Does behaviour match user typologies? An exploratory cluster analysis of behavioural data from a gamified fitness platform

Jeanine Krath[1], Adam Palmquist[2], Izabella Jedel[3], Isak Barbopoulos[4], Miralem Helmefalk[5] and Robin Isfold Munkvold[3]

[1] *University of Koblenz-Landau, Universitaetsstrasse 1, Koblenz, 56070, Germany*
[2] *University of Gothenburg, Forskningsgången 6, Gothenburg, 417 56, Sweden*
[3] *Nord University, Universitetsalléen 11, Bodø, 8026, Norway*
[4] *Insert Coin, Vasagatan 33, Gothenburg, 411 37, Sweden*
[5] *Linnaeus University, Universitetsplatsen 1, Kalmar, 392 31, Sweden*

Abstract

A promising solution to increase user engagement in gamified applications is tailored gamification design. However, current personalisation relies primarily on user types identified through self-reporting rather than actual behaviour. As a novel approach, the present study used an exploratory machine learning analysis to identify seven clusters of users in a gamified fitness application based on their behavioural data (N = 19,576). The clusters were then conceptually compared to common user typologies in gamification, identifying possible relationships between behavioural user clusters and user types motivated by achievement, sociability, and extrinsic incentives. The findings shed light on nuanced behaviour patterns of user types in the fitness context and how knowing these patterns can inform the way in which tailored gamification could be implemented to meet the needs of specific types. Thereby, they contribute to the discussion on utilising behavioural data and user typologies for tailored gamification design.

Keywords

Cluster analysis, user types, tailored gamification design, personalisation, fitness, exploratory machine learning, *k*-means clustering

1. Introduction

Gamification, the application of game elements in a non-game context [1], has been researched within several fields to increase user engagement and motivation [2]. One of the most popular applications of gamification is using game elements in fitness applications [3,4]. However, mixed outcomes have let gamification research both in general [5] and in the fitness context [6,7] question the applicability of universal design and increasingly focus on individual differences in how gamification is perceived and used to guide personalized gamification design [5]. Instead of executing a one-size-fits-all design, the prospect of personalisation and adaptivity affords a design that can be automatically informed, rearranged, and redeployed based on users' actions and reactions [5]. Therefore, personalising gamified fitness applications might present a solution to create improved experiences and simultaneously provide users with a wider array of features that may be of particular interest.

6th International GamiFIN Conference 2022 (GamiFIN 2022), April 26-29, 2022, Finland
EMAIL: jkrath@uni-koblenz.de (A. 1); adam.palmquist@ait.gu.se (A. 2); izabella.jedel@hotmail.com (A. 3); isak@insertcoin.se (A. 4); miralem.helmefalk@lnu.se (A. 5); robin.munkvold@nord.no (A. 6)
ORCID: 0000-0003-4996-1147 (A. 1); 0000-0003-0943-6022 (A. 2); 0000-0001-9212-3259 (A. 3); 0000-0003-2485-0184 (A. 4); 0000-0003-2924-2874 (A. 5); 0000-0002-2524-279X (A. 6)

Previous research on tailored gamification has explored personalised gamification design based on demographic data such as age and gender, as well as personality [5,8]. However, to date, the most widely used approach to personalising gamification interventions has been to classify users based on their needs and motivations via user typologies [5]. A variety of user typologies have already been used in research on tailored gamification design [5,8–10], the most popular among them being Bartle's player types [11], the gamification user types HEXAD [12], the BrainHex typology [13] and Yee's motivations to play MMORPGs [14]. Specifically, Yee's motivations [15] and HEXAD types have been successfully used to personalise gamification for fitness and health applications [6,7].

A particular limitation of current gamification design based on these predefined user typologies is that the design process mainly relies on questionnaires to determine user types based on self-report rather than actual behaviour, which is particularly challenging as user types can change over time [16] and people tend to respond in self-reports in a socially desirable way that does not necessarily reflect their actual behaviour [17]. In addition, users may be classified as multiple or hybrid user types depending on the context [18]. Therefore, further research is needed on how technologies such as machine learning [16] could help identify different user types based on their behaviour and dynamically adapt the gamification system accordingly [16,18].

The present study addresses this gap by analysing user behaviour in a gamified fitness application through a machine learning approach and discussing the relationships between identified clusters of users based on their actions and common user typologies. The contribution of this work is thus twofold:

First, a *k*-means cluster analysis is conducted to identify distinct clusters of users based on their actions in a gamified fitness platform, drawing on a large dataset (n = 19,576) of behavioural data with over 1 million events recorded in 49 weeks.

Second, by exploring the extent to which the identified clusters can be mapped to common user typologies, we contribute to the ongoing discussion of user typologies in terms of tailored, personalised, and adaptive gamification. In summary, the study aims to answer the following research question:

RQ1: Which clusters of users can be identified using exploratory clustering techniques on behaviour data from a gamified fitness platform?

2. Background

2.1. User typologies in gamification

Personalisation of gamification design has recently gained tremendous importance in gamification research [19]. Within this stream, researchers have proposed a variety of classifications or typologies of users [5,8–10] and explored their preferences for game elements [20,21] to inform tailored gamification design.

One of the first typologies was Bartle's typology of players in multi-user dungeons [11]. He distinguished between *Achievers* (who focus on earning points and levelling up), *Explorers* (who enjoy discovering interesting features and exploring the system), *Socializers* (who value relationships with other players), and *Killers* (who like disrupting the experience of others) [11].

Building on Bartle's findings, Yee [14] sought to identify the underlying motivations of users of MMORPGs and found ten motivations categorised into the three components of *Achievement* (including advancement, mechanics and competition), *Social* (including socialising, relationships and teamwork) and *Immersion* (including discovery, role-playing, customisation and escapism) [14].

The BrainHex model by Nacke et al. aimed to combine findings from previous models with neurobiological insights [13]. As a result, they described seven types of players, namely *Seekers* (who are curious and eager to explore), *Survivors* (who enjoy intense fright experiences), *Daredevils* (who enjoy thrilling and risky experiences), *Masterminds* (who like to overcome problems and develop strategies), *Conquerors* (who derive satisfaction from defeating others), *Socializers* (who like to interact with others) and *Achievers* (who are goal-oriented and motivated by long-term success) [13].

While the previous typologies were developed in the context of games, Marczewski designed and developed the HEXAD typology, which builds on extrinsic motivation and the three basic psychological needs [22] as a model specifically for use in gamification [12]. The model consists of six user types: *Philanthropists* (driven by purpose and altruism), *Socialisers* (driven by social relations and interaction with others), *Achievers* (striving for self-improvement through challenges and proficiency), *Players* (striving for external rewards), *Free Spirits* (driven by autonomy and exploration), and *Disruptors*

Table 1
Prevalent user typologies in gamification research and their conceptual relations (based on [8,9,13])

Concept ([9])	Bartle ([11])	HEXAD ([12])	BrainHex ([13])	Yee ([14])
Achievement	Achiever	Achiever	Achiever / Mastermind	Achievement
Exploration	Explorer	Free Spirit	Seeker	Immersion
Sociability	Socializer	Socializer / Philanthropist	Socializer	Social
Domination	Killer	Disruptor	Conqueror	-
Gaming intensity and skill	-	-	Survivor / Daredevil	-
-	-	Player	-	-

(driven by challenging the status quo and participating in disruptive alterations) [12].

All of these typologies share common concepts that are reflected in certain types, such as achievement, exploration or sociability [9], and several researchers have attempted to relate the different user types to each other [8,9,13] (see Table 1 for an overview). For example, the concept of achievement is prevalent in each of the four typologies, while the HEXAD model adds the Player as a type that is extrinsically rather than intrinsically driven [6]. In contrast, the BrainHex model describes *Survivors* and *Daredevils* motivated by intense gaming experiences absent in other typologies [13].

2.2. Personalisation in gamified fitness applications

Health and fitness is the second largest area of research in gamification [2]. Previous studies have shown that gamification in fitness tracking apps can successfully promote physical activity and bodyweight reduction [23]. In the fitness context, personalisation of gamification involves real-time adjustment of difficulty based on physiological parameters such as heart rate [24] or acceleration [25]. However, their usefulness for a thoroughly tailored gamification design that modifies different aspects of gamification with appropriate solutions for each user [26] is limited. Therefore, other studies focusing on personalising gamified fitness applications drew on psychological determinants such as motivations and user typologies. One of the first studies on individual differences in gamified fitness applications was that of Brauner et al. [15], who observed that users' motivations to play [14] significantly influenced performance. Later, Kappen et al. examined various exercise motivations in different age groups and found distinct preferences for intrinsic, extrinsic, and feedback elements [27]. Recently, both Altmeyer et al. [6] and Zhao et al. [7] used the HEXAD typology to personalise gamified fitness apps and found that it led to more positive affective experiences [6], motivation, and satisfaction [7] than a one-size-fits-all approach.

These previous approaches relied on self-report to obtain information about users' motivations and needs, categorised into user types and assumed to influence behaviour. However, further research is warranted on user types that can be identified by analysing actual behaviour in gamified systems rather than relying on self-reported motivations [16], especially since user types can manifest in hybrid or multiple forms depending on the context [18].

A promising approach to identifying different types of users is to use unsupervised machine learning techniques [16] to cluster users based on their recorded behaviours in the system. A recent review has shown that automatically adapting gamification design using machine learning is gaining momentum [28]. It has been suggested that artificial intelligence and machine learning are among the most promising emerging areas in gamification research [29]. Previous studies have applied clustering techniques in education to identify different types of students [30,31]. However, as far as the authors know, such techniques have not yet been applied to gamified fitness applications.

2.3. AUTOMATON project

The current work is based on a university and industry initiative started as an applied artificial intelligence (A.I.) project between the University of Skövde and Insert Coin. The long-term project goal is to design and develop a system of machine learning models that are capable of independently identifying user clusters and their behaviour patterns (as stage 1) and then make personalised suggestions, as well as apply and/or adjust the gamification balance to better fit the users in each segment (as stage 2). The expected project outcome is an adaptive gamified fitness platform based on predicted user preferences toward tailored gamification experiences.

3. Method

3.1. Materials

The cloud-based fitness platform in the study is an iOS/Android application. The platform's principal purpose is to function as a marketplace between individuals interested in fitness and exercise on the one hand (hereinafter users) and various fitness centres and fitness coaches on the other. The platform provides each user with a workout diary to log various physical activities (e.g., cycling, weightlifting, running), log and track their weight, and view other metrics indicating their overall progression. The platform also includes social features, such as reactions, adding and sending messages to friends, posting workouts for others to see, or browse, like or commenting on friends' workout feeds, as well as connecting and interacting with other users or fitness coaches.

The gamification design was focused on the workout diary due to its central position in the platform ecosystem, aiming for a gameful ambience in the whole platform. In order to create this ambience, several *motivational affordances* [2] were used, associated with different features in the platform ecosystem (Table 2), exemplified in Figure 1.

3.2. Procedure

The platform has an event-based architecture, meaning that user-generated events, such as reaching training goals associated with milestones (e.g., 30 mins of activity, lifting a total of 3 tonnes of weight in a workout), recording a new exercise (e.g., power walking or weightlifting), viewing planned activities or responding to a fitness coach trial were logged for analysis. Therefore, the dataset for the cluster analysis consisted of 1,116,126 events recorded on the fitness platform originating from 19,576 unique users collected over 49 weeks.

In order to cluster the user actions, the 1,116,126 events first had to be parsed (i.e., variables and value labels had to be extracted from the event meta-data) and aggregated into a list of events associated with each cluster. As some event types consisted of both predefined activities (e.g., "power walk") and free-text entries (e.g., "I went running"), the total number of unique categorical values exceeded 1,600. Because of the large number of categorical values, we treated all categorical values as text entries and used techniques suitable for this kind of data [31].

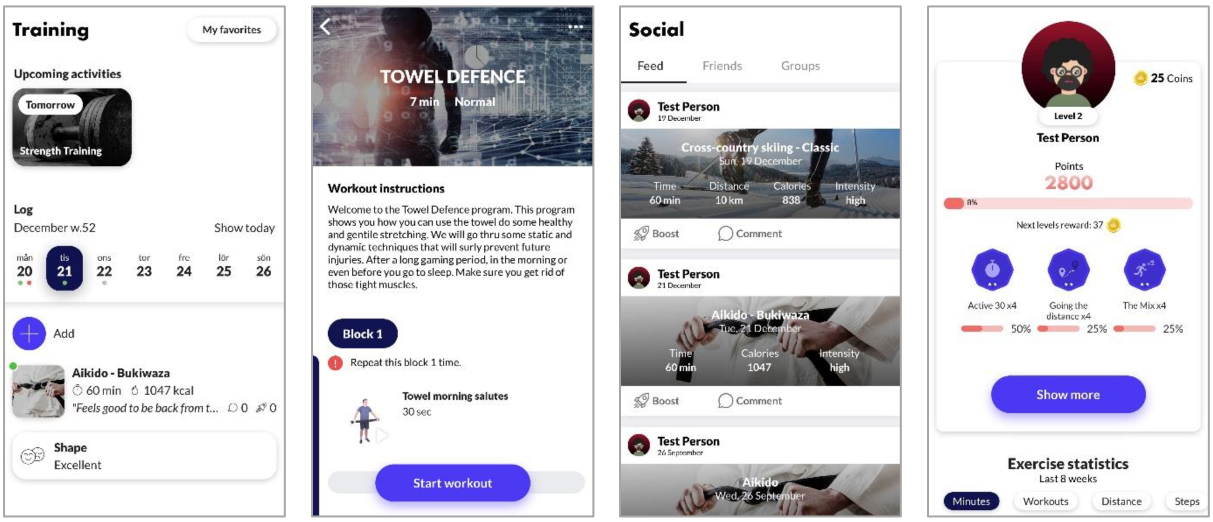

Figure 1: Screenshots of the fitness platform showing the overview of planned and logged activities, a challenge with adaptive difficulty, the social feed with peer rating features and the personal profile with points, virtual coins as currency for real-world rewards, level, milestones and statistics

Table 2
Ecosystem's motivational affordances categorised after Koivisto & Hamari [2]

Affordance	Category
Experience points	Progression
Performance/ Progress stats	Achievement/ Progression
Milestones	Achievement
Levels	Achievement
Team Leaderboards	Achievement/Social
Challenges/Competitions	Achievement/Social
Peer rating	Social
Gameful narrative	Immersion
Adaptive difficulty	Miscellaneous
Onboarding	Miscellaneous
Real-world reward	Miscellaneous
Reminders	Miscellaneous
Virtual currency	Miscellaneous

First, each categorical value was split into separate words. The words were then vectorised with the Term Frequency–Inverse Document Frequency (TF-IDF) statistic[2], a common way to prepare text data for clustering in the field of natural language processing [32]. In order to identify clusters, the vectorised events were then used as input data in a *k*-means cluster analysis [33], conducted using the sci-kit-learn library in Python [32]. The number of clusters (k) to extract is often determined by the elbow method [33]. However, in this case, the ultimate aim of the analysis was to optimise the clusters specifically for use in prediction models [34]. Therefore, a different approach was taken, in which each cluster model (varying between 1 and 12 extracted clusters) was evaluated in terms of how well the overall model could predict what event was most likely to be recorded by users of a specific cluster during their next active day (hereafter the target day), using a Sequential Long Short-Term Memory model. Only users who had recorded events on at least seven unique days were included in the training sample, which brought the sample to a total of 9,667 users. Of these, 60% were used when training the models, and the remaining 40% when testing the trained models and calculating the prediction scores. The clusters were labelled based on *their most frequently recorded events* (in absolute terms), *their most frequently recorded events offset by the frequency of the event across all clusters* (TF-IDF) and *descriptive statistics* (e.g., size of the cluster relative to the whole sample, proportion of total events, and conversion events, i.e., bought subscriptions, sent by users in the cluster).

The clustering procedure and labelling were planned and conducted by author 5, who was not involved with the current research project at the time and therefore did not conduct the analysis or set the labels with the purpose of relating them to existing user typologies in gamified systems.

4. Results

4.1. Cluster model evaluation

The model with the highest percentage of total correct predictions across all its clusters was selected as the final model. As a benchmark, the most common event occurred on ~54% of the target days. The "baseline" model (where all users belonged to the same cluster) achieved a prediction score of 55%. However, the score increased to 68.5% in the cluster model with two clusters, and each subsequent model improved this score further until a peak was reached in the model with 7 clusters, achieving 76% correct predictions. Therefore, the model with 7 clusters was selected.

4.2. Identified clusters

Once the final model had been selected, that model was used to assign the entire sample of 19,576 users to one of the seven identified clusters. The clusters were then compared in terms of most frequent events, most frequent events offset by event frequency across all clusters (TF-IDF), and various descriptive statistics (e.g., size and proportion of events sent by each cluster). Unless otherwise stated, all event frequencies are reported using the TF-IDF statistic [32], not the absolute frequency, as this is generally more useful to differentiate the clusters.

Labels were set accordingly: *Generalists* (7.5% of users, 15.3% of events, 4.0% of conversions[3]), a cluster characterized by recording a relatively even spread between physical activities (TF-IDF: power walk: 12.1% of their total events, lift weights: 8.5%), scheduled

[2] TF-IDF is especially useful for clustering, as the statistic will increase proportionally to the number of times an event is recorded by a user, offset by how many other users have recorded the same event, thereby more effectively differentiating users.

[3] Triggered by paying for subscriptions.

activities (TF-IDF: perform planned activity: 19.8%, view planned activity: 5.0%, which can both be done in the overview of planned and logged activities shown in Figure 1) and attaining achievements (TF-IDF: active 30 minutes: 16.5%, depicted on the personal profile in Figure 1); *Socializers* (6.5% of users, 25.1% of events, 5.7% of conversions), a cluster with a relatively high degree of social activities (TF-IDF: like someone's workout as depicted in the social feed in Figure 1: 15.9% of their events; add friend: 3.4%); *Achievement hunters* (4.9% of users, 6.5% of events, 2.6% of conversions), a cluster characterized by a attaining a high degree of achievements during their workouts (TF-IDF: 30 active minutes: 18.2%, distance 3km: 8.4%); *Organizers* (9.7% of users, 8.2% of events, 22.8% of conversions), a cluster whose most frequent events involved viewing or planning scheduled activities (TF-IDF: view planned activity: 26.0%, perform planned activity: 10.3%); *Heavy lifters* (14.8% of users, 6.7% of events, 21.9% of conversions), a cluster which prioritized weightlifting (TF-IDF: 23.4%) above most other activities, and the only cluster where the achievement for lifting a total of 3000 kg during a workout appeared in the top 5; *Weight watchers* (the most populous cluster by far, at 47.5% of users, 20.6% of events, 37.1% of conversions), a cluster where the users focus on tracking their weight progress more frequently than other users (TF-IDF: 12.8%); and finally *Third-party app. users* (9.1% of users, 17.5% of events, 5.9% of conversions), denote a cluster characterized by a high frequency of events related to third-party devices and apps (41.9% of total events fall in this category according to TF-IDF). The clusters and their most frequent events are shown in Table 3.

Table 3
The clusters with their 5 most frequent events according to TF-IDF and their percentage of the clusters' total events according to TF-IDF and absolute frequency (CF)

Cluster	Top 5 events	TF-IDF	CF
Generalists	Perform planned activity	19.8%	24.9%
n = 1,472	Achievement: 30 active min[4]	16.5%	21.2%
(7,5%)	Activity: Power walk	12.1%	15.1%
	Activity: Lift weights	8.5%	10.5%
	View planned activities	5.0%	4.4%
Socialisers	Like someone's workout	15.9%	30.6%
n = 1,279	Achievement: 30 active min.	7.8%	8.5%
(6,5%)	Add friend	3.4%	2.0%
	Third-party application[5]	3.2%	2.6%
	Activity: Lift weights	2.7%	1.9%
Achievement hunters	Achievement: 30 active min.	18.2%	29.7%
n = 960	Activity: Walking	12.3%	18.4%
(4,9%)	Achievement: Distance 3km	8.4%	10.8%
	Activity: Lift weights	4.4%	5.5%
	View planned activities	2.9%	2.7%
Organizers	View planned activities	26.0%	34.7%
n = 1,901	Perform planned activity	10.3%	13.1%
(9,7%)	Activity: Lift weights	7.9%	6.3%
	Achievement: 30 active min.	7.1%	10.3%
	Activity: Running	4.8%	3.4%
Heavy lifters	Activity: Lift weights	23.4%	29.9%
n = 2,894	Perform planned activity	20.5%	22.8%
(14,8%)	Achievement: 30 active min.	13.7%	20.4%
	View planned activities	5.9%	5.3%
	Achievement: Lift 3000kg total	3.4%	3.4%
Weight watchers	View weight progress	12.8%	2.4%
n = 9,288	Achievement: 30 active min.	9.2%	26.3%
(47,5%)	Third-party application	5.1%	1.2%
	Activity: Lift weights	3.7%	5.5%
	Achievement: Distance 3km	3.4%	7.7%
Third-party app. users	Third-party application	41.9%	60.9%
n = 1,782	Achievement: 30 active min.	5.8%	7.5%
(9,1%)	Activity: Lift weights	3.8%	3.6%
	View planned activities	3.7%	2.7%
	Perform planned activity	3.4%	3.4%

5. Discussion

The results of our exploratory cluster analysis led to the identification of seven distinct clusters of gamified fitness platform users based on their behaviours. By applying k-means clustering, a machine learning technique, to identify these clusters based on over one million events recorded by 19,576 users, the results extend previous research [6,7,15,27] that mainly relied on self-report tools as a basis for personalising gamified systems rather than actual behaviour [16,18]. In order to elaborate on the contribution of this study to the scientific debate on personalisation of gamified systems, which is currently oriented towards needs- and motivation-based user typologies [5,8–10], it is important to discuss how the exploratively identified clusters conceptually

[4] "Achievement" events are generated when a user attains one of the achievements in the gamification design.

[5] The "Third-party application" event refers to events generated by other (third-party) devices/apps (like smart watches, smart scales, etc.). We do not have any information about which apps users interacted with or how they used them.

relate to these typologies, in order to examine how certain needs can manifest themselves in behaviours and how behaviour-based clusters might contribute to effective tailored gamification design, given the high predictability of future user actions based on the identified clusters. The conceptual discussion (see Table 4 for an overview) is based on the most distinguishing events of each cluster (Table 3) and the described characteristics of the user types in the typologies.

From a behavioural perspective, the *Socialiser* cluster differs from the others in the prevalence of social events. The most common event is Like someone's workout (16.7%), probably to encourage others after a workout and show social appreciation for their performance. User types driven by sociability [9], i.e., *Socializers* in Bartle's typology [11], HEXAD [12] and BrainHex [13] and *Social* motivation in Yee's motivations [14], are motivated by relatedness, social connections and interaction [12] and like social networks and social status functions [20,21]. Thus, we argue for a first conceptual relationship between the *Socialiser* cluster and these socially-driven user types.

Achievement hunters show a comparatively dominant number of events related to attaining in-app achievements. They over proportionally earned achievements for 30 active minutes (18.2%) and 3km distance (8.4%), with the most common activity being walking (12.3%), which might be ideal to get these achievements. *Players*, described as users who are motivated by extrinsic affirmations of their achievements [12], are keen on receiving virtual or real-world rewards and incentives for their activities [20,21]. Therefore, another link could be seen between the *Achievement hunters* cluster and the HEXAD *Player*, whereby we can only relate to the HEXAD because extrinsic motivations are not reflected in other typologies [9,11,12]. It should be noted, however, that the causal relationship cannot be clearly determined (i.e., it could also be that they got to the achievements because they mostly preferred walking rather than vice versa).

Next, we see three clusters of users that best relate to the concept of achievement through self-improvement [9]. Associated types are described as motivated by levelling up [11], overcoming challenges [12], advancing and competition orientation [14] and goal orientation [13] and thus report liking features such as progress monitoring and levels as well as challenges [20,21]. In the clusters of *Organisers*, *Heavy lifters*, and *Generalists*, we can observe different constellations of behaviour related to these achievement needs. Comparing the top five events of *Organizers'* and *Heavy lifters* shows that they record similar events, but with different frequencies. The most common events of *Organisers* are View planned activity (26.0%) and Perform planned activity (10.3%), which indicates a desire to plan and track activities and progress towards goals, reflecting the long-term orientation postulated in the BrainHex typology [13]. In turn, the most frequent event of the *Heavy lifters* is Lift weights (23.4%), followed by events relating to planned activities, and it was the only cluster with the achievement Lift 3000kg among their top 5 events (3.4%), suggesting that they might be motivated by the challenge of strenuous physical activities and self-improvement through planning and mastery, which fits with the achievement-oriented user types reflected in the HEXAD typology [12] and Yee's motivations [14]. In comparison to *Organisers, Heavy lifters* represent a more action-oriented cluster, as they seem to be focused on mastering a specific form of physical activity (weightlifting). At the same time, the former performs a more diverse set of physical activities and is more characterised by the planning itself. *Generalists* are distinguished by a more even spread of events between planning (Perform planned activity: 19.8%, View planned activities: 5.0%), attaining achievements (30 active minutes: 16.5%) and performing physical activities (Power walk: 12.1%, Lift weights: 8.5%). Like *Organisers* and *Heavy lifters*, they also seem motivated by goal-setting and challenge but less focused on either aspect. In contrast to these more action- and planning-oriented clusters, *Generalists* seem to combine the short-term challenge and advancement orientation [12,14] with the long-term goal orientation [13].

The *Weight watchers* cluster is fascinating because it is the most populous cluster (47.5% of users) and, at the same time, is more difficult to relate to existing user typologies. While it could be argued that monitoring progress (View weight progress: 12.8% of events) is related to an achievement orientation [14], the cluster lacks events indicating goal-oriented planning [13], which argues against a link to existing user types.

The *Third party app. users* cluster is characterised by events that have been recorded via appliances such as smartwatches and smart scales. However, as we lack information on what events were recorded in them, we cannot conclude the specific behaviours of this cluster and thus cannot relate them to user types.

Table 4
Clusters postulated to be conceptually related to different user types in gamification research

Identified Cluster	Bartle ([11])	HEXAD ([12])	BrainHex ([13])	Yee ([14])
Organizers / Heavy lifters / Generalists	Achiever	Achiever	Achiever / Mastermind	Achievement
-	Explorer	Free Spirit	Seeker	Immersion
Socializers	Socializer	Socializer / Philanthropist	Socializer	Social
-	Killer	Disruptor	Conqueror	-
-	-	-	Survivor / Daredevil	-
Achievement hunters	-	Player	-	-
Weight watchers	-	-	-	-
Third party app. users	-	-	-	-

There are other types from existing user typologies, namely those driven by exploration [9] (*Explorer*, *Free Spirit*, *Seeker*, *Immersion*), and domination [9] (*Killer*, *Disruptor*, *Conqueror*), *Philanthropists* [12] and *Survivors/Daredevils* [13] that could not be identified in the clusters. This is likely because the gamified fitness platform does not offer specific features corresponding to these user types, so no behavioural clusters emerged concerning these needs. For *Philanthropists*, there was no specific altruistic action [12] that users could perform, apart from liking others' workouts, which we deem to be more related to the social aspect. Furthermore, there was no specific way to express autonomy and exploration [11–14], nor to dominate other players [13] or disrupt their experience [11,12]. Concerning *Survivors* and *Daredevils*, the gamified fitness platform as a smartphone app might not have provided intense and thrilling experiences.

The preceding discussion yields interesting contributions to the debate on user typologies and behavioural data for personalised gamification design. First, several distinct clusters of users driven by achievement could be identified, likely due to the fitness platforms nature and possibly enabled by the wide variety of motivational affordances in the Achievement and Progression category (Table 2). In examining the clusters, we observed a general, action-oriented, and planning-oriented expression of achievement motivation types, suggesting that the actual behaviour patterns of these user types may be more nuanced than theory would suggest. Furthermore, since the cluster model that distinguished these three achievement-oriented clusters outperformed simpler models with fewer clusters, personalisation based on these different kinds of achievement-oriented behaviours may be more effective than one that groups them under a single type. This insight is fascinating and calls for further research into the possible multi-dimensionality of user types.

Second, we illustrate the value of behavioural data for personalised gamification design. The *Weight watchers* could not be clearly linked to existing user typologies, yet it is the most populous cluster and exhibits distinct behavioural patterns from other clusters. This result does not allow conclusions to be drawn about general user typologies, as the cluster probably results from a particular type of monitoring that is likely to be relevant only to fitness apps and similar platforms rather than a more general motivation or need. However, it does illustrate the value of basing personalisation not only on user types but also on actual behaviours. For example, progress statistics and milestones may generally appeal to users with a need for achievement [2], while for specific users, as the *Weight watchers* cluster, progress and milestones may be primarily of interest if they related specifically to their weight. Likewise, users in the *Heavy lifters* cluster may prefer them related to their weightlifting goals. Thus, by clustering users based on behavioural data, the emergent behavioural patterns can inform *how* a given gamification mechanic should be implemented to meet their psychological needs.

6. Limitations and outlook

The cluster of *Third-party app. users* was characterised by a high frequency of events recorded via third-party applications (41.9%). Thus, a limitation of the present research is that we lacked information about what those applications were and how they were used, so it is difficult to accurately describe this cluster and discuss to what extent it may relate to other clusters and user typologies. Another limitation of the study is that we did not have self-reported data from questionnaires on user typologies, so the identified clusters could not be directly correlated, but only conceptually related based on their most characteristic behaviours (as calculated using the TF-IDF statistic), which are merely hypotheses to be explored in further studies. Therefore, future research is encouraged to extend the behavioural cluster analysis with data on user types and correlate the self-reported responses with observed behaviours to gain a more nuanced and comprehensive understanding of the relationship between needs-based user typologies and actual behaviour. As the gamification design of the application was not a priori based on matching gamification mechanics with different psychological needs, the results of the behavioural cluster analysis, as well as the relationship to needs-based user typologies, may be different for other applications with different gamification features or other target groups and domains (e.g., gamification in sustainability or education). For example, the gamified fitness platform used in this study did not include features related to exploration and domination, which is likely the reason why the identified behavioural clusters could not be related to user types associated with these motivations. In order to understand the generalisability of the results of our cluster analysis, further research should be conducted using similar methods in gamified applications in different contexts. Finally, since need-based user types have been shown to change over time [16], it would be interesting to further study the behavioural clusters' temporal context and explore whether they are stable over time or transient.

7. References

[1] S. Deterding, D. Dixon, R. Khaled, L.E. Nacke, Gamification: Toward a Definition, in: CHI 2011, 2010: pp. 1–4. doi:10.1007/978-3-642-13959-8_1.

[2] J. Koivisto, J. Hamari, The rise of motivational information systems: A review of gamification research., Int. J. Inf. Manage. 45 (2019) 191–210. doi:10.0.3.248/j.ijinfomgt.2018.10.013.

[3] Y. Wang, W.B. Collins, Systematic evaluation of mobile fitness apps: Apps as the Tutor, Recorder, Game Companion, and Cheerleader, Telemat. Informatics. 59 (2021) 101552. doi:10.1016/j.tele.2020.101552.

[4] J. Koivisto, J. Hamari, Gamification of physical activity: A systematic literature review of comparison studies, in: Proc. 3rd Int. GamiFIN Conf., 2019: pp. 106–117.

[5] A.C.T. Klock, I. Gasparini, M.S. Pimenta, J. Hamari, Tailored gamification: A review of literature, Int. J. Hum. Comput. Stud. 144 (2020) 102495. doi:10.1016/j.ijhcs.2020.102495.

[6] M. Altmeyer, P. Lessel, S. Jantwal, L. Muller, F. Daiber, A. Krüger, Potential and effects of personalizing gameful fitness applications using behavior change intentions and Hexad user types, User Model. User-Adapt. Interact. 31 (2021) 675–712. doi:10.1007/s11257-021-09288-6.

[7] Z. Zhao, A. Arya, R. Orji, G. Chan, Effects of a personalized fitness recommender system using gamification and continuous player modeling: System design and long-term validation study, JMIR Serious Games. 8 (2020) 1–27. doi:10.2196/19968.

[8] F. De Vette, M. Tabak, M. Dekker - van Weering, M. Vollenbroek-Hutten, Engaging Elderly People in Telemedicine Through Gamification, JMIR Serious Games. 3 (2015) e9. doi:10.2196/games.4561.

[9] J. Hamari, J. Tuunanen, Player Types: A Meta-synthesis, Trans. Digit. Games Res. Assoc. 1 (2014) 29–53. doi:10.26503/todigra.v1i2.13.

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

[11] R. Bartle, Hearts, clubs, diamonds, spades: Players who suit MUDs, J. MUD Res. 1 (1996) 19.

[12] A. Marczewski, User Types, in: Even Ninja Monkeys Like to Play, CreateSpace Independent Publishing, 2015: pp. 65–80.

[13] L.E. Nacke, C. Bateman, R.L. Mandryk, BrainHex: A neurobiological gamer typology survey, Entertain. Comput. 5 (2014) 55–62. doi:10.1016/j.entcom.2013.06.002.

[14] N. Yee, Motivations for Play in Online Games, CyberPsychology Behav. 9 (2006) 772–775. doi:10.1089/cpb.2006.9.772.
[15] P. Brauner, A. Calero Valdez, U. Schroeder, M. Ziefle, Increase Physical Fitness and Create Health Awareness through Exergames and Gamification, in: Hum. Factors Comput. Informatics, 2013: pp. 349–362. doi:10.1007/978-3-642-39062-3_22.
[16] A.C.G. Santos, W. Oliveira, J. Hamari, S. Isotani, Do people's user types change over time? An exploratory study, in: Proc. 5th Int. GamiFIN Conf., 2021: pp. 90–99.
[17] D.M. Randall, M.F. Fernandes, The social desirability response bias in ethics research, J. Bus. Ethics. 10 (1991) 805–817. doi:10.1007/BF00383696.
[18] S. Hallifax, A. Serna, J.-C. Marty, G. Lavoué, E. Lavoué, Factors to Consider for Tailored Gamification, in: Proc. Annu. Symp. Comput. Interact. Play, ACM, New York, NY, USA, 2019: pp. 559–572. doi:10.1145/3311350.3347167.
[19] S. Schöbel, M. Schmidt-Kraepelin, A. Janson, A. Sunyaev, Adaptive and Personalized Gamification Designs: Call for Action and Future Research, AIS Trans. Human-Computer Interact. 13 (2021) 479–494. doi:10.17705/1thci.00158.
[20] J. Krath, H.F.O. von Korflesch, Player Types and Game Element Preferences: Investigating the Relationship with the Gamification User Types HEXAD Scale, in: HCI Int. 2021, 2021: pp. 219–238. doi:10.1007/978-3-030-77277-2_18.
[21] G.F. Tondello, A. Mora, L.E. Nacke, Elements of Gameful Design Emerging from User Preferences, in: Proc. Annu. Symp. Comput. Interact. Play, ACM, New York, NY, USA, 2017: pp. 129–142. doi:10.1145/3116595.3116627.
[22] R.M. Ryan, E.L. Deci, Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions, Contemp. Educ. Psychol. 25 (2000) 54–67. doi:10.1006/ceps.1999.1020.
[23] C. González-González, N.G. Río, V. Navarro-Adelantado, Exploring the Benefits of Using Gamification and Videogames for Physical Exercise: a Review of State of Art, Int. J. Interact. Multimed. Artif. Intell. 5 (2018) 46. doi:10.9781/ijimai.2018.03.005.
[24] A. Schwarz, G. Cardon, S. Chastin, J. Stragier, L. De Marez, A. DeSmet, Does Dynamic Tailoring of A Narrative-Driven Exergame Result in Higher User Engagement among Adolescents? Results from A Cluster-Randomized Controlled Trial, Int. J. Environ. Res. Public Health. 18 (2021) 7444. doi:10.3390/ijerph18147444.
[25] S.A. Pérez, A.M. Díaz, D.M. López, Personalized Tracking of Physical Activity in Children Using a Wearable Heart Rate Monitor, Int. J. Environ. Res. Public Health. 17 (2020) 5895. doi:10.3390/ijerph17165895.
[26] V.M. García-Barrios, F. Mödritscher, C. Gütl, Personalisation versus Adaptation? A User-centred Model Approach and its Application, in: Proc. Int. Conf. Knowl. Manag., 2005: pp. 120–127. doi:10.1.1.161.7376.
[27] D.L. Kappen, P. Mirza-Babaei, L.E. Nacke, Gamification through the application of motivational affordances for physical activity technology, in: Proc. Annu. Symp. Comput. Interact. Play, 2017: pp. 5–18. doi:10.1145/3116595.3116604.
[28] A. Khakpour, R. Colomo-Palacios, Convergence of Gamification and Machine Learning: A Systematic Literature Review, Technol. Knowl. Learn. 26 (2021) 597–636. doi:10.1007/s10758-020-09456-4.
[29] A. Mazarakis, Gamification Reloaded, I-Com. 20 (2021) 279–294. doi:10.1515/icom-2021-0025.
[30] D. Codish, E. Rabin, G. Ravid, User behavior pattern detection in unstructured processes – a learning management system case study, Interact. Learn. Environ. 27 (2019) 699–725. doi:10.1080/10494820.2019.1610456.
[31] A.H. Nabizadeh, J. Jorge, S. Gama, D. Goncalves, How Do Students Behave in a Gamified Course?—A Ten-Year Study, IEEE Access. 9 (2021) 81008–81031. doi:10.1109/ACCESS.2021.3083238.
[32] G. Varoquaux, L. Buitinck, G. Louppe, O. Grisel, F. Pedregosa, A. Mueller, Scikit-learn, GetMobile Mob. Comput. Commun. 19 (2015) 29–33. doi:10.1145/2786984.2786995.
[33] B. Purnima, K. Arvind, EBK-means: A clustering technique based on elbow method and k-means in WSN, Int. J. Comput. Appl. 105 (2014) 17–24.
[34] W. Jenq-Haur, L. Ting-Wei, L. Xiong, W. Long, An LSTM approach to short text sentiment classification with word embeddings, in: 2018 Conf. Comput. Linguist. Speech Process, 2018: pp. 214–223.