HexArcade: Predicting Hexad User Types By Using Gameful Applications

Maximilian Altmeyer
maximilian.altmeyer@dfki.de
German Research Center for Artificial Intelligence (DFKI)
Saarland Informatics Campus, Germany

Gustavo F. Tondello
gustavo@tondello.com
HCI Games Group, Games Institute and School of Computer Science
University of Waterloo, ON, Canada

Antonio Krüger
krueger@dfki.de
German Research Center for Artificial Intelligence (DFKI)
Saarland Informatics Campus, Germany

Lennart E. Nacke
lennart.nacke@acm.org
HCI Games Group, Games Institute and Stratford School of Interaction Design and Business
University of Waterloo, ON, Canada

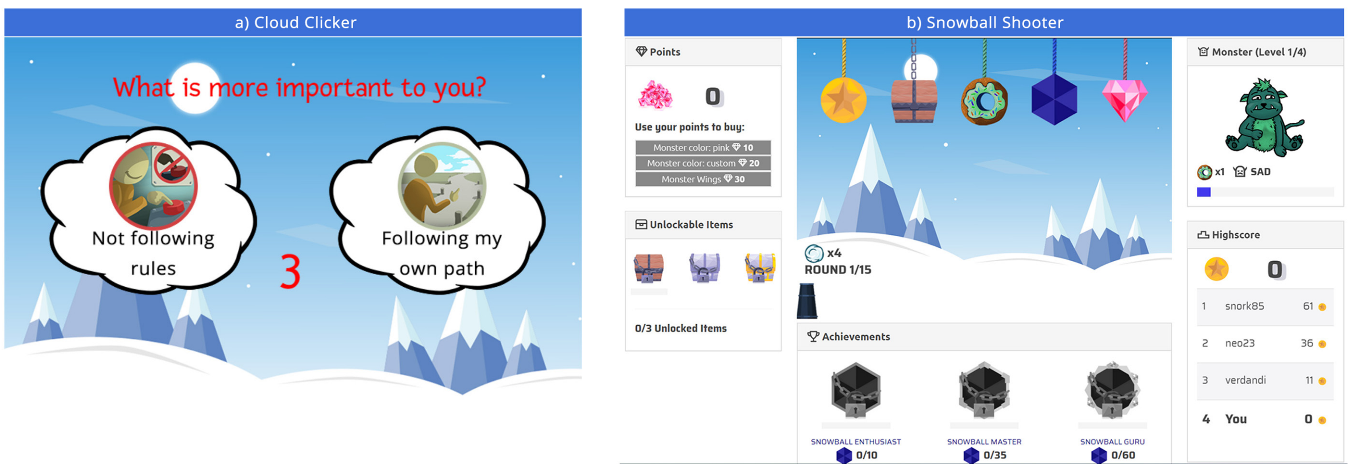

Figure 1: Gameful applications used in the user study. Cloud Clicker (a) asks users to decide which statement is more important to them. Snowball Shooter (b) provides several gamification elements, which can be interacted with.

ABSTRACT

Personalization is essential for gameful systems. Past research showed that the Hexad user types model is particularly suitable for personalizing user experiences. The validated Hexad user types questionnaire is an effective tool for scientific purposes. However, it is less suitable in practice for personalizing gameful applications, because filling out a questionnaire potentially affects a person's gameful experience and immersion within an interactive system negatively. Furthermore, studies investigating correlations between Hexad user types and preferences for gamification elements were survey-based (i.e., not based on user behaviour). In this paper, we improve upon both these aspects. In a user study (N=147), we show that gameful applications can be used to predict Hexad user types and that the interaction behaviour with gamification elements corresponds to a users' Hexad type. Ultimately, participants perceived our gameful applications as more enjoyable and immersive than filling out the Hexad questionnaire.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in HCI**.

KEYWORDS

Gamification; Hexad; Personalization; Prediction

CHI PLAY '20, November 2–4, 2020, Virtual Event, Canada

ACM ISBN 978-1-4503-8074-4/20/11.
https://doi.org/10.1145/3410404.3414232

ACM Reference Format:
Maximilian Altmeyer, Gustavo F. Tondello, Antonio Krüger, and Lennart E. Nacke. 2020. HexArcade: Predicting Hexad User Types By Using Gameful Applications. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play (CHI PLAY '20), November 2–4, 2020, Virtual Event, Canada.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3410404.3414232

1 INTRODUCTION

Gamification, *the use of game design elements in non-game contexts* [8], has been used in various contexts, including health, education, commerce or crowdsourcing [11]. Researchers have used gamification to motivate people to reach their goals, enhance their user experience or turn unpleasant tasks into fun ones [11, 37]. Often, this was done by adopting a "one-size-fits-all" approach (i.e., using a static set of gamification elements) [11, 16, 37]. However, interpersonal differences in the perception of gamification elements exist according to prior research [44]. These interpersonal differences could threaten static gamification approaches. For instance, researchers found that demographic factors such as age [5], gender [28], and personality traits [16] influence the perception of gameful elements.

The influence of demographic factors prompted gamification research to study and model user preferences in gamified systems. As a result, Marczewski [20, 44] proposed the Hexad user types model—a model that has been developed to understand and explain user preferences and behaviour in gameful systems [30, 43]. The model consists of six user types, which differ in the degree to which they need to experience autonomy, relatedness, competence, and purpose (which are core pillars of Self-Determination Theory [35]). Tondello et al. [44] developed a survey, refined it, and demonstrated its reliability and validity [42]. Subsequently, the Hexad user types model was used to investigate user preferences in gamified systems across different contexts, including physical activity [3], education [24], energy conservation [19], health [30], and others. These investigations highlighted the usefulness of the Hexad user types model for personalizing gameful systems. The correlations between the perception of gamification elements and Hexad user types [3, 44] enable dynamic adjustments to the elements of a gameful system.

However, determining the Hexad user type requires people to fill out a 24-item questionnaire. While this is appropriate and necessary in academic contexts to not break the psychometric properties of the Hexad model and ensure scientific rigour, it may be disadvantageous when using the Hexad model to tailor gamified systems dynamically to its users. Because gamified systems usually aim at providing an enjoyable and gameful user experience, requiring users to fill out a survey upfront may break immersion, lead to frustration and thus detrimentally affect the overall user experience of a gamified system, as suggested by previous work in the context of gamified surveys [12, 13, 17, 45]. Therefore, researching ways to tailor gamified systems without affecting the user experience of a gameful system negatively is quintessential.

We contribute to this tailoring of gamified systems by investigating whether gameful applications may be used to predict Hexad user types, such that gamified systems, which aim to personalize the set of gamification elements to their users, do not require users to fill out questionnaires. We analyze whether the interaction behaviour of users in gameful applications corresponds to their Hexad user type as a secondary aspect of our work. Specifically, this means whether users choose to interact with gamification elements that should—according to their Hexad user type—be particularly relevant to them. Previous studies investigating correlations between preferences for gamification elements and Hexad user types were survey-based. This survey methodology implied that participants did not interact with gameful applications but instead rated their perception subjectively. Thus, investigating whether we can replicate these correlations within gameful applications is a necessary next step.

We implemented two gameful applications to investigate the aspects above in an online study with 147 participants. In "Cloud Clicker" (see Figure 1a), participants were asked to select one statement that is more important to them out of two statements in total and then received gameful feedback. In the online study, we expected that the choice that participants made in Cloud Clicker corresponds to their Hexad user type. In "Snowball Shooter" (see Figure 1b), participants experienced gamification elements (Points, Achievements, Leaderboard, a Virtual Character, and Unlockables) and could decide with which gamification elements they wanted to interact. They could interact with the gamification elements by shooting snowballs at different virtual items, which in turn increased their progress in the corresponding gamification elements. We expected that the amount of interactions with each gamification element would reflect a participants' Hexad user type.

We found that Cloud Clicker can be used to predict Hexad user types in a gameful way. We also found that Snowball Shooter does not explain enough variance to predict Hexad user types reliably, but that participants indeed mostly interacted with gamification elements that correspond to their user type. Our results also show that both gameful applications are more enjoyable and immersive than completing the Hexad questionnaire.

Taking these findings into account, we contribute an effective, gameful way of assessing Hexad user types that is beneficial for enjoyment and immersion. In addition, our system shows that interaction behaviour in gameful applications indeed corresponds to a user's Hexad type.

2 RELATED WORK

We contribute to the field of personalized gamification. Consequently, we begin our discussion of the literature by motivating the field's importance. Next, we show that the Hexad model is frequently used, provides consistent findings across domains and was found to be the most favourable model to personalize gameful systems. Afterwards, we demonstrate that gamifiying surveys was successfully used to enhance the user experience while maintaining the validity of the collected data. We conclude by situating our work within previous literature.

2.1 Personalization in Gameful Systems

To account for the interpersonal differences in the perception of motivational affordances, researchers studied which factors moderate their perception. For instance, Jia et al. [16] investigated personality traits and to which extend they can be used to tailor gamification elements to users. Their results show that personality traits influence the perception of certain gamification elements, suggesting that tailoring gamified systems to the personality of users is beneficial. Orji et al. [29]'s research—which focused on persuasive strategies and how personality traits influence their perception—supports these results. They found correlations between persuasive strategies and personality traits.

In addition to research on personality, the motivational impact of demographic factors (e.g., age or gender) on game elements has been studied. Birk et al. [5] found that play motives change with increasing age when older adults focus more on enjoyment instead of performance (i.e., it seems like—with increasing age—the relevance of game performance decreases because enjoyment becomes paramount). Findings from Altmeyer et al. [2] support the decreased importance of game performance among older adults, showing the main reason to play is that older adults enjoy spending time with other people. In addition, the effect of gender has been studied by Oyibo et al. [31], who found that competition and rewards were perceived as more persuasive by male participants. Similarly, effects of age and gender have also been described by Orji et al. [28], who investigated differences in the perception of Cialdini's persuasion strategies in a study with 1,108 participants. The authors found that females are more responsive to most of the strategies.

2.2 Hexad User Types and Gameful Elements

Although the factors mentioned above are useful to personalize gameful systems, none of those were explicitly developed for this purpose. The Hexad user type model [20] bridges this gap, because it lays out ways to cluster users of gameful systems and provides recommendations to inform and tailor a system's design. It consists of six user types that differ in how much they are driven by their needs for autonomy, relatedness, competence and purpose (from Self-Determination Theory (SDT) [35]).

Philanthropists ("PH") are socially-minded, like to take responsibility, and share their knowledge with other users. Overall, they are driven by *purpose*.

Socialisers ("SO") are also socially-minded but are more interested in user interaction. Consequently, *relatedness* is their main motivation.

Free Spirits ("FS") strive for exploration and acting without external control, with *autonomy* being most important.

Achievers ("AC") are driven by overcoming obstacles and mastering difficult challenges. They are motivated by *competence*.

Players ("PL") —maybe the least aptly named of the user types—are focused on their benefits, and are motivated by the will to win and earn rewards. Hence, *extrinsic rewards* are most important for them.

Disruptors ("DI") like to test a system's boundaries and are driven by triggering *change*, either positive or negative.

Tondello et al. [44] developed a questionnaire to assess Hexad user types and thereby provide the foundation for further research. More recently, the authors [42] made slight adjustments to the initial questionnaire and showed its reliability and validity. Additionally, Tondello [41, chapter 3] proposed a method for personalized gameful design that can use the Hexad user types to aid in the selection of the preferred game design elements. As a result, the Hexad user types model has been used successfully in various domains, showing that the model can explain preferences for gamification elements. For instance, the Hexad model was used to investigate the perception of gamification elements in the physical activity domain by using a storyboards-based approach [3]. The authors found that a considerable majority of correlations between the Hexad user types and preferences for gamification elements established in [44] could be replicated. Similarly, Mora et al. [24] investigated whether using the Hexad model to personalize learning experiences helps to motivate and engage students better. They found that the personalized approach yielded higher engagement of the students, underlining the usefulness of the Hexad model for tailoring gameful systems.

Besides physical activity and education, Orji et al. [30] examined the Hexad model in the domain of unhealthy alcohol consumption. The authors found that the Hexad user type influences the perceived persuasiveness of strategies. Overall, the reported effects relate to the user type definitions and underline Hexad's applicability. Kotsopoulos et al. [19] investigated the perception of certain gamification elements and potential correlations to Hexad user types in the domain of energy efficiency applications at the workplace. The authors validated the Hexad user model in another domain by showing that the user types can be used to explain preferences towards gamification elements. They found similar correlations between gamification elements and user types as Tondello et al. [44]. Tondello et al. [43] proposed a conceptual framework for classifying game elements based on an exploratory factor analysis of people's preferences in a general context. Supporting previous results, expected correlations to the Hexad user types were shown.

Lastly, Hallifax et al. [10] analyzed which factors should be considered when personalizing gamified systems. They investigated which user models should be used, and compared the BrainHex [25] model, the Hexad model, and the Big-5 personality model [22]. They concluded that the Hexad model is the most suitable typology and should be used to tailor gamified systems as most of the results that were found by the authors are in line with the definitions of the Hexad user types. The authors state that this is likely because the Hexad model was specifically designed for gamification and most of its user types are based on SDT [35].

2.3 Gamified Questionnaires

We reviewed literature in the domain of gamified questionnaires because one of the goals of this research is to assess Hexad user types in a gameful way. In this context, Triantoro et al. [45] compared non-gamified surveys against gamified ones. They used the well-established Big-5 personality model [22] and transformed the Big-5 survey items from Likert scales into a binary, gameful decision for the gamified version of the questionnaire. In contrast, the non-gamified questionnaire used the traditional scale. The results revealed that the Big-5 responses that were assessed in a gameful way could be used to predict the actual Big-5 responses of the traditional survey. This prediction strategy is an essential precursor to our work because it shows that transforming a validated survey into a gamified version using a binary choice is possible without compromising the integrity of the collected data. The authors also found positive effects of gamification on enjoyment and attention, suggesting a better user experience and advantages for data quality.

Guiding the realization of gamified surveys, Harms et al. [14] suggests a four-step design process. In a follow-up work [12], this design process is extended, used and evaluated. Two HCI designers went through the aforementioned process transforming a traditional sports survey into its gamified counterpart. As a result, they came up with multiple sports-related minigames. The feedback of both designers supported the practical usefulness of the proposed process. Next, the gamified survey was compared to a non-gamified counterpart. Similar to the study from McCrae and John [22], it was found that participants considered the gamified version as more fun, provided more qualitative feedback and spent more time filling out the survey. Another work by Harms et al. [13] even shows that simplistic approaches to gamify questionnaires lead to positive effects. In this work, badges could be unlocked by completing the survey and results showed that participants preferred such a simple gamified survey over traditional ones, that the gamified design was found to be more attractive and stimulating, and that participants perceived the use of badges very positively. Lastly, Keusch et al.[17] conducted a literature review on comparative studies of gamified surveys. They support the aforementioned positive outcomes, outlining that gamified surveys were perceived as more interesting, more fun and that participants perceived them as more enjoyable and more comfortable to fill out. They also found that most of the survey questions were not negatively affected by gamification.

2.4 Summary

To sum up, related work has demonstrated that personalization is a vital topic in gamification research, emphasizing the relevance of our work. Previous research has identified numerous factors to consider when personalizing gameful systems, including age [2, 5], gender [28, 31] and personality traits [16, 29]. To better understand gamification preferences and how to tailor gameful systems, the Hexad user types model has been developed [20]. It is the only model specifically designed for the gamification domain and a validated questionnaire assessing Hexad user types reliably has been provided by Tondello et al. [42].

Although being comparably new, the Hexad model has been investigated and used across various domains and contexts and has been consistently shown to explain user preferences for gamification elements [3, 19, 24, 30, 43, 44]. These findings support the relevance of the Hexad model for personalizing gameful systems and consequently researching alternative ways of assessing Hexad user types. A study comparing the importance of different user typologies for personalization of gameful systems conducted by Hallifax et al. [10], highlights the importance and applicability of the Hexad model even further, showing that the Hexad model explains preferences of users most reasonably and thus should be used to effectively tailor gamified systems.

As a third fact, related work has shown that turning surveys into gameful experiences has positive effects on cognitive and affective reactions, including enjoyment, attention, involvement, and ease-of-use [12, 13, 17, 45]. It was shown that gamification mostly had no or little effect on the validity of the questionnaire [17, 45] and that turning Likert-scale questions into gameful binary-choice decisions provides comparable data that has favourable construct reliability [45]. This evidence compellingly supports our endeavour of assessing Hexad user types using gameful applications.

Taking all these findings into account, we extend previous work by investigating whether interaction behaviour with gamification elements in gameful applications corresponds to Hexad user types. Given that previous work considered self-reported preferences based on storyboards [3, 30] or textual descriptions [19, 44], allowing users to perceive, experience and interact with gamification elements is a necessary extension. Next, we build upon previous literature showing the positive effects of gamifying questionnaires by investigating a gameful way of assessing Hexad user types. Given that the Hexad model is meant to be used within gamified systems, providing a gameful way of assessing user types seems to be imperative to turn the high potential of the Hexad model for tailoring gamified systems into practice.

3 GAMEFUL APPLICATIONS

We implemented two gameful web applications using AngularJS[1] and Phaser[2]. In the following, the concepts behind both gameful applications are explained and discussed. For both applications, interactive step-by-step tutorials were developed, explaining how to interact with them. We used the systematic literature review by Keusch et al. [17] to ensure that most relevant papers on gamified surveys have been considered to inform the design of the gameful applications.

3.1 Gameful Application 1: Cloud Clicker

Similar to Triantoro et al. [45], we decided to transfer the 7-point Likert scales, which the Hexad questionnaire uses into binary choice questions in the gameful application. The whole design of the first gameful application ("Cloud Clicker") followed the process proposed by Harms et al. [12, 14] and considered recommendations and lessons learned from relevant previous work [12, 13, 17, 45], as described in the following. In Cloud Clicker, users see two statements in each of 15 rounds and then have to decide which of these two statements is more relevant to them (see Figure 1a). We selected one statement for each Hexad user type, which grasps its underlying motivation. We based the statements on the definitions given by Tondello et al. [44] and on items with a substantial factor load in the confirmatory factor analysis in the validation study of the Hexad user types questionnaire [42]. This follows the same procedure as Triantoro et al. [45] proposed to translate the Big-5 survey into a gamified counterpart.

Figure 2: Illustrations and statements used in Cloud Clicker.

Next, as part of the "aesthetics and relationship" layer of the design process for gamified surveys by Harms et al. [12, 14], we

[1]https://angularjs.org/, last accessed August 5, 2020
[2]https://phaser.io/, last accessed August 5, 2020

decided to present the two statements shown to the user using cloud visualizations to create visual sensation [12] and a convincing and motivational environment [17]. To enhance questions and support the comprehensibility of the statements [12, 17], we created visual illustrations for each statement, explaining the statement by using a gender-neutral avatar (Figure 2). For the Philanthropist, we decided to focus on the aspect of helping others because the statements *It makes me happy if I am able to help others* and *I like helping others to orient themselves in new situations* had the highest factor load for this type [42]. For Socialisers, we focused on being part of a team because the corresponding statement *I like being part of a team* had the highest factor load among this trait [42].

Similarly, we used the statements having the highest factor load in a trait for the Free Spirit (*It is important to follow my own path*), the Achiever (*I like overcoming obstacles*) and the Player (*Rewards are a great way to motivate me*) [42]. Because the item having the highest factor load for the Disruptor type (*I see myself as a rebel*) was hard to illustrate visually, we decided to use the item *I dislike following rules* instead, which had the second highest factor load [42]. Similar to Hallifax et al. [10], statements were presented to users using a full paired-comparison design (i.e., each user was asked to evaluate all possible pairs of statements, resulting in 15 rounds of comparisons in Cloud Clicker). The order of the statements as well as whether a cloud with a statement was shown on the left or on the right side of the screen in a round was randomized to avoid biasing results [12]. Cloud Clicker provides a ranking of the statements, in which scores of 0–5 are distributed across each statement representing its corresponding user type, because each statement is compared to every other statement.

Clouds were shaking and dropping coins when being clicked (i.e., when a user decided on a particular statement) to increase the gameful experience of Cloud Clicker. Also, we added sound effects to indicate interactions with the gameful applications. Both follow recommendations by Harms et al. [12], stating that gameful feedback in surveys should be provided by using indicators such as coins as rewards and supported by using auditory feedback. The coins were colour-coded to represent the corresponding Hexad user type and were showing a miniature version of the illustration of the related user type. Similar to Triantoro et al. [45], we introduced time pressure when participants were asked to decide between two statements in each round. This mechanic had three reasons: First, it emphasizes and stimulates the gameful experience of the application [45]. Second, it supports spontaneous responses, which was shown to increase the reliability of responses [26]. Third, it limits the time it takes to complete the application and thus to assess the Hexad user type, which might be important as we aim to provide a practical way of assessing Hexad user types in gamified systems. To allow researchers and practitioners to use Cloud Clicker, we published the source code as well as all graphical assets on GitHub[3].

3.2 Gameful Application 2: Snowball Shooter

Cloud Clicker aims at providing a gameful way of assessing Hexad user types and, thus, builds on the original items from the Hexad questionnaire [42]. In contrast, Snowball Shooter focuses on user behaviour when interacting with gameful elements and whether it is possible to use this input to predict Hexad user types. As such, Snowball Shooter provides some gamification elements that users can interact with (see Figure 1b): The user controls a snowball cannon and shoots snowballs at items representing each gamification element. These items are randomly positioned in each round. Shooting at an item increases the internal score of the corresponding gamification element. Similar to Cloud Clicker, the application consists of 15 rounds in which users may shoot five snowballs. It uses feedback sounds when shooting.

Figure 3: Visualizations of the score levels for Unlockables, Achievements and Virtual Character in Snowball Shooter.

Following design suggestions from Harms et al. [12], we integrated progression loops in each gamification element. Consequently, there are three score thresholds, which lead to a state change of the corresponding element (e.g., unlocking a virtual item). To ensure comparability, these thresholds were kept constant across all gamification elements (given that a maximum score of 75 can be reached, the first state change happens at a score of 10, the second at 35 and the third at 60). As a result, it is impossible to complete all gamification elements. This was explained to users in the tutorial. It is important to note that—in contrast to previous work—users could experience the gamification elements instead of being given storyboards or textual descriptions only. Based on the user type descriptions by Marczewski [20] and the proposed gamification elements by Tondello et al. [44], we integrated the following gamification elements (Figure 1b):

Unlockables: Unlockables are expected to motivate **Free Spirits** because they are mainly driven by autonomy [20, 44]. In Snowball Shooter, we decided to provide treasure chests which could be unlocked to obtain virtual items. Reflecting the score thresholds, there are three different types of treasure chests (wooden, silver, golden) unlocking items of different rarity (common, rare, epic; see Figure 3).

Achievements: This element was shown to be especially suitable for **Achievers** as it supports mastery [20, 44]. In Snowball Shooter,

[3]https://github.com/m-altmeyer/cloud-clicker, last accessed August 5, 2020

three Achievements (using the score thresholds mentioned before) can be unlocked: "Snowball Enthusiast" (bronze frame), "Snowball Master" (silver frame) and "Snowball Guru" (golden frame), see Figure 3.

Points: Points have been shown to positively affect **Players** [20, 44]. To underline the value of points as virtual currency (which is important for Players [44]), points can be used to buy modifications for the Virtual Character (see below). The amount of points that needs to be spent to buy all modifications (change or customize the colour of and add wings for the virtual character, see Figure 1b) equals the maximum score threshold described above.

Leaderboard: Social gamification elements such as Leaderboards are relevant for **Socialisers** [20, 44]. However, findings by Tondello et al. [44] and other researchers [3, 19, 30] consistently demonstrate that Leaderboards are also positively correlated to **Players**, **Achievers** and **Disruptors**, which is why we expect to also find such correlations in Snowball Shooter. Similar to Mekler et al. [23], we decided to show fictitious users, having scores based on the thresholds established before, to ensure that all participants have same chances to rise in ranks.

Virtual Character: **Philanthropists** are driven by purpose and like to care for others [20, 44]. Although no significant correlations have been shown, we expect that this gamification element should be particularly relevant for Philanthropists because it may induce feelings of care-taking. We used a virtual monster whose emotional state is coupled to the amount of snowballs shot at the respective item, a green doughnut. The three changes in its emotional state (Figure 3) are coupled to the thresholds described before.

4 METHOD

We used the aforementioned gameful applications to investigate the following hypotheses stemming from our review of the related literature:

H1: Gameful applications can be used to predict Hexad types.

- **H1a**: The score of the statements in Cloud Clicker is correlated to the corresponding Hexad user types and thus may be used to predict them.
- **H1b**: The amount of interactions with gameful elements in Snowball Shooter is correlated to the corresponding Hexad user types and thus may be used to predict them.

H2: The users' perception of the gameful applications differs compared to their perception of the Hexad questionnaire.

- **H2a**: Both applications are perceived as more enjoyable (measured by the IMI enjoyment subscale) than the Hexad questionnaire.
- **H2b**: Participants feel more competent (measured by the IMI competence subscale) using both applications than using the Hexad questionnaire.
- **H2c**: Participants feel more pressure (measured by the IMI pressure subscale) in both applications than in the Hexad questionnaire.
- **H2d**: Both applications are perceived as more immersive (measured by the PXI immersion subscale) than the Hexad questionnaire.

H1 is motivated by previous work showing that questionnaires can be transformed into gameful applications without heavily affecting their validity [45]. Triantoro et al. [45] demonstrated that the Big-5 personality traits can be predicted based on gameful, binary choices in their survey, which is similar to our approach and thus motivates **H1a**. The subjective assessments of preferences for gamification elements using textual descriptions [44] or storyboards [3, 30] are correlated to the Hexad user types, which motivated us to find similar correlations when investigating actual interaction with implemented gamification elements (**H1b**).

H2a is mainly based on previous work in the domain of gamified surveys [12, 13, 45], where positive effects on enjoyment-related measures have been demonstrated. **H2b** relates back to feedback provided by gamification elements having been shown to increase perceived competence [36]. We expect to see an increase in perceived pressure in the gameful applications mainly because of the time pressure that is induced by Cloud Clicker and because of the round-based nature of both applications (**H2c**). An increase in pressure does not necessarily affect user experience negatively but might help to shape optimally challenging systems [12]. This supports users in reaching a flow state, which is described as a state of increased concentration and enjoyment [7]. To better understand whether the perceived pressure related to feelings of flow and immersion, we also evaluated immersion as part of the PXI questionnaire and expected it to be higher in the gameful applications (**H2d**).

4.1 Procedure

We conducted an online study on Prolific[4], an online platform specifically targeted at recruiting participants for scientific research studies. The only requirement was an understanding of the English language. The study has been reviewed and received ethics clearance through a University of Waterloo Research Ethics Committee (ORE#41608). It took approximately 15–20 minutes to complete and participants were paid £2 GBP. After giving informed consent, they were asked to provide demographic data including age and gender. Next, the 24-item Hexad user types questionnaire [42] was administered. The questionnaire consists of four items for each of the six user types, being measured on 7-point scales. To obtain a baseline for how participants perceived filling out the Hexad questionnaire, they were asked to fill out the 22-item task evaluation questionnaire of the Intrinsic Motivation Inventory ("IMI") [21, 34] as well as the "Immersion" subscale of the Player Experience Inventory ("PXI") [1]. Both the IMI items and the PXI items are measured on 7-point scales.

Next, participants were asked to interact with the gameful applications. The order of the gameful applications was randomized. Before starting the actual application, participants had to complete a tutorial explaining how to interact with them. In Cloud Clicker, we measured how often participants chose each statement. Similarly, the number of interactions with each gamification element in Snowball Shooter was measured. After interacting with each

[4]https://www.prolific.co/, last accessed August 5, 2020

gameful application, participants were asked to fill out the IMI questionnaire and the PXI “Immersion” subscale.

4.2 Participants

After removing participants who preferred not to answer questions of the Hexad questionnaire, 147 participants were considered for the analysis. Of those, 49% self-reported their gender as female, 49% as male, 0.7% as non-binary and 1.3% preferred not to answer this question. The mean age was 33 years (SD=11.5, Mdn=30, Min=18, Max=66). The Hexad user types average scores are similar to the averages reported in the validation study of the Hexad questionnaire by Tondello et al. [42]. Achievers showed the highest average scores (M=23.6, SD=2.98), followed by Philanthropists (M=22.8, SD=3.18), Players (M=22.8, SD=3.53) and Free Spirits (M=22.3, SD=3.52). Socialisers (M=18.7, SD=4.89) and Disruptors (M=15.0, SD=4.49) followed with lower average scores.

5 RESULTS

In this section, we present results related to predicting Hexad user types based on each gameful application as well as findings related to the enjoyment and perception of them.

5.1 Cloud Clicker and Hexad User Types

To analyze whether Cloud Clicker may be used to predict Hexad user types, a canonical correlation analysis (“CCA”) was conducted using the score of the six statements in the gameful application as predictors of the six Hexad user types measured by the Hexad user types questionnaire. A CCA is preferable when analyzing the association strength between two sets of variables and allows to evaluate the multivariate shared variance between them (i.e., between the six statement scores of the gameful application and the six scores of the Hexad subscales) [38]. Next, this method is explained based on Sherry and Hanson's guide on using CCA [38].

The core idea of CCA is that the set of predictor variables and the set of criterion variables are combined into a synthetic variable each (i.e., there is a synthetic predictor and a synthetic criterion variable). The canonical correlation is the correlation between these synthetic variables. Each pair of synthetic variables is called a canonical function (“CF”). Canonical functions are comparable to principal components in Principal Component Analyses (PCA) with the main difference that the CFs are composed of two different variable sets and thus can be seen as an extension of PCA [46]. In line with this, CCA was loosely defined as “a double-barreled principal components analysis” [40].

As long as there is residual variance left in the two variable sets which cannot be explained by the already derived canonical functions, the above process is repeated. This continues until either no residual variance is left to be explained or there are as many canonical functions as there are variables in the smaller variable set. Although CCA can accommodate variables without relying strictly on multivariate normality [46], multivariate normality was assessed by inspecting univariate Q-Q plots, skewness, and kurtosis of each variable included in the CCA. The Q-Q plots mainly supported the assumptions of normality, whereas some variables were shown to be slightly skewed. However, all skewness and kurtosis values were within the acceptable thresholds of skewness < 3 and kurtosis < 8 [18], given that the maximum absolute values of skewness and kurtosis were found to be 2.5 and 6.8 respectively such that the CCA could be conducted. Given that 10 participants per observed variable are recommended to reach a reliability of 80% [39], our sample size can be considered as adequate.

Overall, the full model across all CF was statistically significant using the Wilks's λ=.256 criterion, $F(36, 595.59) = 6.01, p < .001$. This shows that the variance unexplained by the model is 25.63%. Consequently, the full model is able to explain 74.37% of the variance (the r^2 type effect size is .74) shared between the two variable sets. Given that the recommended threshold for strong effects was derived to be $r^2 = .64$ [9], the model can be considered to explain a substantial amount of variance between the two variable sets. Based on this, we derive **R1: The score of statements in the gameful application is substantially associated to the scores of the Hexad user type questionnaire.** This result shows that the two variable sets are strongly related. As a next step of the CCA, we will consider the results of the dimension reduction analysis to analyze whether the predictor variables (the score of the statements in Cloud Clicker) load on the same canonical functions as the corresponding Hexad user types. This is important to investigate whether the statements we have chosen for a certain user type actually represent this user type, given our data.

The dimension reduction analysis yielded six canonical functions (CF1–CF6) with squared canonical correlations of .36, .33, .22, .17, .09 and .00 each. The first five canonical functions were statistically significant whereas CF6 did not explain a statistically significant amount of shared variance between the variable sets (CF1–CF4: $p < .001$, CF5: $p = .011$, CF6: $p = .99$). Therefore, CF6 will not be interpreted as part of the analysis. Figure 4 presents the structure coefficients for CF1–CF5 being stronger than |.35| (an upper threshold for weak factor loads established in [47]). Dotted and transparent lines indicate relationships which differed between the predictor and criterion variables. All standardized canonical function coefficients and structure coefficients can be found in Table 1.

	CF 1		CF2		CF3		CF4		CF5	
Pred.	*co*	*rs*	*co*	*rs*	*co*	*rs*	*co*	*rs*	*co*	*rs*
G_DI	.10	**.46**	.04	-.21	.19	-.03	-.06	-.10	1.34	**.76**
G_FS	.38	**.73**	.11	-.12	.32	**.36**	-.12	-.22	.49	-.17
G_AC	-.53	**-.37**	.21	.01	.89	**.85**	.34	**.36**	.76	-.07
G_PL	-.11	.18	.99	**.91**	-.13	-.19	.45	.31	.90	.01
G_PH	-.22	-.30	-.21	**-.64**	-.29	**-.46**	.72	**.45**	.69	.10
G_SO	-.64	**-.67**	.40	-.04	.10	-.30	-.57	**-.65**	.98	.04
Crit.										
Hex_DI	.30	**.41**	-.31	**-.37**	-.18	-.08	-.08	-.17	1.00	**.78**
Hex_FS	.75	**.39**	-.01	-.34	-.09	.19	-.50	**-.41**	-.70	-.08
Hex_AC	-.56	-.33	-.07	-.05	1.24	**.59**	-.11	-.19	.37	.30
Hex_PL	.24	-.21	.92	**.60**	-.40	-.12	.26	.05	.16	**.36**
Hex_PH	-.16	-.31	-.57	**-.51**	-.21	-.12	.99	**.36**	-.11	.07
Hex_SO	-.64	**-.63**	-.19	-.17	-.58	**-.35**	-.87	**-.46**	.03	.13

Table 1: Structure coefficients (rs) and standardized canonical function coefficients (co) for predictor variables (statement scores in Cloud Clicker: G_DI etc.) and criterion variables (user type scores: Hex_DI etc.) for the canonical functions. Bold entries represent loads higher than |.35|, underlined entries represent loads higher than |.50|.

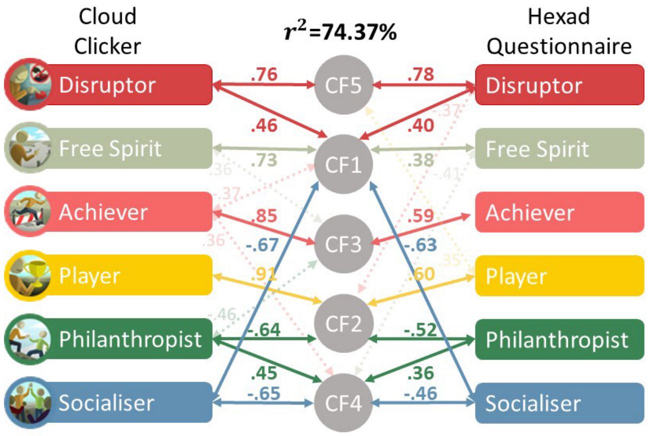

Figure 4: Structure coefficients for CF1–CF5 being stronger than |.35| for Cloud Clicker. Dotted and transparent lines indicate relationships which differed between the predictor and criterion variables.

While standardized canonical function coefficients represent the weights applied to the observed variables to combine the unobserved synthetic variables, structure coefficients are simple bivariate correlations between observed variables and synthetic variables [27]. It can be seen that most predictor and criterion variables have large structure coefficients loading substantially (i.e., > |.5| according to [6]) on the same canonical functions. This is supported by the symmetry of the relationships, which can be seen in Figure 4. In terms of strength of the correlation, the Free Spirit subscale is an exemption to this as it is the only variable having no structure coefficient higher than |.5| but loads moderately [6] on CF1.

Thus, we formulate **R2: All predictor and criterion variables have medium to large structure coefficients loading on the same canonical functions**. This result shows that there is not only a substantial relationship between the variable sets (**R1**) but also that the predictor variables load on the same canonical functions as the criterion variables. This means that the statements we have chosen represent the corresponding Hexad user types.

	DI_R	FS_R	AC_R	PL_R	SO_R	PH_R
G_DI	**.32**				**-.23**	
G_FS	.21	**.36**			-.36	
G_AC			**.35**			
G_PL				**.50**		
G_SO	**-.32**				**.44**	
G_PH				**-.22**		**.56**
DI	**.44**			-.22	**-.37**	
FS		**.52**		-.38		
AC			**.45**			
PL		-.37		**.60**		-.32
SO	**-.52**	-.31			**.64**	
PH		-.31	-.23	**-.26**		**.54**

Table 2: Spearman's rank correlation coefficients between the ranked Hexad user type scores (DI_R etc.) and the score of each statement in Cloud Clicker (G_DI etc.) as well as between the ranked Hexad user type scores and the absolute Hexad user type scores (DI etc.). All $p < .01$.

To further analyze the correlations between the Hexad user types questionnaire scores and the scores obtained through the gameful application, we calculated bivariate Spearman's rank correlation coefficients. Since the gameful application requires users to make a binary choice whereas the original Hexad user types questionnaire allows to have the same score in multiple user types, we ranked the Hexad user type scores by assigning values from 0–5 to ensure comparability. These ranked Hexad scores are denoted by "_R" in the results. As a reference for interpretation, we also added the correlations between the absolute score of each user type of the Hexad and the ranked Hexad user type. The results are shown in Table 2.

Providing further support for the results of the CCA, it can be seen that there are medium to large size correlations [6] between the scores of each statement of the gameful application and the ranked Hexad user types. This leads to **R3: The ranked Hexad scores are positively correlated to the corresponding scores of each statement of Cloud Clicker having medium to large effect sizes**. This provides further support for the suitability of the statements and visualizations used in Cloud Clicker. It can also be seen that the correlations between the ranked Hexad user type scores and the score of each statement in the gameful application are similar to the correlations between the absolute Hexad user type scores calculated using the Hexad questionnaire and the ranked Hexad user type scores concerning both strength and direction of the correlations.

Thus, we formulate **R4: The correlations between the scores of the statements of Cloud Clicker and the ranked Hexad scores are similar to the correlations between the absolute Hexad user type scores and the ranked Hexad user type scores**. **R4** is reflected visually by the two highlighted diagonals in Table 2. In line with the results of the CCA, this indicates that assessing the ranking of Hexad user types with Cloud Clicker is comparable to assessing the ranking of Hexad user types based on the questionnaire. Consequently, taking **R1–R4** together, our results demonstrate that Cloud Clicker explains a substantial amount of shared variance between the predictor and criterion variable sets, that the statements and visualizations used in Cloud Clicker successfully represent their corresponding Hexad user types and that Cloud Clicker can be used to assess the ranking of a user's Hexad type scores.

5.2 Snowball Shooter and Hexad User Types

Again, a CCA was conducted to investigate the shared variance between the amount of interactions with each gamification element in Snowball Shooter as predictor variables and the Hexad scores of each user type. The analysis yielded five canonical functions with squared canonical correlations of .20, .12, .07, .01, and .00. The full model was statistically significant (Wilks's $\lambda = .659$ criterion, $F(30, 546.00) = 2.00, p = .001$). This leads to result **R5: The amount of interactions with gamification elements in Snowball Shooter is moderately associated to the scores of the validated Hexad user type questionnaire.** Similar to Cloud Clicker, this means that the amount of interactions with gamification elements and the Hexad user types are related. However, the variance shared between the two sets of variables was considerably

	CF 1	
Predictor	*co*	*rs*
Collectibles	.42	-.05
Achievements	.20	.01
Points	.37	.06
Leaderboard	1.08	**<u>.89</u>**
Virtual Character	-.06	**<u>-.72</u>**

	CF1	
Criterion	*co*	*rs*
Hex_DI	.37	**.37**
Hex_FS	-.35	.10
Hex_AC	.62	**<u>.69</u>**
Hex_PL	.51	**<u>.79</u>**
Hex_PH	-.40	.09
Hex_SO	.22	**.46**

Table 3: Structure coefficients (rs) and standardized canonical function coefficients (co) for predictor variables (number of interactions with gamification elements in Snowball Shooter) and criterion variables (score of each user type: Hex_DI etc.) for CF1. Bold entries represent loads higher than $|.35|$, underlined ones higher than $|.50|$.

lower than in Cloud Clicker as the model of Snowball Shooter accounts for 34.1% of the shared variance. This indicates a moderate effect size [9]. As part of the dimension reduction analysis, it was found that solely CF1 was explaining a significant amount of variance. Therefore, only CF1 was considered for the interpretation of the canonical correlation analysis. An overview of the structure coefficients and standardized canonical function coefficients for CF1 can be found in Table 3.

Looking at the CF1 coefficients, it can be seen that the score in the gamification elements Leaderboard and Virtual Character strongly contribute to the synthetic predictor variable. While the score in Leaderboard contributes positively to CF1, Virtual Character contributes negatively. Regarding the criterion variable set in CF1, Achiever and Player were the primary contributors to the criterion synthetic variable, with secondary contributions by Socialiser and Disruptor. All of the aforementioned variables add positively to CF1. Thus, the amount of interactions with the Leaderboard is positively related to the score in the Achiever, Player, Disruptor and Socialiser factors of the Hexad. This is in line with previous results based on self-reported preferences for gamification elements [44]. That the Virtual Character is negatively contributing to CF1 indicates that participants interacting with it were likely not interested in interacting with Leaderboards and tended to score lower on the Player, Achiever, Socialiser and Disruptor types.

Analyzing the relationships between interaction with gamification elements and Hexad user types further, we again calculated bivariate Spearman's rank correlations. Similar to Cloud Clicker, we considered the absolute score in each Hexad user type as well as the ranked Hexad user type. Table 4 shows significant correlation coefficients.

	DI	AC	PL	SO	PH_R
Collectibles					
Achievements		$.19^{*}$			
Points					
Leaderboard	$.19^{*}$	$.30^{**}$	$.33^{**}$	$.23^{**}$	$-.23^{**}$
Virtual Character		$-.25^{**}$	$-.22^{**}$		$.19^{*}$

Table 4: Spearman's rank correlation coefficients between the amount of interactions with each gamification element in the Snowball Shooter application and the absolute Hexad user type scores (DI etc.) as well as the ranked Hexad scores (PH_R). Hexad user types having at least one significant correlation are shown. $^{*}p < .05$, $^{**}p < .01$

Overall, the correlations support the results from the canonical correlation analysis. It can be seen that most correlations were found for the Leaderboard and Virtual Character gamification elements. The positive correlation between the ranked Philanthropist score and the amount of interactions with the Virtual Character gamification element indicates that, as expected based on the definition of the user type [44], the Virtual Character seems to be particularly relevant for the Philanthropist. Also, the positive correlations between the Achiever, Player, Disruptor and Socialiser user types and the Leaderboard are in line with the results from the canonical correlation analysis and were expected based on previous work [44]. In addition, the positive correlation between Achievements and the Achiever was expected and is in line with previous findings [44].

Taking these results all into account together, we establish **R6: The amount of interactions with gamification elements correlates to their corresponding Hexad user types**. On a more abstract level, **R5** and **R6** mean that users interact with gamification elements that correspond to their Hexad user types. This is an important result for the validity of the Hexad model, as previous research did not consider actual user behaviour within gameful applications, as far as we know. However, it should be noted that we could not find correlations between Collectibles and Free Spirits as well as between Points and Players.

5.3 Perception of Cloud Clicker and Snowball Shooter

To analyze the perception of the gameful applications compared to completing the Hexad user types questionnaire, a repeated measures Friedman ANOVA was calculated for the IMI and PXI factors (the responses in the IMI and PXI responses were not normally distributed). The Durbin-Conover method was used for post-hoc analysis and the Benjamini-Hochberg false discovery rate [4] was used to adjust significance values for multiple comparisons. For this analysis, all participants who decided to not answer either the PXI or IMI questions after the Hexad questionnaire, Cloud Clicker, or Snowball Shooter, were excluded. Thus, the responses of 113 participants were considered.

Table 5 provides an overview of the mean and standard deviations of the IMI subscales and the PXI Immersion subscale in each condition (Hexad questionnaire, Cloud Clicker, Snowball Shooter). When analyzing the Competence subscale of the IMI, we did not find significant differences between the conditions ($\chi^2(2) = 5.54, p = .063$). Similarly, no significant differences were found for the Choice subscale of the IMI ($\chi^2(2) = 4.94, p = .085$). However, the Enjoyment score differed significantly ($\chi^2(2) = 35.8, p < .001$). The post-hoc analysis revealed that both gameful applications were significantly more enjoyable than completing the Hexad questionnaire ($p < .001$ each). Also, Snowball Shooter was perceived as more enjoyable than Cloud Clicker ($p = 0.012$).

We summarize these effects by **R7: Both gameful applications are perceived as more enjoyable than completing the Hexad questionnaire**. This result indicates that gameful approaches might be more suitable to be used within gameful systems than the Hexad questionnaire, when a gameful experience is important. In addition, we found that the perceived pressure differed significantly between

	Hexad	G1	G2
IMI Competence	5.00 /1.05	5.08 / 1.13	4.84 / 1.32
IMI Choice	5.29 / 1.47	5.29 / 1.37	5.48 / 1.36
IMI Enjoyment*	3.87 / 1.34	4.56 / 1.45	4.81 / 1.62
IMI Pressure*	1.96 / 0.98	2.44 / 1.30	2.34 / 1.27
IMI Immersion*	4.91 / 1.10	5.24 / 1.21	5.40 / 1.23

Table 5: Mean / Standard Deviation for each condition of the study. All variables are measured on 7-point scales. Variables for which the Friedman ANOVA was significant are marked (*). G1=Cloud Clicker, G2=Snowball Shooter.

the conditions ($\chi^2(2) = 11.7, p = .003$). Both gameful applications scored higher in the Pressure factor of the IMI (both $p < .001$), whereas no difference was found between the gameful applications themselves ($p = 1.00$), leading to **R8: The perceived pressure is significantly higher in both gameful applications**. This finding is likely related to the timed and round-based nature of the gameful applications (i.e., the fact that we used a timer in Cloud Clicker and 15 rounds of interaction in both gameful applications).

Finally, we analyzed whether the immersion, as measured by the Immersion subscale of the PXI, differs across the conditions. The Friedman ANOVA revealed a significant effect ($\chi^2(2) = 26.5, p <$ .001). Both gameful applications scored significantly higher on the Immersion subscale of the PXI (each p<.001) while there were no differences between the two gameful applications ($p = .127$). Thus, we derive **R9: Both gameful applications were perceived as more immersive than completing the Hexad questionnaire**. Taking **R7–R9** together, it seems like the higher pressure is not perceived negatively but may cause a feeling of higher immersion leading to a more enjoyable experience [7].

6 DISCUSSION

Our results show the scores of the statements ("predictor variables") in the Cloud Clicker application and the Hexad user type scores ("criterion variables") are substantially related to each other. They share 74.37% of their variance (**R1**). We also found the structure coefficients of the predictor variables and the structure coefficients of the criterion variables load on the same canonical functions, with large effect sizes (**R2**). In addition, we found there are medium-to-large correlations between the ranked Hexad scores and the scores of the statements in Cloud Clicker (**R3**), which were shown to be comparable to the correlations between the absolute Hexad scores and the ranked Hexad scores (**R4**). Taking **R1–R4** together, we conclude that Cloud Clicker can be used to predict Hexad user types in a gameful way, when an order of user types is sufficient (which is likely the case when personalizing gamification elements set in a gameful system).

We suggest using the order of scores, since Cloud Clicker uses a binary choice instead of allowing users to rate their agreement with statements on an ordinal scale (as was done in the validated Hexad questionnaire). Ultimately, these results support **H1a: The score of the statements in Cloud Clicker is correlated to the corresponding Hexad user types and thus may be used to predict them**. This is explainable because we used statements which were similar to the Hexad questionnaire items with the highest factor load.

Regarding the Snowball Shooter application, in which we analyzed whether the amount of interactions with gamification elements ("predictor variables") could be used to predict Hexad user types ("criterion variables"), we found that in general, there is a relationship between the predictor and criterion variables (**R5**). This is an important finding on its own because it shows that the correlations between the Hexad user types and preferences for gamification elements—which have been identified based on self-reports using textual descriptions or storyboards in previous work [3, 44]—can be replicated based on actual user behaviour. This finding supports the suitability of the Hexad model to explain user behaviour in gameful systems. When analyzing correlations between the amount of interactions with gamification elements and Hexad user types further, we found correlations that were expected based on previous work (**R6**). This is in line with the findings by Hallifax et al. [10], and supports the validity of the Hexad model.

However, it should be noted that two correlations that were expected could not be found (between Free Spirits and Unlockables as well as between Players and Points). A reason might be that Snowball Shooter did not motivate a specific behaviour (as gameful systems usually do [11]) but rather encouraged users to try out different gamification elements. This would be likely for Free Spirits who like to explore [44] and thus might explain the absence of correlations for this user type. The unlocked items could only be collected and not be used for anything else. This might have affected the engagement of users negatively. Also, an incentive for collecting points was missing, which might have been detrimental to Players' motivation to collect points. Considering that the shared variance between predictor and criterion variable sets was moderate (34.11%), we do not recommend deriving Hexad user types based on interaction behaviour alone in Snowball Shooter. However, the amount of interaction with gamification elements might still be a useful factor for dynamic adjustments of gameful systems. Based on **R5** and **R6**, we consider **H1b: The amount of interactions with gamification elements in Snowball Shooter is correlated to the corresponding Hexad user types and thus may be used to predict them** partially supported. Although we found that there are correlations to gamification elements that match the corresponding Hexad user types, the amount of shared variance between the two sets of variables is too low to reliably predict Hexad user types.

Furthermore, our results show that both gameful applications were perceived as more enjoyable (**R7**) and more immersive (**R9**) than the traditional Hexad questionnaire. Based on these results, **H2a: Both applications are perceived as more enjoyable** and **H2d: Both applications are perceived as more immersive** are supported. We also found a significant effect on perceived pressure (**R8**), which might be related to the higher immersion and arguably a higher sense of flow [7, 12]. Based on this, **H2c: Participants feel more pressure in both applications** is supported.

We conclude that both gameful applications provide a more pleasurable gameful experience than completing the Hexad questionnaire. **H2b** is not supported because no effects were found regarding perceived competence. Contrary to previous work [36], it seems like the gamification elements did not enhance the perceived competence through feedback as much as expected. A potential explanation might be that the user interface elements (such as radio buttons) provide visual feedback on their own, potentially enhancing the perceived competence in the baseline condition.

On a more abstract level, our results show that Hexad user types can be assessed by using binary choices. These could be easily adapted to different contexts or could even be turned into concrete choices a player needs to make in a more game-like setting. Also, we show that interacting with gameful design elements is related to a user's Hexad type. This could be used to infer Hexad user types dynamically when interacting with a gameful system and provides huge potential for further research.

6.1 Limitations and Future Work

Our work has several limitations. First, transforming the 7-point Likert scales into binary decisions and considering only one particularly relevant item per Hexad user type as was done in Cloud Clicker unavoidably leads to a loss of information. Consequently, we recommend using Cloud Clicker as a practical tool when personalizing gameful systems while ensuring a gameful user experience and to prevent a loss of immersion. For scientific purposes, we acknowledge that Cloud Clicker cannot replace the validated Hexad questionnaire [42]. Second, we used statements in Cloud Clicker that were similar to the statements with high factor loads in the Hexad questionnaire validation study (but not the same). We decided for one statement instead of all four statements. This means that even though we used statements that had a substantial factor load, using other statements for the corresponding user types might lead to different results.

Next, regarding the Snowball Shooter application, it did not motivate a real-life behaviour but allowed users to interact with the gamification elements. Although this allows users to experience how certain gamification elements work, their perception might be different when motivating concrete real-life goals in specific domains such as physical activity. The free exploration of gamification elements within Snowball Shooter could be particularly appreciated by Free Spirits, which might explain why we could not find any correlations for this user type (because they might have tried several different gamification elements).

Also, we decided to randomize the order of the gameful applications, but not of the Hexad questionnaire. This was done to avoid detrimental effects of removing gamification on the perception of the Hexad questionnaire [11]. Since the IMI and PXI scales are the only shared dependent variables across these conditions, potential ordering effects would not primarily concern the main goal of the study, i.e. predicting Hexad user types from interaction behaviour. In addition, we assume the chance of ordering effects is low since filling out a questionnaire and interacting with gameful applications can be considered different tasks, reducing the chance of practice effects [32]. Nevertheless, the fact that participants always started by completing the Hexad questionnaire should be considered.

Since participants were asked to answer roughly 100 items in total, we cannot rule out fatigue effects. However, considering the number of items and that the duration of the study is within a maximum length of 20 minutes [33], no practically relevant effects on data quality are to be expected [15].

Last, we acknowledge that, although the design of the applications is based on previous research, certain decisions are inherently a matter of interpretation, which might affect the external validity.

Future work should investigate whether different game controls (e.g., allowing for continuous user input) will enhance the accuracy of using gameful applications to predict Hexad user types. Further research should be conducted into correlations between actual user behaviour and Hexad user types to replicate the findings of previous research. Also, our findings should be validated in different domains.

7 CONCLUSION

To reduce the barrier of having to fill out Hexad questionnaires for user types assessments, we conceptualized and implemented two gameful applications to predict Hexad user types in a gameful way. We contribute two main findings: First, we show that the first gameful application, Cloud Clicker, indeed could be used to predict Hexad user types. The interaction within the gameful application and the validated Hexad scores are correlated for all user types and share a substantial amount of variance. Cloud Clicker could be used to tailor gameful systems without having to complete a questionnaire.

Second, we showed that the interaction behaviour with gameful elements correlates to the corresponding Hexad user types. This is important because previous work only used self-reported measures based on storyboards or textual descriptions explaining the gameful elements. Our results highlight that enjoyment and immersion are significantly higher in both gameful applications than when filling out the Hexad questionnaire.

8 ACKNOWLEDGEMENTS

We thank Karina Arrambide, Cayley MacArthur and the anonymous reviewers for their valuable feedback to improve this paper.

REFERENCES

[1] Vero Vanden Abeele, Katta Spiel, Lennart Nacke, Daniel Johnson, and Kathrin Gerling. 2020. Development and Validation of the Player Experience Inventory: A Scale to Measure Player Experiences at the Level of Functional and Psychosocial Consequences. *International Journal of Human Computer Studies* 135 (2020). https://doi.org/10.1016/j.ijhcs.2019.102370

[2] Maximilian Altmeyer and Pascal Lessel. 2017. The Importance of Social Relations for Well-Being Change in Old Age - Do Game Preferences Change As Well? *Proceedings of the Positive Gaming: Workshop on Gamification and Games for Wellbeing. Amsterdam, The Netherlands* (2017).

[3] Maximilian Altmeyer, Pascal Lessel, Linda Muller, and Antonio Krüger. 2019. Combining Behavior Change Intentions and User Types to Select Suitable Gamification Elements for Persuasive Fitness Systems. In *International Conference on Persuasive Technology*. Springer.

[4] Yoav Benjamini and Yosef Hochberg. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society* 58, 4 (1995), 619–678.

[5] Max V. Birk, Maximilian A. Friehs, and Regan L. Mandryk. 2017. Age-Based Preferences and Player Experience: A Crowdsourced Cross-sectional Study. *Proceedings of the Annual Symposium on Computer-Human Interaction in Play - CHI PLAY '17* (2017), 157–170. https://doi.org/10.1145/3116595.3116608

[6] Jacob Cohen. 1988. Statistical Power Analysis for the Behavioral Sciences, Hillsdale. *NJ: Lawrence Erlbaum Associates* (1988).

[7] Mihaly Csikszentmihalyi. 1997. *Finding Flow: The Psychology of Engagement with Everyday Life.* Basic Books, New York, NY, US. ix, 181–ix, 181 pages.

[8] Sebastian Deterding, Dan Dixon, Rilla Khaled, and Lennart Nacke. 2011. From Game Design Elements to Gamefulness: Defining Gamification. *Proceedings of the 15th International Academic MindTrek Conference. ACM, 2011.* (2011), 9–15. https://doi.org/10.1145/2181037.2181040

[9] Christopher J. Ferguson. 2009. An Effect Size Primer: A Guide for Clinicians and Researchers. *Professional Psychology: Research and Practice* 40, 5 (2009), 532–538. https://doi.org/10.1037/a0015808

[10] Stuart Hallifax, Audrey Serna, Jean-charles Marty, Guillaume Lavoué, and Elise Lavoué. 2019. Factors to Consider for Tailored Gamification. *Proceedings of the Annual Symposium on Computer-Human Interaction in Play - CHI PLAY '19* (2019).

[11] Juho Hamari and Harri Sarsa. 2014. Does Gamification Work? - A Literature Review of Empirical Studies on Gamification. *Hawaii International Conference on System Sciences.* (2014), 3025–3034. https://doi.org/10.1109/HICSS.2014.377

[12] Johannes Harms, Stefan Biegler, Christoph Wimmer, Karin Kappel, and Thomas Grechenig. 2015. Gamification of Online Surveys: Design Process, Case Study, and Evaluation. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 9296, Section 4 (2015), 219–236. https://doi.org/10.1007/978-3-319-22701-6_16

[13] Johannes Harms, Dominik Seitz, Christoph Wimmer, Karin Kappel, and Thomas Grechenig. 2015. Low-Cost Gamification of Online Surveys: Improving the User Experience through Achievement Badges. *CHI PLAY 2015 - Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play* (2015), 109–114. https://doi.org/10.1145/2793107.2793146

[14] Johannes Harms, Christoph Wimmer, Karin Kappel, and Thomas Grechenig. 2014. Gamification of Online Surveys: Conceptual Foundations and a Design Process based on the MDA Framework. *Proceedings of the NordiCHI 2014: The 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational* (2014), 565–568. https://doi.org/10.1145/2639189.2639230

[15] A. Regula Herzog and Jerald G. Bachman. 1981. Effects of Questionnaire Length on Response Quality. *Public Opinion Quarterly* 45, 4 (1981), 549–559. https://doi.org/10.1086/268687

[16] Yuan Jia, Bin Xu, Yamini Karanam, and Stephen Voida. 2016. Personality-Targeted Gamification: A Survey Study on Personality Traits and Motivational Affordances. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16* (2016), 2001–2013. https://doi.org/10.1145/2858036.2858515

[17] Florian Keusch and Chan Zhang. 2017. A Review of Issues in Gamified Surveys. *Social Science Computer Review* 35, 2 (2017), 147–166. https://doi.org/10.1177/0894439315608451

[18] Rex B. Kline. 2011. *Principles and Practice of Strctural Equation Modeling.* Vol. 53. 1689–1699 pages. https://doi.org/10.1017/CBO9781107415324.004 arXiv:arXiv:1011.1669v3

[19] Dimosthenis Kotsopoulos, Cleopatra Bardaki, Stavros Lounis, and Katerina Pramatari. 2018. Employee Profiles and Preferences towards IoT-enabled Gamification for Energy Conservation. *International Journal of Serious Games* 5, 2 (2018), 65–85. https://doi.org/10.17083/ijsg.v5i2.225

[20] Andrzej Marczewski. 2015. *Even Ninja Monkeys Like to Play: Gamification, Game Thinking and Motivational Design.* CreateSpace Independent Publishing Platform.

[21] Edward D. McAuley, Terry Duncan, and Vance V. Tammen. 1989. Psychometric Properties of the Intrinsic Motivation Inventoiy in a Competitive Sport Setting: A Confirmatory Factor Analysis. *Research Quarterly for Exercise and Sport* 60, 1 (1989), 48–58. https://doi.org/10.1080/02701367.1989.10607413

[22] Robert R. McCrae and Oliver P. John. 1992. An Introduction to the Five-Factor Model and Its Applications. *Journal of Personality* 60, 2 (1992), 175–215. https://doi.org/10.1111/j.1467-6494.1992.tb00970.x

[23] Elisa D. Mekler, Florian Brühlmann, Alexandre N. Tuch, and Klaus Opwis. 2017. Towards Understanding the Effects of Individual Gamification Elements on Intrinsic Motivation and Performance. *Computers in Human Behavior* 71 (2017), 525–534. https://doi.org/10.1016/j.chb.2015.08.048

[24] Alberto Mora, Gustavo F. Tondello, Lennart E. Nacke, and Joan Arnedo-Moreno. 2018. Effect of Personalized Gameful Design on Student Engagement. *IEEE Global Engineering Education Conference, EDUCON* 2018-April (2018), 1925–1933. https://doi.org/10.1109/EDUCON.2018.8363471

[25] Lennart E. Nacke, Chris Bateman, and Regan L. Mandryk. 2014. BrainHex: A Neurobiological Gamer Typology Survey. *Entertainment Computing* 5, 1 (2014), 55–62. https://doi.org/10.1016/j.entcom.2013.06.002

[26] Aljoscha C. Neubauer and Bertram F. Malle. 1997. Questionnaire Response Latencies: Implications for Personality Assessment and Self-Schema Theory. *European Journal of Psychological Assessment* 13, 2 (1997), 109–117. https://doi.org/10.1027/1015-5759.13.2.109

[27] Kim Nimon, Robin K. Henson, and Michael S. Gates. 2010. Revisiting Interpretation of Canonical Correlation Analysis: A Tutorial and Demonstration of Canonical Commonality Analysis. *Multivariate Behavioral Research* 45, 4 (2010), 702–724. https://doi.org/10.1080/00273171.2010.498293

[28] Rita Orji, Regan L Mandryk, and Julita Vassileva. 2015. Gender, Age, and Responsiveness to Cialdini's Persuasion Strategies. *International Conference on Persuasive Technology* 9072, June (2015). https://doi.org/10.1007/978-3-319-20306-5

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

[30] Rita Orji, Gustavo F Tondello, and Lennart E Nacke. 2018. Personalizing Persuasive Strategies in Gameful Systems to Gamification User Types. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '18* (2018). https://doi.org/doi.org/10.1145/3173574.3174009

[31] Kiemute Oyibo, Rita Orji, and Julita Vassileva. 2017. The Influence of Culture in the Effect of Age and Gender on Social Influence in Persuasive Technology. *Adjunct Publication of the 25th Conference on User Modeling, Adaptation and Personalization - UMAP '17* (2017), 47–52. https://doi.org/10.1145/3099023.3099071

[32] Paul Christopher Price, I-Chant A. Chiang, and Rajiv Jhangiani. 2015. Research Methods in Psychology. *BCcampus* 40, 5 (2015), 551–552. https://openlibrary-repo.ecampusontario.ca/jspui/handle/123456789/246

[33] Melanie Revilla and Carlos Ochoa. 2017. Ideal and Maximum Length for a Web Survey. *International Journal of Market Research* 59, 5 (2017), 557–565. https://doi.org/10.2501/IJMR-2017-039

[34] Richard M. Ryan. 1982. Control and Information in the Intrapersonal Sphere: An Extension of Cognitive Evaluation Theory. *Journal of Personality and Social Psychology* 43, 3 (1982), 450–461.

[35] Richard M Ryan and Edward L Deci. 2000. Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being. *American Psychologist* 55, 1 (2000), 68–78. https://doi.org/10.1037/0003-066X.55.1.68

[36] Michael Sailer, Jan Ulrich Hense, Sarah Katharina Mayr, and Heinz Mandl. 2017. How Gamification Motivates: An Experimental Study of the Effects of Specific Game Design Elements on Psychological Need Satisfaction. *Computers in Human Behavior* 69 (2017), 371–380. https://doi.org/10.1016/j.chb.2016.12.033 arXiv:arXiv:1011.1669v3

[37] Katie Seaborn and Deborah Fels. 2015. Gamification in Theory and Action: A Survey. *International Journal of Human-Computer Studies* 74 (2015), 14–31. https://doi.org/10.1016/j.ijhcs.2014.09.006

[38] Alissa Sherry and Robin K. Henson. 2005. Conducting and Interpreting Canonical Correlation Analysis in Personality Research: A User-Friendly Primer. *Journal of Personality Assessment* 3891 (2005), 37–48. https://doi.org/10.1207/s15327752jpa8401_09

[39] Barabara G. Tabachnick and Linda S. Fidell. 2013. *Using Multivariate Statistics* (sixth edit ed.). Pearson. 1–1018 pages.

[40] Bruce Thompson. 1984. *Canonical Correlation Analysis: Uses and Interpretation.* Number 47. Sage.

[41] Gustavo Fortes Tondello. 2019. *Dynamic Personalization of Gameful Interactive Systems.* Ph.D. Dissertation. University of Waterloo. http://hdl.handle.net/10012/14807

[42] Gustavo F. Tondello, Alberto Mora, Andrzej Marczewski, and Lennart E. Nacke. 2018. Empirical Validation of the Gamification User Types Hexad Scale in English and Spanish. *International Journal of Human-Computer Studies* (2018). https://doi.org/10.1016/j.ijhcs.2018.10.002

[43] Gustavo F. Tondello, Alberto Mora, and Lennart E. Nacke. 2017. Elements of Gameful Design Emerging from User Preferences. *Proceedings of the Annual Symposium on Computer-Human Interaction in Play - CHI PLAY '17* (2017), 129–142. https://doi.org/10.1145/3116595.3116627

[44] Gustavo F Tondello, Rina R Wehbe, Lisa Diamond, Marc Busch, Andrzej Marczewski, and Lennart E Nacke. 2016. The Gamification User Types Hexad Scale. *The ACM SIGCHI Annual Symposium on Computer-Human Interaction in Play - CHI PLAY '16* (2016). https://doi.org/10.1145/2967934.2968082

[45] Tamilla Triantoro, Ram Gopal, Raquel Benbunan-Fich, and Guido Lang. 2019. Would You Like to Play? A Comparison of a Gamified Survey with a Traditional Online Survey Method. *International Journal of Information Management* 49, April (2019), 242–252. https://doi.org/10.1016/j.ijinfomgt.2019.06.001

[46] Hao-Ting Wang, Jonathan Smallwood, Janaina Mourao-Miranda, Cedric Huchuan Xia, Theodore D. Satterthwaite, Danielle S. Bassett, and Danilo Bzdok. 2018. Finding the Needle in High-Dimensional Haystack: A Tutorial on Canonical Correlation Analysis. *arXiv preprint arXiv:1812.02598* (2018). arXiv:1812.02598 http://arxiv.org/abs/1812.02598

[47] Carmen Ximénez. 2007. Effect of variable and subject sampling on recovery of weak factors in CFA. *Methodology* 3, 2 (2007), 67–80. https://doi.org/10.1027/1614-2241.3.2.67