Games Usability Heuristics (PLAY) [32]	A set of 48 principles aimed at evaluating action-adventure, RTS, and FPS games. The heuristics are organized according to three categories: gameplay, coolness/entertainment/humor/emotional immersion, and usability & game mechanics
Game Approachability Principles (GAP) [33]	A set of guidelines to create better tutorials or experiences for new players
Playability Heuristics for Mobile Games [34]	A set of heuristics for mobile games comprising three categories: game usability, mobility, and gameplay
Networked Game Heuristics (NGH) [35]	A set of heuristics that consider specific issues related to group play over a network
Heuristics for Social Games [36]	A set of heuristics created from a critical review of prior video game evaluation heuristics
GameFlow [37, 38]	A comprehensive heuristic set designed as a tool to evaluate player enjoyment in eight dimensions: concentration, challenge, player skills, control, clear goals, feedback, immersion, and social interaction

Game design heuristics do not cover the full range of common motivational affordances used in gamification. For example, meaning, rewards, and scarcity are dimensions of motivational affordances often used in gameful design that are not covered by existing game heuristics. This makes it difficult to use game design heuristics to evaluate gameful applications. In order to do so, an evaluator would have to decide first which dimensions from the game heuristics should be used and which should not; next, they would also have to be concerned with motivational issues that are not currently covered by game heuristics. Consequently, we conclude that we need an inspection method better suited to assess gameful applications.

Before creating our set of gameful design heuristics, we reviewed the abovementioned game heuristics and considered the possibility of extending the existing models rather than proposing a new one. However, we encountered the same issues mentioned above: we would have to separate which heuristics from the existing models are applicable to gameful design and which are not. The resulting model would be confusing and difficult to apply. Therefore, we decided to create a new set of gameful design heuristics by analyzing existing gameful design methods rather than analyzing and extending existing game design heuristics.

2.2   Heuristic Evaluation for Playful Design

The Playful Experiences (PLEX) Framework [16, 17] provides an understanding of pleasurable user experience, which can be applied to both games and gameful applications. It classifies playful experiences according to 22 categories (see Table 2).

Table 2. The 22 categories of the PLEX framework.

Experience	Description
Captivation	Forgetting one's surroundings
Challenge	Testing abilities in a demanding task
Competition	Contest with oneself or an opponent
Completion	Finishing a major task, closure
Control	Dominating, commanding, regulating
Cruelty	Causing mental or physical pain
Discovery	Finding something new or unknown
Eroticism	A sexually arousing experience
Exploration	Investigating an object or situation
Expression	Manifesting oneself creatively
Fantasy	An imagined experience
Fellowship	Friendship, communality, or intimacy
Humor	Fun, joy, amusement, jokes, gags
Nurture	Taking care of oneself or others
Relaxation	Relief from bodily or mental work
Sensation	Excitement by stimulating senses
Simulation	An imitation of everyday life
Submission	Being part of a larger structure
Subversion	Breaking social rules and norms
Suffering	Experience of loss, frustration, anger
Sympathy	Sharing emotional feelings
Thrill	Excitement derived from risk, danger

The PLEX framework can be used as a tool for heuristic evaluation of gameful interactive systems, similar to the gameful design heuristics we are presenting. Nevertheless, PLEX is focused on classifying the types of experiences that the system can afford, rather than the motivational potential of these experiences. Therefore, the PLEX framework and the gameful design heuristics are two complementary tools, which can each provide insights into different characteristics of interactive systems that work together to afford an enjoyable user experience.

2.3 Review of Gameful Design Methods

To the best of our knowledge, no extant set of heuristics is available for evaluating motivation in gameful design. Some of the existing gameful design methods, namely Octalysis [18], HEXAD [19], and Lens of Intrinsic Skill Atoms [11], suggest procedures to evaluate an existing system. Nevertheless, these procedures only provide a starting point for the design process. They are less suited for being used as an evaluation tool by a quality control team because they lack a concise set of heuristics with brief descriptors which could be quickly checked by a UX practitioner. Moreover, the lack of a succinct rubric implies that an evaluator would need to study the methods intensively before being able to conduct an evaluation. Therefore, presently, there is no evaluation method for gameful applications that can be easily learned by UX professionals who are not familiar with gameful design. Our research fills this gap.

Several gameful design frameworks and methods are currently available (see [11, 12, 20] for comprehensive reviews). Therefore, we decided to review these existing methods to extract the different dimensions of motivational affordances that need to be considered in gameful design. Since the reviewed methods synthesize the current set of best practices in gameful design, we considered that they could provide an adequate starting point to identify motivational dimensions of concern. However, only a few of the reviewed methods feature a classification of motivational affordances in different dimensions, which we could use as a theoretical background to devise our heuristics. This was unfortunate, since our goal was to use these dimensions of motivational affordances as the starting point for the development of our framework. Gameful design methods that do not provide a classification of dimensions of motivational affordances would not be helpful in creating our gameful design heuristics. Therefore, we expanded the scope of our analysis to include methods that presented some sort of classification of motivational affordances. Table 3 lists the frameworks and methods we considered, as well as the rationale for their inclusion or otherwise in our analysis.

After reviewing the frameworks and methods and selecting six of them for further analysis (see Table 3), we conducted a comparison of the motivational dimensions in each model to map the similarities between them, using the following procedure:

1. The first framework was added as the first column of a table, with each one of its suggested motivational dimensions as separate rows. We chose the Octalysis framework as the first one because it comprised the highest number of dimensions (eight), which facilitated subsequent procedures/steps, but we could have chosen any of the frameworks as a starting point.
2. Next, we added each one of the remaining models as additional columns into the table. For each added model, we compared each one of its suggested dimensions with the rows that already existed in the table. When the new dimension to be added corresponded to one of the dimensions already in the table, we added it to the relevant existing row. Otherwise, we added a new row to the table creating a new dimension. In some cases, the addition of a new dimension also prompted the subdivision of an existing row. For example, the competence dimension was split into challenge/competence and completeness/mastery.

3. After adding all the models to the table, we observed the characteristics of the dimensions named in the rows and created for each of the latter a unique label, comprising the meaning of all the dimensions it encompassed.

Table 3. A summary of the gameful design frameworks & methods considered in our research.

Framework or Method	References	Included in the analysis?	Rationale
Gamification by Design	Zichermann and Cunningham [39]	No	Does not provide a classification of dimensions of motivational affordances
Gamification Framework	Francisco-Aparicio et al. [40]	No	Does not provide a classification of dimensions of motivational affordances
Gamification Model Canvas	Jiménez [41]	No	Does not provide a classification of dimensions of motivational affordances
Gamify	Burke [42]	No	Does not provide a classification of dimensions of motivational affordances
User Types HEXAD	Tondello et al. [19]	Yes	Provides a classification with six user types that are further used to classify sets of game elements for each type
The Kaleidoscope of Effective Gamification (KEG)	Kappen and Nacke [43]	Yes	Provides a classification with several layers of motivational affordances that can be used to design or evaluate gameful systems
Motivational Design Lenses (MDL)	Deterding [11]	Yes	Provides a classification of motivational design lenses that can be used to evaluate gameful systems
Loyalty 3.0	Paharia [44]	No	Does not provide a classification of dimensions of motivational affordances
The RECIPE for Meaningful Gamification	Nicholson [45]	Yes	Provides six different motivational dimensions for gameful design
Octalysis Framework	Chou [18]	Yes	Provides a classification with eight dimensions of motivation that can be used to design or evaluate gameful systems
Six Steps to Success	Werbach and Hunter [46]	No	Does not provide a classification of dimensions of motivational affordances
Super Better	McGonigal [47]	Yes	Proposes a gameful design method based on seven steps, which can be mapped as motivational dimensions

The resulting model consists of twelve common dimensions of motivational affordances (see Table 4). The similarity analysis between dimensions of different models was conceptual, meaning that we studied the description of each dimension as presented by their original authors and decided whether they represented the same core construct as any of the dimensions already present in the table. Similarly, we derived the labels for each one of the twelve resulting dimensions (first column of Table 4) by identifying the core concepts of each dimension. In the resulting classification, we noted that these dimensions were strongly based on: (1) the theories of intrinsic and extrinsic motivation (SDT; [7–9]), (2) behavioural economics [21], and (3) the practical experience of the authors of the analyzed frameworks. The entire initial analysis was conducted by one of the researchers; next, three other researchers (co-authors) also analyzed the resulting table. We then conducted an iterative loop of feedback and editing until none of the researchers had additional suggestions to improve the final model.

Table 4. Dimensions of motivational affordances from the reviewed gameful design methods.

Dimension	Octalysis [18]	HEXAD [19]	KEG [43]	MDL [11]	RECIPE [45]	Super Better [47]
Purpose and Meaning	Epic Meaning & Calling	Philanthropist			Information; Reflection	Epic win
Challenge and Competence	Development & Accomplishment	Achiever	Competence; Challenge	Challenge lenses; Intrinsic rewards	Engagement	Challenge; Bad guys
Completeness and Mastery	Development & Accomplishment	Achiever	Competence; Achievements	Goal and Action lenses; Intrinsic rewards		Complete quests
Autonomy and Creativity	Creativity & Feedback	Free Spirit	Autonomy	Object lenses; Intrinsic rewards	Play; Choice	
Relatedness	Social Influence & Relatedness	Socialiser	Relatedness	Intrinsic rewards	Engagement	Recruit allies
Immersion			Perceived Fun		Exposition	Secret identity
Ownership and Rewards	Ownership & Possession	Player	Extrinsic motivation	Intrinsic rewards		Power-ups
Unpredictability	Unpredictability & Curiosity	Free Spirit		Varied challenge; Varied feedback; Secrets	Play	
Scarcity	Scarcity & Impatience					
Loss avoidance	Loss & Avoidance					
Feedback	Creativity & Feedback			Feedback lenses		
Change and Disruption		Disruptor				

3   Gameful Design Heuristics

Our set of heuristics enables experts to identify gaps in a gameful system's design. This is achieved by identifying missing affordances from each of the dimensions.

Prior to creating the heuristics, we reviewed the research on motivation [7, 8] to help categorize the twelve dimensions into *intrinsic*, *extrinsic*, and *context-dependent* motivational categories. This is a common practice in gameful design and many of the reviewed methods also employ a similar classification. Although it is a simplification of the underlying theory, this simple categorization helps designers and evaluators better understand the guidelines and focus their attention on specific motivational techniques. We chose SDT as the theoretical background for this classification because it is the motivational theory most frequently employed in gameful design methodologies [11, 12].

We used the following criteria to split our heuristics into categories:

- *Intrinsic motivation* includes affordances related to the three intrinsic needs introduced by SDT [7, 8] (competence, autonomy, and relatedness), as well as 'purpose' and 'meaning' as facilitators of internalization [22–24] and 'immersion', as suggested by Ryan and Rigby [9, 25] and Malone [26].
- *Extrinsic motivation* includes affordances that provide an outcome or value separated from the activity itself as suggested by SDT [8] and Chou [18]: ownership and rewards, scarcity, and loss avoidance.
- *Context-dependent* motivation includes the feedback, unpredictability, and disruption affordances, which can afford either intrinsic or extrinsic motivation depending on contextual factors. For example, the application can provide feedback to the user regarding either intrinsically or extrinsically motivated tasks; therefore, feedback might afford intrinsic or extrinsic motivation according to the type of task with which it is associated.

We constructed the heuristics based on an examination of the literature cited in Table 4, by writing adequate guidelines for each of the twelve identified dimensions. Following the literature review, we created these guidelines by studying the descriptions of each dimension in the original models, identifying the main aspects of each dimension, and writing concise descriptions of each aspect to assist expert evaluation. We employed the following procedure:

1. For each one of the twelve motivational dimensions, we first studied the underlying concepts and wrote a short description of the dimension itself, aimed at guiding expert evaluators' understanding of each dimension.
2. Next, for each dimension, we identified the main aspects of concern, meaning the aspects that should be considered by designers when envisioning a gameful system, as suggested by the reviewed frameworks or methods. We argue that these aspects of concern, when designing a system, should also be the main points of evaluation.
3. For each aspect of concern, we then wrote a concise description aimed at guiding experts in evaluating whether the aspect being scrutinized was considered in the evaluated system's design.

Tables 5, 6 and 7 present the final set of 28 heuristics organized within the 12 dimensions, following the initial analysis, framing, and iterative feedback mentioned above, and which have been presented previously in a work-in-progress [27].

Table 5. Intrinsic motivation heuristics.

Intrinsic motivation heuristics
Purpose and Meaning: Affordances aimed at helping users identify a meaningful goal that will be achieved through the system and can benefit the users themselves or other people
I1. *Meaning*: The system clearly helps users identify a meaningful contribution (to themselves or to others)
I2. *Information and Reflection*: The system provides information and opportunities for reflection towards self-improvement
Challenge and Competence: Affordances aimed at helping users satisfy their intrinsic need of competence through accomplishing difficult challenges or goals
I3. *Increasing Challenge*: The system offers challenges that grow with the user's skill
I4. *Onboarding*: The system offers initial challenges for newcomers that help them learn how it works
I5. *Self-challenge*: The system helps users discover or create new challenges to test themselves
Completeness and Mastery: Affordances aimed at helping users satisfy their intrinsic need of competence by completing series of tasks or collecting virtual achievements
I6. *Progressive Goals*: The system always presents the next actions users can take as tasks of immediately doable size
I7. *Achievement*: The system lets users keeps track of their achievements or advancements
Autonomy and Creativity: Affordances aimed at helping users satisfy their intrinsic need of autonomy by offering meaningful choices and opportunities for self-expression
I8. *Choice*: The system provides users with choices on what to do or how to do something, which are interesting but also limited in scope according to each user's capacity
I9. *Self-expression*: The system lets users express themselves or create new content
I10. *Freedom*: The system lets users experiment with new or different paths without fear or serious consequences
Relatedness: Affordances aimed at helping users satisfy their intrinsic need for relatedness through social interaction, usually with other users
I11. *Social Interaction*: The system lets users connect and interact socially
I12. *Social Cooperation*: The system offers the opportunity of users working together towards achieving common goals
I13. *Social Competition*: The system lets users compare themselves with others or challenge other users
I14. *Fairness*: The system offers similar opportunities of success and progression for everyone and means for newcomers to feel motivated even when comparing themselves with veterans
Immersion: Affordances aimed at immersing users in the system in order to improve their aesthetic experience [48], usually by means of a theme, narrative, or story, which can be real or fictional
I15. *Narrative*: The system offers users a meaningful narrative or story with which they can relate to
I16. *Perceived Fun*: The system affords users the possibility of interacting with and being part of the story (easy fun; [49])

Table 6. Extrinsic motivation heuristics.

Extrinsic motivation heuristics
Ownership and Rewards: Affordances aimed at motivating users through extrinsic rewards or possession of real or virtual goods. Ownership is different from Competence when acquiring goods is perceived by the user as the reason for interacting with the system, instead of feeling competent
E1. *Ownership*: The system lets users own virtual goods or build an individual profile over time, which can be developed by continued use of the system and to which users can relate
E2. *Rewards*: The system offers incentive rewards for interaction and continued use, which are valuable to users and proportional to the amount of effort invested
E3. *Virtual Economy*: The system lets users exchange the result of their efforts with in-system or external rewards
Scarcity: Affordances aimed at motivating users through feelings of status or exclusivity by means of acquisition of difficult or rare rewards, goods, or achievements.
E4. *Scarcity*: The system offers interesting features or rewards that are rare or difficult to obtain
Loss Avoidance: Affordances aimed at leading users to act with urgency, by creating situations in which they could lose acquired or potential rewards, goods, or achievements if they do not act immediately
E5. *Loss Avoidance*: The system creates urgency through possible losses unless users act immediately

Table 7. Context-dependent heuristics.

Context-dependent heuristics
Feedback: Affordances aimed at informing users of their progress and the next available actions or challenges
C1. *Clear and Immediate Feedback*: The systems always inform users immediately of any changes or accomplishments in an easy and graspable way
C2. *Actionable Feedback*: The system always informs users about the next available actions and improvements
C3. *Graspable Progress*: Feedback always tells users where they stand and what is the path ahead for progression
Unpredictability: Affordances aimed at surprising users with variable tasks, challenges, feedback, or rewards
C4. *Varied Challenges*: The system offers unexpected variability in the challenges or tasks presented to the user
C5. *Varied Rewards*: The system offers unexpected variability in the rewards that are offered to the user
Change and Disruption: Affordances aimed at engaging users with disruptive tendencies [19] by allowing them to help improve the system, in a positive rather than destructive way
C6. *Innovation*: The system lets users contribute ideas, content, plugins, or modifications aimed at improving, enhancing, or extending the system itself
C7. *Disruption Control*: The system is protected against cheating, hacking, or other forms of manipulation from users

Additionally, we have extended the gameful design heuristics by writing a set of questions for each heuristic. These questions inquire about common ways of implementing each guideline, helping the evaluators assess whether the guideline is implemented in the system at all. We do not include the complete set of questions here because of space constraints, but we provide them in our website[1].

3.1 Using the Gameful Design Heuristics

Similar to previous heuristic UX evaluation methods, gamification heuristics should be used by experts to identify gaps in a gameful system's design. Experts should consider each guideline to evaluate whether it is adequately implemented into the design. Prior studies have shown that evaluations conducted by many evaluators are more effective in finding issues than those conducted by an individual evaluator [3, 28, 29]. Thus, we recommend the evaluation to be conducted by two or more examiners.

When applying the heuristics, the evaluators should first familiarize themselves with the application to be analyzed and its main features. Then, for each heuristic, they should read the general guideline and observe the application, identifying and noting what the application does to implement this guideline. Next, they should read the questions associated with the heuristic and answer them to identify possible gaps in the application's design. The evaluation is focused on observing the presence or absence of the motivational affordances and, if the evaluator has enough expertise, in evaluating their quality. However, it does not aim to observe the actual user experience, which is highly dependent on the users themselves in addition to the system. Therefore, this method cannot evaluate the user experience; its goal is to evaluate the system's *potential* to afford a gameful, engaging experience. As we have stated before, the heuristic evaluation should be subsequently validated by user studies to establish whether the observed potential translates into actual gameful experiences.

It is important to note that the questions associated with each heuristic act as guidelines to facilitate the evaluation process. They are not intended to represent every aspect related to the heuristic. Therefore, it is important that the evaluator also thinks beyond the suggested questions and considers other issues that might be present in the application regarding each heuristic.

After evaluating all the dimensions, a count of the number of issues identified in each dimension can help identify which motivational issues (from the heuristics) require more attention in improving the system's potential to engage users.

3.2 Turning the Evaluation Results into Actionable Design

Since the gameful design heuristics are an evaluation method, they do not provide the means to turn the identified issues into actionable design ideas to improve the application's design. Although the heuristics identify what dimensions of motivational affordances are implemented in, or excluded from the system, they do not provide any information about the need (or otherwise) to implementing the missing dimensions.

[1] http://gamefuldesign.hcigames.com/.

Depending on the goals of the gameful software being developed, including motivational affordances for all dimensions might be either necessary or unimportant. Therefore, we suggest that the identified design gaps should be considered within an iterative gameful design method, which can then provide the tools to assess the need for including new motivational affordances into the system to address the gaps. The methods used to inform the development of the heuristics (see Table 4) are adequate for this goal because they make it easy to map the dimensions where gaps are identified to the design element categories suggested by these gameful design methods.

4   Evaluation

We conducted a summative study with five UX or HCI experts to evaluate the gameful design heuristics. We asked participants to evaluate two online gameful applications: Habitica[2] and Termling[3]. Data were collected between August and December 2016.

Three participants (P1, P2, P3) conducted the evaluation using the heuristics and the remaining two (P4, P5) without it, enabling us to compare how many motivational design issues were found by experts with and without the heuristics. Furthermore, three participants (P1, P3, P4) had expertise in gamification or games, whereas two (P2, P5) were knowledgeable in UX or HCI, but did not have a specific background in gamification. This enabled us to assess if prior gamification expertise would influence the evaluators' ability to identify motivational design issues.

4.1   Participants

We initially invited 18 experts in UX, HCI, or gamification to participate in the study. Potential participants were selected from the authors' acquaintances and from previous project collaborators. The criterion was that potential participants should have an expertise either in gamification or games (including design practice or research experience) or in using other UX or HCI methods to evaluate interactive digital applications. Potential participants were contacted by email or in person. No compensation was provided for participation.

From the 18 invited participants, 10 initially agreed to participate and were sent the instructions; of these only five participants completed the procedures (likely because of scheduling difficulties and the lack of compensation). Of these five, two participants completed the evaluation of Habitica only; however, we decided to include their feedback in the study anyway. This meant that we collected five evaluations for Habitica, but only three for Termling. Table 8 summarizes the demographics of the participants.

---

[2] http://habitica.com/, last accessed December 2016.

[3] http://www.termling.com/, last accessed December 2016.

4.2   Procedure

Initially, participants read and signed a consent form and filled out a short demographic information form (see Table 8). Next, the instructions to evaluate the two applications were sent out. Since both applications were free and available online, participants were instructed to create a free account to test them. We instructed participants P1 and P2 to carry out the evaluation without the gameful design heuristics and participants P3, P4, and P5 to use the heuristics. Assignment to experimental conditions was not random because we needed to ensure that we had participants with and without gamification expertise in both conditions (with or without the heuristics).

Table 8. Participant demographics.

#	Gender	Role	Gamification expertise?	Has studied gamification before?
P1	Male	Graduate Student (HCI)	No	Yes (4 months)
P2	Male	Creative Director	Yes	Yes (3 years)
P3	Female	Professor (HCI)	Yes	Yes
P4	Male	Creative Lead	Yes	Yes
P5	Female	Graphic Designer	No	No

The instructions for P1 and P2 contained a one-page summarized introduction about gamification and motivation, followed by instructions requesting them to reflect on the applications' design and motivational affordances, try to understand how they afford intrinsic and extrinsic motivation, and then list any issue they identified related to the motivational affordances (or lack of the same).

Participants P3, P4, and P5 received information that contained the same introduction about gamification and motivation, followed by an introduction to the gameful design heuristics, and instructions that asked them to reflect on the applications and identify motivational issues using the gameful design heuristics. Participants were given a complete copy of the gameful design heuristics to guide them during the evaluation, including the full list of heuristics with all the accompanying questions to guide the evaluation (see Sect. 3). The heuristics were formatted as a fillable form, which offered an additional column where participants could take notes about the issues observed in the applications. After receiving the instructions, participants could conduct their evaluations at their own pace and discretion; they were not supervised by the researchers. After completing the evaluation, participants emailed the forms back to the researchers.

4.3   Results

Table 9 shows the number of issues found in the two evaluated applications by the participants. Overall, participants who used the gameful design heuristics identified more issues than those who did not use any heuristics.

Table 9. Number of issues found by participants.

Participant	Habitica					Termling				
	P1	P2	P3	P4	P5	P1	P2	P3	P4	P5
Used heuristics?	No	No	Yes	Yes	Yes	No	No	-	Yes	-
I1. Meaning	1	3	0	0	0	0	2	-	1	-
I2. Information and Reflection	0	0	1	0	1	0	0	-	1	-
I3. Increasing Challenge	0	0	1	0	0	0	0	-	1	-
I4. Onboarding	2	1	1	1	1	1	3	-	1	-
I5. Self-challenge	0	0	0	0	0	0	0	-	0	-
I6. Progressive Goals	0	0	0	1	1	0	0	-	1	-
I7. Achievement	0	0	1	0	0	0	1	-	1	-
I8. Choice	0	0	2	1	0	0	0	-	2	-
I9. Self-expression	1	0	0	0	0	1	2	-	0	-
I10. Freedom	0	0	0	0	1	0	0	-	1	-
I11. Social Interaction	0	1	2	0	0	0	1	-	0	-
I12. Social Cooperation	0	0	1	0	1	0	0	-	0	-
I13. Social Competition	0	0	0	1	0	0	0	-	2	-
I14. Fairness	0	0	0	1	0	0	0	-	0	-
I15. Narrative	0	0	1	2	0	0	0	-	1	-
I16. Perceived Fun	1	0	0	1	0	0	0	-	1	-
E1. Ownership	0	0	0	0	0	0	0	-	1	-
E2. Rewards	1	0	0	0	0	1	0	-	1	-
E3. Virtual Economy	1	0	0	2	0	0	0	-	2	-
E4. Scarcity	0	0	0	0	0	0	1	-	0	-
E5. Loss Avoidance	0	1	0	1	1	0	1	-	1	-
C1. Clear & Immediate Feedback	0	0	0	0	1	1	0	-	1	-
C2. Actionable Feedback	0	0	0	0	0	0	0	-	1	-
C3. Graspable Progress	0	0	0	1	1	0	0	-	1	-
C4. Varied Challenges	0	1	1	1	1	0	1	-	0	-
C5. Varied Rewards	0	1	1	1	1	0	0	-	1	-
C6. Innovation	0	0	0	0	0	0	0	-	2	-
C7. Disruption Control	0	0	0	2	0	0	0	-	0	-
Total motivational issues	7	8	12	16	10	4	12	-	24	-
Other issues (usability, bugs)	3	2	-	-	-	4	2	-	-	-

The number of issues identified by the participant who had no prior gamification expertise and used the heuristics (P5) was just slightly higher than the participants who did not use the heuristics (P1 and P2), whether they had gamification expertise or not. However, it is noteworthy that the heuristics helped P5 identify issues in more dimensions than did P1 and P2: while P5 identified issues in 10 different dimensions for Habitica, P1's and P2's issues were concentrated in only six dimensions.

Moreover, congruent to our intentions, the heuristics helped evaluators focus their analyses on the motivational affordances instead of other usability issues or bugs.

This is demonstrated by the fact that P1 and P2 both reported some issues that were not related to the motivational affordances at all (e.g., usability issues or bugs), whereas P3, P4, and P5 only reported motivational issues.

Furthermore, a qualitative comparison of participants' responses shows that when they used the heuristics, their comments were generally more focused on the motivational aspects, whereas the comments from participants who did not use the heuristics were more general. For example, regarding Habitica's onboarding, P1, P2, and P3 mostly recognized the fact that some information or tutorial material is missing or hidden. However, they do not comment on how this would affect the user's motivation. On the other hand, P4 and P5 could point out that, although a set of instructions existed, it did not motivate the user because it was not challenging or fun. Thus, it seems that the heuristics are useful in focusing the evaluator's attention into the motivational issues of the application.

Additionally, the participants who had prior gamification expertise and used the heuristics could identify approximately twice as many motivational issues as the participant who also had gamification expertise but did not use the heuristics. In comparison, P3 found 12 and P4 found 16 motivational issues in Habitica, whereas P2 found only eight. In Termling, P4 found 24 motivational issues while P2 found only 12.

We can also observe that the motivational dimensions where participants classified the issues sometimes differ. However, this is not a characteristic specific to our tool, but it is a known fact of heuristic evaluation in general that a single evaluator usually does not notice all the existing issues. This is why it is recommended that a heuristic evaluation should be conducted by a number of experts instead of only one [3, 28, 29]. This way, by combining all the issues identified by the different experts, good coverage of the total issues existing in the system will be achieved.

In summary, the results provided the following evidence:

- A participant who had no prior gamification expertise, but used the gameful design heuristics, could find as many motivational issues as participants who did not use the heuristics (with or without prior gamification expertise), but in a broader range of motivational dimensions;
- Participants who had prior gamification expertise and used the gameful design heuristics could find twice as many motivational issues than participants who did not use the heuristics or did not have prior expertise;
- Using the gameful design heuristics helped participants focus their analyses on the motivational issues, avoiding any distraction with other types of problems.

5 Discussion

We have created a set of 28 gameful design heuristics for the evaluation and identification of design gaps in gameful software. Due to the lack of direct applicability of existing heuritics from game design, we deliberately decided to create a new set of heuritics specific to gameful design, based on motivational theories and gameful

design methods, rather than extending the existing heuristics for game design. By deriving our set of heuristics from common dimensions of motivational affordances employed by different gameful design methods, we have presented a novel and comprehensive approach that encompasses a broad range of motivational affordances. Furthermore, to enable expert evaluation, the heuristics are written in a concise form, together with supportive questions for reflection.

Our study with five UX and HCI experts provided empirical evidence that:

- gameful design heuristics can help UX evaluators who are not familiar with gamification to evaluate a gameful system at least as well as a gamification expert who does not use the heuristics; and
- gameful design heuristics can greatly improve the ability of gamification experts to perform a heuristic evaluation, leading them to find twice as many issues as they would find without the heuristics.

The implications of our findings are twofold. First, we provide evidence that evaluation of gameful applications without a support tool is subjective; therefore, even gamification experts might miss important issues. A probable reason for this is the complexity of gameful design and the number of motivational dimensions involved. Second, we demonstrate that usage of the gameful design heuristics can significantly improve the results of heuristic evaluations conducted both by gamification experts and non-experts. Considering that gameful design still suffers from difficulties in reproducing some of the successful results and that several studies have reported mixed results [5, 30], our work sheds light on one of the probable causes for this. Consequently, the gameful design heuristics represent an important instrument, which can be used to improve the chances of building effective gameful applications.

Nevertheless, the study was limited by the small sample size. Thus, although these initial results seem promising, future studies will be needed to support them. Additionally, even though the proposed method was meant to be generic enough to work in any heuristic evaluation of gameful applications, future studies will need to consider diverse usage scenarios to investigate if adaptations are needed for specific purposes.

6   Conclusion

Evaluation using heuristics is a way of identifying issues during various stages of software development, ranging from ideation, design, and prototyping to implementation and tests. While many heuristics exist in various fields such as usability and game design, we still lacked guidelines specific to gameful design due to the differences in types of solutions emergent from this domain. Therefore, our work addresses this gap and contributes to gameful design research and practice by identifying key motivational dimensions and presenting a novel evaluation tool specific for gameful systems. This gameful design heuristics provides a method of evaluating interactive systems in various stages of their development. The suggested method fulfills a need for UX evaluation tools specific to gameful design, which could help evaluators assess the potential UX of a gameful application in the early phases of the software project. The expert evaluation of the gameful design heuristics provided information that the

heuristics enabled experts to identify the presence of motivational affordances from several dimensions, as well as the absence of specific affordances from other dimensions. This is valuable information, which could help software developers and systems designers to incorporate the missing elements. We expect the gameful design heuristics to be of use to both researchers and practitioners who design and evaluate gameful software, whether in research studies or in industry applications.

References

1. Deterding, S., Dixon, D., Khaled, R., Nacke, L.E.: From game design elements to gamefulness: defining "Gamification." In: Proceedings of the 15th International Academic MindTrek Conference, Tampere, Finland, pp. 9–15. ACM (2011). https://doi.org/10.1145/2181037.2181040
2. Smith, S.L., Mosier, J.N.: Guidelines for Designing User Interface Software. Mitre Corporation, Bedford (1986)
3. Nielsen, J.: Finding usability problems through heuristic evaluation. In: Proceedings of the SIGCHI Conference on Human Factors in Computer Systems - CHI 1992, pp. 373–380 (1992). https://doi.org/10.1145/142750.142834
4. Nielsen, J.: Heuristic evaluation. In: Nielsen, J., Mack, R.L. (eds.) Usability Inspection Methods, pp. 25–62. Wiley, New York (1994)
5. Hamari, J., Koivisto, J., Sarsa, H.: Does gamification work? - A literature review of empirical studies on gamification. In: Proceedings of the Annual Hawaii International Conference on System Sciences, pp. 3025–3034 (2014). https://doi.org/10.1109/hicss.2014.377
6. White, G.R., Mirza-Babaei, P., McAllister, G., Good, J.: Weak inter-rater reliability in heuristic evaluation of video games. In: Proceedings of the CHI 2011 Extended Abstracts on Human Factors in Computing Systems - CHI EA 2011, pp. 1441–1446. ACM (2011). https://doi.org/10.1145/1979742.1979788
7. Ryan, R.M., Deci, E.L.: Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am. Psychol. **55**, 68–78 (2000). https://doi.org/10.1037/0003-066X.55.1.68
8. Ryan, R.M., Deci, E.L.: Intrinsic and extrinsic motivations: classic definitions and new directions. Contemp. Educ. Psychol. **25**, 54–67 (2000). https://doi.org/10.1006/ceps.1999.1020
9. Ryan, R.M., Rigby, C.S., Przybylski, A.: The motivational pull of video games: a self-determination theory approach. Motiv. Emot. **30**, 347–363 (2006). https://doi.org/10.1007/s11031-006-9051-8
10. Deci, E.L., Ryan, R.M.: Intrinsic Motivation and Self-Determination in Human Behavior. Plenum, New York and London (1985)
11. Deterding, S.: The lens of intrinsic skill atoms: a method for gameful design. Hum. Comput. Interact. **30**, 294–335 (2015). https://doi.org/10.1080/07370024.2014.993471

12. Mora, A., Riera, D., González, C., Arnedo-Moreno, J.: Gamification: a systematic review of design frameworks. J. Comput. High. Educ. **29**(3), 516–548 (2017). https://doi.org/10.1007/s12528-017-9150-4

13. Deterding, S.: Situated motivational affordances of game elements: a conceptual model. In: Gamification: Using Game Design Elements in Non-Gaming Contexts, a Workshop at CHI 2011. ACM (2011)

14. Zhang, P.: Motivational affordances: reasons for ICT design and use. Commun. ACM **51**, 145–147 (2008). https://doi.org/10.1145/1400214.1400244

15. Nielsen, J.: Enhancing the explanatory power of usability heuristics. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI 1994, pp. 152–158 (1994). https://doi.org/10.1145/191666.191729

16. Lucero, A., Holopainen, J., Ollila, E., Suomela, R., Karapanos, E.: The playful experiences (PLEX) framework as a guide for expert evaluation. In: Proceedings of the 6th International Conference on Designing Pleasurable Products and Interfaces - DPPI 2013, pp. 221–230. ACM (2013). https://doi.org/10.1145/2513506.2513530

17. Lucero, A., Karapanos, E., Arrasvuori, J., Korhonen, H.: Playful or gameful? Creating delightful user experiences. Interactions **21**, 34–39 (2014). https://doi.org/10.1145/2590973

18. Chou, Y.: Actionable Gamification - Beyond Points, Badges, and Leaderboards. Octalysis Media (2015)

19. Tondello, G.F., Wehbe, R.R., Diamond, L., Busch, M., Marczewski, A., Nacke, L.E.: The gamification user types hexad scale. In: Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play - CHI PLAY 2016, Austin, TX, USA , pp. 229–243. ACM (2016). https://doi.org/10.1145/2967934.2968082

20. Morschheuser, B., Werder, K., Hamari, J., Abe, J.: How to gamify? A method for designing gamification. In: Proceedings of the 50th Annual Hawaii International Conference on System Sciences (HICSS), Hawaii, USA. IEEE (2017). https://doi.org/10.24251/hicss.2017.155

21. Hamari, J., Huotari, K., Tolvanen, J.: Gamification and economics. In: Walz, S.P., Deterding, S. (eds.) The Gameful World: Approaches, Issues, Applications, pp. 139–161. The MIT Press, Cambridge (2015)

22. Deci, E.L., Eghrari, H., Patrick, B.C., Leone, D.R.: Facilitating internalization: the self-determination theory perspective. J. Pers. **62**, 119–142 (1994). https://doi.org/10.1111/j.1467-6494.1994.tb00797.x

23. Huta, V., Waterman, A.S.: Eudaimonia and its distinction from Hedonia: developing a classification and terminology for understanding conceptual and operational definitions. J. Happiness Stud. **15**, 1425–1456 (2014). https://doi.org/10.1007/s10902-013-9485-0

24. Peterson, C., Park, N., Seligman, M.E.P.: Orientations to happiness and life satisfaction: the full life versus the empty life. J. Happiness Stud. **6**, 25–41 (2005). https://doi.org/10.1007/s10902-004-1278-z

25. Rigby, S., Ryan, R.M.: Glued to Games: How Video Games Draw Us In and Hold Us Spellbound. Praeger, Santa Barbara (2011)

26. Malone, T.W.: Toward a Theory of Intrinsically Motivating Instruction. Cogn. Sci. **4**, 333–369 (1981)

27. Tondello, G.F., Kappen, D.L., Mekler, E.D., Ganaba, M., Nacke, L.E.: Heuristic evaluation for gameful design. In: Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play Extended Abstracts - CHI PLAY EA 2016, pp. 315–323. ACM (2016). https://doi.org/10.1145/2968120.2987729

28. Nielsen, J., Molich, R.: Heuristic evaluation of user interfaces. In: Proceedings of the SIGCHI Conference on Human Factors in Computer Systems - CHI 1990, pp. 249–256 (1990). https://doi.org/10.1145/97243.97281

29. Nielsen, J., Landauer, T.K.: A mathematical model of the finding of usability problems. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI 1993, pp. 206–213. ACM (1993). https://doi.org/10.1145/169059.169166
30. Seaborn, K., Fels, D.I.: Gamification in theory and action: a survey. Int. J. Hum. Comput. Stud. **74**, 14–31 (2014). https://doi.org/10.1016/j.ijhcs.2014.09.006
31. Desurvire, H., Caplan, M., Toth, J.A.: Using heuristics to evaluate the playability of games. In: CHI 2004 Extended Abstracts on Human Factors in Computing Systems, pp. 1509–1512 (2004). https://doi.org/10.1145/985921.986102
32. Desurvire, H., Wiberg, C.: Game usability heuristics (PLAY) for evaluating and designing better games: the next iteration. In: Ozok, A.A., Zaphiris, P. (eds.) OCSC 2009. LNCS, vol. 5621, pp. 557–566. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-02774-1_60
33. Desurvire, H., Wiberg, C.: User experience design for inexperienced gamers: GAP – Game Approachability Principles. In: Bernhaupt, R. (ed.) Evaluating User Experience in Games, pp. 131–148. Springer, London (2010). https://doi.org/10.1007/978-1-84882-963-3_1
34. Korhonen, H., Koivisto, E.M.I.: Playability heuristics for mobile multi-player games. In: Proceedings of the 2nd International Conference on Digital Interactive Media in Entertainment and Arts, pp. 28–35 (2007). https://doi.org/10.1145/1306813.1306828
35. Pinelle, D., Wong, N., Stach, T., Gutwin, C.: Usability heuristics for networked multiplayer games. In: Proceedings of the ACM 2009 International Conference on Supporting Group Work, pp. 169–178 (2009). https://doi.org/10.1145/1531674.1531700
36. Paavilainen, J.: Critical review on video game evaluation heuristics: social games perspective. In: Proceedings of the International Academic Conference on the Future of Game Design and Technology, pp. 56–65 (2010). https://doi.org/10.1145/1920778.1920787
37. Sweetser, P., Wyeth, P.: GameFlow: a model for evaluating player enjoyment in games. Comput. Entertain. **3**, 3 (2005). https://doi.org/10.1145/1077246.1077253
38. Sweetser, P., Johnson, D., Wyeth, P.: Revisiting the GameFlow model with detailed heuristics. J. Creat. Technol. **3** (2013). https://ojs.aut.ac.nz/journal-of-creative-technologies/index.php/JCT/article/view/16
39. Zichermann, G., Cunningham, C.: Gamification by Design: Implementing Game Mechanics in Web and Mobile Apps. O'Reilly, Sebastopol (2011)
40. Francisco-Aparicio, A., Gutiérrez-Vela, F.L., Isla-Montes, J.L., Sanchez, J.L.G.: Gamification: analysis and application. In: Penichet, V., Peñalver, A., Gallud, J. (eds.) New Trends in Interaction, Virtual Reality and Modeling, pp. 113–126. Springer, London (2013). https://doi.org/10.1007/978-1-4471-5445-7_9
41. Jiménez, S.: Gamification Model Canvas. http://www.gameonlab.com/canvas/
42. Burke, B.: Gamify: How Gamification Motivates People to Do Extraordinary Things. Bibliomotion, Brookline (2014)
43. Kappen, D.L., Nacke, L.E.: The kaleidoscope of effective gamification: deconstructing gamification in business applications. In: Proceedings of Gamification 2013, Stratford, ON, Canada, pp. 119–122. ACM (2013). https://doi.org/10.1145/2583008.2583029
44. Paharia, R.: Loyalty 3.0: How to Revolutionize Customer and Employee Engagement with Big Data and Gamification. McGraw-Hill, New York (2013)
45. Nicholson, S.: A RECIPE for meaningful gamification. In: Reiners, T., Wood, L.C. (eds.) Gamification in Education and Business, pp. 1–20. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-10208-5_1
46. Werbach, K., Hunter, D.: For the Win: How Game Thinking Can Revolutionize Your Business. Wharton Digital Press, Philadelphia (2012)
47. McGonigal, J.: SuperBetter: A Revolutionary Approach to Getting Stronger, Happier, Braver and More Resilient. Penguin Books, New York (2015)

48. Hunicke, R., LeBlanc, M., Zubek, R.: MDA: a formal approach to game design and game research. In: Workshop on Challenges in Game AI, pp. 1–4 (2004)
49. Lazzaro, N.: The four fun keys. In: Isbister, K., Schaffer, N. (eds.) Game Usability: Advancing the Player Experience, pp. 315–344. Elsevier, Burlington (2008)