



# Potential and effects of personalizing gameful fitness applications using behavior change intentions and Hexad user types

Maximilian Altmeyer<sup>1,2</sup> · Pascal Lessel<sup>1,2</sup> · Subhashini Jantwal<sup>2</sup> · Linda Muller<sup>2</sup> · Florian Daiber<sup>1,2</sup> · Antonio Krüger<sup>1,2</sup>

Received: 16 June 2020 / Accepted in revised form: 6 January 2021  
© The Author(s) 2021

## Abstract

Personalizing gameful applications is essential to account for interpersonal differences in the perception of gameful design elements. Considering that an increasing number of people lead sedentary lifestyles, using personalized gameful applications to encourage physical activity is a particularly relevant domain. In this article, we investigate behavior change intentions and Hexad user types as factors to personalize gameful fitness applications. We first explored the potential of these two factors by analyzing differences in the perceived persuasiveness of gameful design elements using a storyboards-based online study ( $N = 178$ ). Our results show several significant effects regarding both factors and thus support the usefulness of them in explaining perceptual differences. Based on these findings, we implemented “Endless Universe,” a personalized gameful application encouraging physical activity on a treadmill. We used the system in a laboratory study ( $N = 20$ ) to study actual effects of personalization on the users’ performance, enjoyment and affective experiences. While we did not find effects on the immediate performance of users, positive effects on user experience-related measures were found. The results of this study support the relevance of behavior change intentions and Hexad user types for personalizing gameful fitness systems further.

**Keywords** Gamification · Persuasive technology · Behavior change · Personalization · Hexad user types · Stage of change · TTM · Running · Sports · Step count · Physical activity

---

✉ Maximilian Altmeyer  
maximilian.altmeyer@dfki.de

<sup>1</sup> German Research Center for Artificial Intelligence (DFKI), Saarbrücken, Germany

<sup>2</sup> Saarland Informatics Campus, Saarbrücken, Germany

## 1 Introduction

Our daily life is more and more susceptible to physical inactivity caused by an increasing number of people leading sedentary lifestyles (Rajaratnam and Arendt 2001). This lack of physical activity leads to numerous health issues, including cardiovascular diseases, obesity and many other chronic illnesses (Bravata et al. 2007). Therefore, motivating people to lead an active lifestyle is important for public and private health and has been targeted by several interventions in the past (Aldenaini et al. 2020). Often, such interventions employ gameful design elements by using gamification, *the use of game design elements in non-game contexts* (Deterding et al. 2011). Mostly, a “one-size-fits-all” approach (i.e., using a static set of gamification elements) is used (Hamari and Sarsa 2014; Jia et al. 2016; Seaborn and Fels 2015). However, previous research has shown that there are interpersonal differences in the perception of gameful design elements (Tondello et al. 2016), which poses a threat to such static gamification approaches. Consequently, research has been carried out to investigate which factors moderate the perception of gameful design elements or persuasive strategies. For instance, demographic factors such as age (Birk et al. 2017), gender (Orji et al. 2015) or personality traits (Jia et al. 2016) have been shown to play a role in this context.

However, none of these factors is particularly suitable or has been specifically developed personalizing gameful systems and maximizing their motivational impact. To bridge this gap, Marczewski (Marczewski 2015) proposed the Hexad user type model—a model that has been developed to explain user preferences in gameful systems (Orji et al. 2018; Tondello et al. 2017). It consists of six user types, which differ in the degree to which they are driven by autonomy, relatedness and competence, which are core aspects of Self-Determination Theory (Ryan and Deci 2000). Although the Hexad model has been subsequently used successfully in various domains including education (Mora et al. 2018), energy conservation (Kotsopoulos et al. 2018) or alcohol consumption (Orji et al. 2018), the applicability in the fitness context has not been shown, as far as we know.

Moreover, most aforementioned factors (including the Hexad model) are static, i.e., they usually do not change over time. Considering that research has demonstrated that goal completion and motivation is affected by task-related self-efficacy and an individual’s belief that the goal can be achieved (Cham et al. 2019; Locke and Latham 2002), considering dynamic factors to personalize gameful systems encouraging physical activity is important.

Ultimately, previous studies investigating factors for personalization were survey-based, which means that participants did not have the chance to interact with applications but instead rated their perception by imagining how the gameful design elements would look like in a real system. This survey methodology is appropriate to recruit a large amount of participants and investigate the potential of such factors for personalization (Orji et al. 2018). However, it is quintessential to also investigate the actual effect of these factors.

We contribute to the aforementioned research gaps. First, we show the applicability of the Hexad model in the fitness context and replicate previously found

correlations between gameful design elements and Hexad user types in other domains, supporting the usefulness of the Hexad model. Second, we demonstrate that behavior change intentions have an impact on the perception of gameful design elements. This shows the importance of dynamic factors in the context of tailored gameful design for behavior change. While we use a storyboards-based approach to investigate the *potential* of Hexad user types and behavior change intentions ( $N = 178$ ), we contribute to the third aspect by applying our findings in the context of “Endless Universe,” a gameful application encouraging physical activity on a treadmill. In a laboratory experiment ( $N = 20$ ), we show that in general, Endless Universe significantly increased the performance of users, supporting its validity. While we found no immediate effects on performance improvement when personalizing Endless Universe based on Hexad user types or behavior change intentions, improvements on user experience-related measures were found. Our results show that adapting gameful applications to the behavioral intention of users leads to stronger affective experiences. Also, we show that tailoring for Hexad user types has a positive effect on users’ motivation to run, whereas counter-tailoring has detrimental effects. Summing up the findings from both user studies, we demonstrate that Hexad user types and behavior change intentions are important factors for personalizing gameful applications encouraging physical activity.

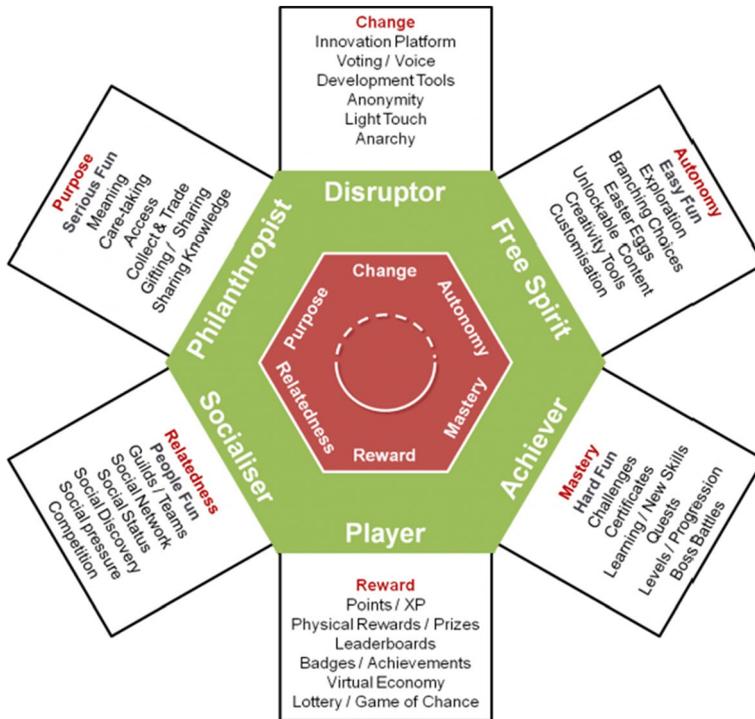
This article is structured as follows: In Sect. 2, we introduce the Hexad user types model and the concept of behavior change intentions, utilizing the “stages of change” theory of the transtheoretical model by Prochaska and Velicer (1997). Next, we present related work and frame our contribution in Sect. 3. Sections 4.1 and 5 explain the storyboards-based approach we followed to investigate interpersonal differences in the perception of gameful design elements in the course of an online study. Sections 4.1 and 5 are based on our previously published work (Altmeyer et al. 2019). In Sect. 6, we describe the design of a personalized gameful application (Fig. 1). This application is used to investigate the effects of personalization in Sect. 7. Both the storyboards-based online study and the laboratory study are discussed in Sect. 8. Finally, we summarize our findings and outline directions for future work in Sect. 9.

## 2 Background

Before presenting relevant literature in the context of encouraging physical activity and personalization of gameful systems, we explain and define the two factors that we are considering in this article.

### 2.1 Hexad user type model

The Hexad user types model (Marczewski 2015) was specifically developed to understand and explain user preferences within gameful systems (Orji et al. 2018; Tondello et al. 2016). It consists of six user types that differ in the degree to which



**Fig. 1** Gamification user types Hexad, taken from Marczewski (2015). Copyright Andrzej Marczewski (CC BY-NC-ND)

they are driven by their needs for autonomy, relatedness and competence as defined by the Self-Determination Theory (SDT) (Ryan and Deci 2000). In HCI research, SDT is widely used to explain motivation and behavior when interacting with technology (Tyack and Mekler 2020). According to SDT, the motivation to engage in a task is located on a spectrum ranging from extrinsic (the task is pursued because of factors outside of the task) to intrinsic (the task is enjoyable on its own) motivation. SDT further posits that a task is more enjoyable (and thus more intrinsically motivating), when three basic psychological human needs are fulfilled: competence, the feeling of acting skillfully and having an effect; autonomy, a feeling of being in control and that actions are self-endorsed; and relatedness, a sense of belonging and a feeling of involvement with others. Based on the type of motivation and on needs satisfaction, the Hexad model establishes the following user types:

- Philanthropists* (“PH”) Are socially minded, like to bear responsibility and share knowledge with other users. They are driven by *purpose*.
- Socialisers* (“SO”) Are also socially minded but are more driven by interacting with other users. Therefore, *relatedness* is their main motivation.

<i>Free Spirits</i> (“FS”)	Are satisfied when acting without external control, with <i>autonomy</i> being most important for them.
<i>Achievers</i> (“AC”)	Are driven by overcoming obstacles and mastering difficult challenges. <i>Competence</i> is most important for them.
<i>Players</i> (“PL”)	Are out for their own benefits and will do their best to earn rewards. Hence, <i>extrinsic rewards</i> are most important for them.
<i>Disruptors</i> (“DI”)	Like to test a system’s boundaries and are driven by triggering <i>change</i> , either positive or negative.

Tondello et al. (2016) developed a questionnaire to assess Hexad user types, and more recently, the authors (Tondello et al. 2018) made slight adjustments to it and showed its reliability and validity. It should be noted that users do not have one specific user type but that the Hexad model is a traits model, which means that users are characterized by their distribution of scores across the six user types (Tondello et al. 2016).

## 2.2 Behavior change intentions

To formalize the intention to change behavior of users, we utilized the “stages of change” concept of the Transtheoretical Model by Prochaska and Velicer (1997). It describes the process of intentional behavior change, stating that behavior change involves progress through five so-called stages of change. These stages are characterized in the following:

<i>Precontemplation</i>	The subject has no intention to take action in the foreseeable future (usually 6 months).
<i>Contemplation</i>	The subject intends to take action within the foreseeable future (6 months).
<i>Preparation</i>	The subject intends to take action in the immediate future (usually 30 days) and has taken some behavioral steps in this direction.
<i>Action</i>	The subject has changed their behavior for less than 6 months.
<i>Maintenance</i>	The subject has changed their behavior for more than 6 months.

When individuals progress through these stages, their motivation becomes more intrinsic as behavioral regulation becomes more self-determined (Mullan and Markland 1997). We expect that this has an effect on the perception of gameful design elements and on aspects related to the user experience and motivation within gameful systems.

## 3 Related work

We contribute to the fields of physical activity encouragement and individualization of gameful applications for behavior change. Therefore, we start by presenting relevant research that has been carried out in the field of gameful applications

encouraging physical activity. Next, we discuss why personalization is essential to gameful systems and which factors have been considered. We conclude the related work section by summarizing the key findings and by framing the contribution of this article.

### 3.1 Encouraging physical activity through gameful design

Encouraging people leading an active lifestyle has been the goal of numerous interventions in the past and is an ongoing research field (Aldenaini et al. 2020; Hamari and Sarsa 2014; Seaborn and Fels 2015). There is a wide spectrum of approaches regarding how to motivate people being more physically active using gameful design (Aldenaini et al. 2020; Hamari et al. 2014; Hamari and Sarsa 2014). For instance, UbiFit Garden (Consolvo et al. 2008a, b), an application showing a virtual garden on participants' mobile phones, has been shown to increase their activity levels. The system uses activity goals and conveys progress through flowers and butterflies growing and appearing. Similarly, goals and progression are used as motivational affordances in a system investigated by Consolvo et al. (2006). The authors present "Houston," a fitness app available in two versions. The "personal" version uses daily step goals and visualizes progression towards these goals for the last seven days. In the "sharing" version, users are additionally able to see progress made towards goals by others. Results demonstrated that participants in the "sharing" version were more likely to reach their daily step goal. Similarly, Zuckerman and Gal-Oz (2014) developed a research prototype called "StepByStep" to motivate people to walk more. In contrast to Consolvo et al. (2006), the study comparing two gamified versions against a non-gamified version revealed that the gamified versions were just as successful as non-gamified one. The authors state that social comparison was effective for some, but not all participants and that interpersonal differences might explain the absence of effects in the gamified conditions. This underlines the need to understand which factors explain such interpersonal differences, to which this article contributes. StepStream (Miller and Mynatt 2014) establishes goals based on the performance of other users. The system uses a social stream on a website, showing achievements when users reach their daily step goals. The user study revealed that the system did not lead to an increase in step counts. As reported by the authors, participants were living in an urban community with low walkability. Thus, their intention to perform physical activity might have been low and social comparison might have been unsuitable to motivate this population effectively.

To better understand user behavior in different group settings within gameful applications, Chen and Pu (2014) investigate the effectiveness of using social collaboration, competition or hybrid settings to encourage physical activity. They developed "HealthyTogether," a smartphone application that pairs users to exercise together. Differing from the findings before, the results show that collaboration and hybrid settings outperformed competition. Similarly, Gui et al. (2017) investigate social comparison strategies in preexisting social networks. Instead of pairing unknown users, the authors analyzed whether existing social peers stimulate an engaging environment motivating physical activity. Their results show that sharing

fitness data with established social networks motivates users to keep tracking their steps and has the potential to improve their social relationships. The impact of social game design elements on walking behavior was also part of “Active2Gether” (Klein et al. 2017), a smartphone application using social comparison, tailored coaching messages and self-monitoring. In a user study by Middelweerd et al. (2020), three conditions were compared. In one condition, the full range of game design elements was used, whereas only self-monitoring and social comparison were used in a second condition. In the third condition, participants were given a commercially available fitness application using self-monitoring only. When comparing both versions of the system against the commercially available application, the effect sizes for active minutes per day were larger in the second condition and smaller in the first condition. However, no significant differences were found between the conditions. As a result, understanding behavioral determinants and studying personalized interventions to increase physical activity is explicitly stated as important future work.

Further investigating the role of social factors in public spaces, Cercos and Mueller (2013) report findings from a public display visualizing each participant’s step count in a graph. It was found that participants started socializing and that the public display led to an increased usage of the pedometers and more motivation. These results are similar to a more recent study by Altmeyer et al. (2018b) who investigated the effect of showing gameful feedback about step counts publicly in addition to showing them in a mobile application. They found that showing each users’ progress toward step goals publicly led to a significant increase in step counts. The acceptance of a public system to encourage stair climbing was investigated by Meyer et al. (2018). They developed the “ActiStairs” system and found that it was successful in increasing awareness for stair climbing. Fish’n’Steps (Lin et al. 2006) links users’ step counts to the growth and emotional state of a virtual fish to encourage them to walk more. In a user study in an office environment, all participants were able to see their personal fish tank, while half of the participants additionally were grouped in teams. Teams have their own fish tanks, in which the virtual fish of all team members are living. These team fish tanks were shown on a public display, thus introducing social comparison. The study revealed that there were no differences in the amount of steps walked between these two conditions. As stated by the authors, this might have been due to the fact that participants had little chance to socialize. Nakajima and Lehdonvirta (2013) investigated virtual ambient paintings which change their appearance based on the amount of physical activity a user performs. In two user studies, no effects could have been found. The authors speculate about the role of behavior change intentions in this context and state that the type of motivational affordance might need to be tailored to the stage of behavior change of a users, which motivates the relevance of our research.

While the aforementioned approaches are build upon static goals, research has demonstrated that designing for dynamic goals, i.e., goals which may change over time, is important in the context of encouraging physical activity. This motivates investigating ways to formalize these dynamic goal adjustments, to which we contribute by analyzing the role of behavior change intentions as a factor for personalization. As such, Niess and Woźniak (2018) emphasize that fitness tracker goals are evolving. To explain these dynamic transition of goals, they define the “Tracker

Goal Evolution Model.” It states that qualitative goals (doing more sports) are built upon internalized user needs, which can be translated into quantitative fitness goals. Similar to this, Epstein et al. (2015) proposed a lived model of personal informatics. The model supports the fact that the motivations, goals and needs, while self-tracking dynamically changes over time. Also, Li et al. (2010) emphasize that fitness tracker users progress through five phases, which pose different challenges to the user. In line with Epstein et al. (2015) and Niess and Woźniak (2018), they state that the motivation of users changes when progressing through these stages. The fact that these stages are based on the Transtheoretical Model of Behavior change supports the relevance of considering behavior change intentions as potential moderators of how certain gameful design elements are perceived.

### 3.2 Personalization of gameful applications

The previous section has demonstrated that interventions aiming at encouraging physical activity lead to a wide spectrum of positive, neutral or even negative outcomes. A recent literature review by Aldenaini et al. (2020) supports this finding. The authors reviewed 170 papers regarding the effectiveness of gameful interventions in encouraging physical activity and found that 49% of them were partially successful or even unsuccessful. Therefore, understanding which factors influence the perception and effectiveness of interventions encouraging physical activity is important.

Jia et al. (2016) investigated the influence of personality traits on the perception of gameful design elements. They used videos of a researcher interacting with gameful design elements, provided textual descriptions of the presented gameful design elements and asked participants to rate their perception in a survey. Their results show that personality traits influence the perception of certain gameful design elements, e.g., the authors found that “extroversion” positively impacts the perception of points and levels. In a follow-up work (Jia et al. 2017), the authors demonstrate that the perception of several ways to represent leaderboards is moderated by personality traits. They use storyboards to explain the different types of leaderboards and ask participants to rate their perceived enjoyment. Besides other results, they found that more extroverted users perceived leaderboards more positively, independent of their ranking. Also, Orji et al. (2017) investigated the role of personality traits to explain the perceived persuasiveness (defined as “an individual’s favorable impressions toward the system” (Drozd et al. 2012)) of 11 persuasive strategies including social comparison, rewards or goal setting. The authors created storyboards explaining each strategy in the context of unhealthy alcohol behavior and found similar effects as Jia et al. (2016). In another work by Orji et al. (2014), the impact of BrainHex (Nacke et al. 2014) player types on the perception of persuasive strategies was investigated and several correlations were found. However, subsequent research revealed severe issues regarding the reliability and validity of BrainHex (Busch et al. 2016b) and the effectiveness of personalizing persuasive systems using BrainHex has been questioned (Busch et al. 2016a). Therefore, BrainHex should not be used for personalization purposes, especially for gameful systems (Hallifax et al. 2019).

In line with Orji et al. (2017), Halko and Kientz (2010) used storyboards explaining persuasive strategies to investigate potential relationships between personality traits and users' perceived enjoyment. The authors focused on the domain of encouraging physical activity by using mobile devices. Their results revealed significant correlations between the factors of the Big-5 personality traits and the perception of the persuasive strategies, further emphasizing the importance of personalization of systems encouraging physical activity.

In addition to personality, research has been carried out to understand age as a potential factor for personalization. As such, Birk et al. (2017) investigated play habits and play preferences among older adults. They found changes in these aspects, i.e., that with increasing age participants focus more on enjoyment instead of performance. This is supported by Altmeyer and Lessel (2017), showing that the main reason to play is that older adults enjoy spending time with other people and focus less on performance in games. In a follow-up study with participants being older than 75 years, Altmeyer et al. (2018a) use storyboards to explain commonly used gameful design elements to older adults. They used this approach to investigate the perception of these gameful design elements and found that the most commonly used elements—points, badges and leaderboards—are perceived negatively among older adults. Similarly, Kappen et al. (2016) focused on barriers and challenges in designing gameful applications encouraging physical activity among older adults. They conclude that personalizing gameful applications to support physical activity is important, as age-specific challenges need to be considered. In addition, the impacts of age and gender on the perception of Cialdini's persuasion strategies have been investigated by Orji et al. (2015). Regarding age, they found that the principle of scarcity is more valuable to younger people, while older adults are more driven by consistent commitment. Regarding gender, their results indicate that females are more responsive to most of the strategies. Gender-wise differences have also been demonstrated by Oyibo et al. (2017), who found that competition and virtual rewards are perceived as more persuasive by male participants. Furthermore, Oyibo and Vassileva (2019) investigated whether there are differences between collectivist and individualist cultures regarding the relevance of persuasive features in the physical activity domain. Their results show that collectivist cultures are more susceptible to persuasive features in general, whereas individualist cultures are more affected by personal persuasive features.

Albeit showing that personalization is essential for gameful applications encouraging physical activity, none of the factors presented above (personality traits, age and gender) is particularly suitable or was specifically developed for the purpose of personalizing gameful applications. The Hexad user-type model (Marczewski 2015) bridges this gap. It was specifically developed to cluster users of gameful systems and personalize the gameful design elements of a system. Establishing the basis for further research, Tondello et al. (2016) created a questionnaire to assess Hexad user types, which has been slightly adjusted and shown to be reliable and valid more recently (Tondello et al. 2018). In addition, the practicability of the Hexad model for personalizing gameful systems has been demonstrated by Tondello (2019), chapter 3. Here, a method for personalized gameful design based on the Hexad user types to select which gameful design elements to use was proposed. Consequently, the Hexad

user-type model has been utilized across different contexts and domains, showing that it is able to explain preferences for and perceptions of gameful design elements. In the health domain, Orji et al. (2018) examined the suitability of the Hexad model in the context of unhealthy alcohol consumption. The authors found that a users' Hexad type influences the perceived persuasiveness of persuasive strategies. Their results are in line with the Hexad user type definitions and support the applicability of the Hexad model in the health domain. The applicability of the Hexad model has also been demonstrated in an educational context by Mora et al. (2018). The authors investigated the potential of using the Hexad model to personalize learning experiences in order to motivate and engage students. They found that the approach that utilized the Hexad model to personalize the game design elements yielded higher engagement of the students. This underlines the usefulness of the Hexad model for tailoring gameful systems. In the context of energy efficiency at the workplace, Kotsopoulos et al. (2018) investigated the perception of certain gameful design elements and correlations to Hexad user types. As such, the authors showed the validity of the Hexad user model in another domain since they found similar correlations between gameful design elements and user types as Tondello et al. (2016). Tondello et al. (2017) proposed a conceptual framework based on an exploratory factor analysis of people's preferences in a general context, which allows to classify game design elements systematically. In line with previous results, expected correlations to the Hexad user types have been found. Further supporting the suitability of the Hexad model for explaining user preferences in gameful systems, Hallifax et al. (2019) found that the Hexad model is the most suitable typology for this purpose. They investigated which user models should be used and compared the BrainHex model, the Hexad model and the Big-5 personality model (McCrae and John 1992). They ran a study utilizing storyboards to explain game design elements to participants and found that 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 potentially because the Hexad model was specifically designed for gamification (which is not the case for other factors), and most of its user types are based on the well-established SDT (Ryan and Deci 2000).

### 3.3 Summary

Related work shows that there is increasing evidence that gameful design elements contribute positively to motivational and behavioral aspects in the context of physical activity (Aldenaini et al. 2020). However, research has also shown that roughly half of the interventions relying on a "one-size-fits-all" approach are only partially successful or even unsuccessful (Aldenaini et al. 2020). Similarly, the results of the presented papers show that the success of different game design elements differs substantially across interventions. Such contradictory findings pose the question of which factors moderate the perception of gameful interventions to encourage physical activity, to which we contribute in this article. It has been shown that static

factors such as personality traits (Halko and Kientz 2010; Jia et al. 2016, 2017; Orji et al. 2017) or demographic data such as age, culture or gender (Birk et al. 2017; Altmeyer and Lessel 2017; Altmeyer et al. 2018a; Kappen et al. 2016; Orji et al. 2015; Oyibo et al. 2017; Oyibo and Vassileva 2019) play a role in the perception of gameful design elements. However, none of these factors has been specifically developed for the purpose of personalizing gameful applications. In fact, the Hexad user-type model is the only model specifically designed for this purpose Orji et al. (2018). It has been shown to be reliable across various domains (Tondello et al. 2016; Mora et al. 2018; Orji et al. 2018; Kotsopoulos et al. 2018; Tondello et al. 2017), which emphasizes its relevance for personalizing gameful systems. We contribute the first investigation of the Hexad user-type model in the domain of encouraging physical activity, as far as we know.

Also, the aforementioned factors are static, i.e., do not change over time. This is contrary to findings by Niess and Woźniak 2018; Li et al. 2010; Epstein et al. 2015, who consistently provided evidence for the dynamic nature of goals and motivations. Therefore, it is important to find a way to formalize these dynamic processes and integrate them into a personalization approach when aiming at encouraging physical activity through gameful design. We contribute to this by investigating behavior change intentions as one way to deal with these dynamics, which has not been investigated before.

Furthermore, previous research has considered self-reported preferences based on storyboards (Oyibo and Vassileva 2019; Orji et al. 2018; Halko and Kientz 2010; Jia et al. 2017; Altmeyer et al. 2018a), textual descriptions (Tondello et al. 2016; Kotsopoulos et al. 2018) or videos (Jia et al. 2016). Since most previous studies used storyboards successfully to assess perceived preferences for game design elements, we follow a similar approach to investigate the potential of behavior change intentions and Hexad user types as factors for personalization in the context of physical activity. Using storyboards allows to recruit a large amount of participants from diverse populations as well as provides a common visual language that is easy to understand (Orji et al. 2018). However, in contrast to previous work, we additionally investigate whether personalizing a real gameful application based on the findings of the first, storyboards-based study, has an effect on affective experiences, user performance and enjoyment. This is an important contribution, as none of the previous works (also outside the physical activity context) allowed participants to interact with an implemented, personalized application.

## 4 Storyboards for gameful design elements

To investigate the *potential* of behavior change intentions and Hexad user types as factors for personalizing gameful applications encouraging physical activity, we follow the approach of illustrating gameful design elements by using storyboards (Orji et al. 2018; Halko and Kientz 2010).

**Table 1** Gameful design elements, a short textual description explaining what is depicted in the corresponding storyboard and the user types (“UT”) that we expect to be positively correlated to their perceived persuasiveness based on Marczewski (2015), Tondello et al. (2016)

Gamef. Des. Elem.	Short storyboard description	Expected UT
Virtual character	The appearance of a virtual character is linked to the amount of steps walked	AC, PL
Custom goal	The user sets herself a custom step goal	AC, FS
Personalized goal	The system personalizes the users’ step goal	AC
Challenge	The user manages to reach a demanding goal	AC
Badges	The user reaches her goal three times, unlocking a new badge	AC, PL
Points	The system rewards the user with points for walking steps	PL, AC
Rewards	After reaching the step goal three times, the user receives a coupon code	PL
Knowledge sharing	The user helps another user in a forum by answering a question	PH
Unlockable content	After reaching the step goal three times, the app unlocks a new feature	FS
Cheating	The user decides to cheat by driving a car to reach her step goal	DI
Social collaboration	A group of users have to collaborate, to reach their shared step goal	SO
Social competition	A group of users are shown on a leaderboard, competing for the top position	SO, PL

#### 4.1 Selection of gameful design elements

We selected commonly used gameful design elements based on the literature reviews by Seaborn and Fels (2015), Hamari and Sarsa (2014) and ensured to include at least one gameful design element for each Hexad user type, based on the correlations established by Marczewski (2015), Tondello et al. (2016). We ended up with a selection of twelve commonly used gameful design elements and created storyboards for each one, illustrating the respective elements as explained in Table 1. The design process of the storyboards followed the guidelines established by Truong et al. (2006), i.e., we used short texts to demonstrate novel aspects, included people to explain the interactive experience, indicated the passage of time only when necessary, and used the minimum level of detail required to understand the gameful design elements. We used walking as a concrete contextualization of physical activity in the storyboards and focused on encouraging a users’ step count. The context of step counting was used because it is among the most frequently used ones (Koivisto and Hamari 2019; Aldenaini et al. 2020) and is relevant to the general public (King et al. 2009). In the storyboards, a character was shown, interacting with a gameful application employing the specific gameful design element. Two exemplary storyboards (for Badges and Social Competition) can be seen in Fig. 2. All created storyboards are freely available on figshare.<sup>1</sup>

<sup>1</sup> <https://doi.org/10.6084/m9.figshare.7380902.v1>.

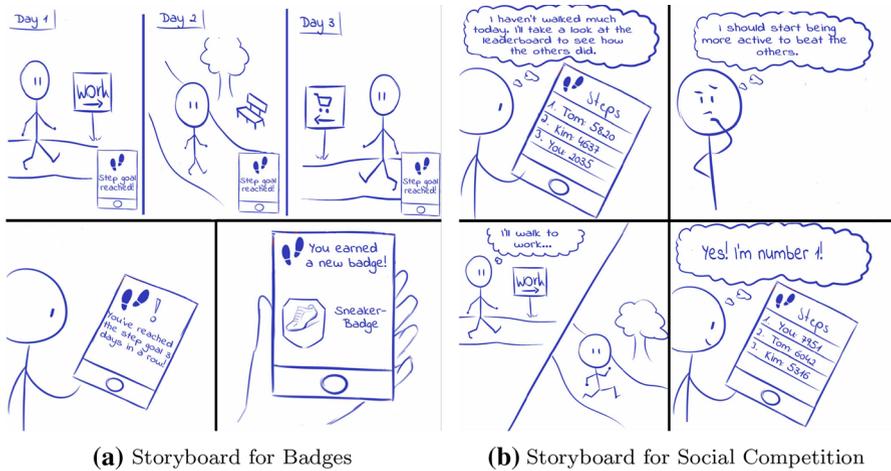


Fig. 2 Storyboards for gameful design elements

## 4.2 Storyboard validation

Before using the storyboards in the online study, we wanted to ensure that they actually explain the intended gameful design elements and are understandable to participants. Therefore, we conducted a qualitative pre-study in the laboratory.

### 4.2.1 Method

After answering demographic questions, the printed storyboards were shown to participants in random order. To understand whether users have problems understanding which gameful design element is illustrated by the storyboards, we conducted semi-structured interviews. The interview sessions were conducted by one researcher, and audio recordings were made. As a first step, participants were asked to describe the storyboards in their own words. When necessary, the interviewer asked questions to prompt participants to state which activities are shown in the storyboards. Questions included: “What is the character’s goal?” and “What means does the character use to achieve her goal?”. Afterwards, participants were given a short printed textual summary of each gameful design element. They were asked to assign these printed statements to each of the storyboards by placing them next to the respective storyboard. This was done to investigate whether the storyboards can be mapped to the respective gameful design elements and thus are successful in conveying them.

Finally, interviews were transcribed and analyzed by two independent raters (“R1,” “R2”). The raters received the transcriptions for each storyboard, without revealing which gameful design element was described by the participants. Their task was to evaluate which element was being described. This was also done to ensure that the storyboards explain the intended gameful design elements. Also,

raters were asked to rate how well the gameful design element was understood, based on the explanation provided by the participant, on a 5-point scale (1—very poor to 5—very well).

#### 4.2.2 Results

Eight German participants took part (four females, average age 21.75). To ensure that the ratings can be interpreted objectively, we calculated the inter-rater agreement and found it to be  $Kappa=0.75$ , which is considered as substantial (McHugh 2012). Analyzing the ratings of the two independent raters, we found that the participants understood the storyboards very well ( $M_{R1} = 4.90$ ,  $Min_{R1} = 4$ ;  $M_{R2} = 4.86$ ,  $Min_{R2} = 4$ ). This was supported by the fact that both raters successfully assigned the correct game element based on participants' storyboard descriptions. Regarding users assigning the textual summaries to the respective storyboard, only one assignment was incorrect. However, this wrong assignment was not due to a misunderstanding of the game element, but due to the participant misreading the descriptions of one of the game elements. The participant assured us that the storyboard and respective game element were clear to him.

### 5 Online study: potential of behavior change intentions and Hexad user types

After showing that the storyboards that were created for the twelve commonly used gameful design elements are comprehensible and successfully explain the intended gameful design elements, we used them to conduct an online study. Here, we were interested in the perceived persuasiveness of each gameful design element and potential differences related to behavior change intentions and Hexad user types. The results presented in this section were already published in a previous paper that we authored (Altmeyer et al. 2019).

#### 5.1 Procedure and method

The online survey was available in English and German. Participants were recruited via social media and Academic Prolific (paid £1.50 GBP). The study took between 10 and 15 minutes to complete and has been reviewed and received ethics clearance through an institutional Research Ethics Committee (#18-6-4).<sup>2</sup> After giving informed consent, participants were asked to provide demographic data and rate their gaming behavior on 5-point Likert scales (5=strong agreement). Behavioral intentions were operationalized by using a validated scale assessing the stage of change ("SoC") within the context of physical activity (Marcus et al. 2008). To analyze the effect of behavior change intentions on the perceived persuasiveness of the

<sup>2</sup> <https://erb.cs.uni-saarland.de/>, last accessed 2021/01/10 20:20:33.

gameful design elements, participants were split into two groups: “Low-SoC” (participants who did not take action so far, having a SoC  $\leq 3$ ) and “High-SoC” (participants who did take action, having a SoC  $\geq 4$ ). This follows the same procedure as was done by Xiao et al. (2004), who split participants in preaction (not achieving their goal) and action (at their goal) groups. Next, participants’ Hexad user type was determined using the Hexad user type scale (Tondello et al. 2016).

Afterward, the main part of the online survey started. Here, participants were shown the 12 storyboards in a randomized order. To measure the persuasiveness of each gameful design element depicted in the storyboards, we adapted the perceived persuasiveness scale by Drozd et al. (2012) in the same way as was done by Orji et al. (2017). The scale consists of four items being measured on 7-point Likert scales. In line with previous research using this scale (Orji et al. 2014, 2017), the internal consistency is excellent, with Cronbach’s alpha = .97. Since a Shapiro–Wilk test revealed that the perceived persuasiveness responses were not normally distributed, we used nonparametric tests for the analysis. Consequently, the effect of behavior change intentions on the perceived persuasiveness of the gameful design elements was assessed by using Mann-Whitney U tests. For correlation analyses, Kendall’s  $\tau$  was used, since it is well-suited for nonparametric data (Howell 2002). For the interpretation of the correlations, it should be considered that Kendall’s  $\tau$  is usually lower than Pearson’s  $r$  for the same effect sizes. Therefore, we transformed interpretation thresholds for Pearson’s  $r$  to Kendall’s  $\tau$ , according to Kendall’s formula (Walker 2003) (small effect:  $\tau = 0.2$ ; medium effect:  $\tau = 0.3$ ; large effect:  $\tau = 0.5$ ).

## 5.2 Results

After excluding nine participants who were either unable to exercise or answered all gaming-related questions with “Strongly disagree,” we considered 179 valid responses. Of those participants, 55.3% self-reported their gender as female, 44.1% as male and 0.6% as non-binary. Most participants (38%) were aged 18–24 years, followed by 25–31 (34.1%), 32–38 (17.3%), 39–45 (6.7%) and younger than 18 (1.7%). The remaining participants were aged 45 and older (1.7%). Participants stated that they had a passion for video games ( $M = 3.70$ ,  $SD = 1.11$ ,  $Mdn = 4.00$ ) and that they frequently play video games ( $M = 3.58$ ,  $SD = 1.24$ ,  $Mdn = 4.00$ ). Seventy-two participants were in the Low-SoC and 107 participants in the High-SoC group. Regarding the average scores of the Hexad user types questionnaire, Free Spirits ( $M=22.75$ ,  $SD=3.43$ ) and Achievers ( $M=22.09$ ,  $SD=3.40$ ) showed the highest and second-highest average scores, followed by Players ( $M=21.56$ ,  $SD=4.14$ ) and Socialisers ( $M=19.27$ ,  $SD = 4.77$ ). Philanthropists ( $M = 17.50$ ,  $SD = 2.49$ ) and Disruptors ( $M = 16.30$ ,  $SD = 4.76$ ) followed with lower average scores.

**Table 2** Perceived persuasiveness of gameful design elements in the Low- and High-SoC groups and results of Mann–Whitney U tests (“Diff. sig.”)

	Low-SoC	High-SoC	Diff. sig.
Virtual character	$M = 4.05$ , $SD = 1.77$ , Mdn = 4.50	$M = 3.94$ , $SD = 1.81$ , Mdn = 4.25	–
Custom goal	$M = 4.34$ , $SD = 1.49$ , Mdn = 4.63	$M = 4.70$ , $SD = 1.55$ , Mdn = 5.25	–
Personalized goal	$M = 4.88$ , $SD = 1.44$ , Mdn = 5.00	$M = 4.93$ , $SD = 1.38$ , Mdn = 5.25	–
Challenge	$M = 4.32$ , $SD = 1.65$ , Mdn = 4.75	$M = 4.88$ , $SD = 1.27$ , Mdn = 5.00	$p = 0.045$ $U = 3173.50$
Badges	$M = 3.95$ , $SD = 1.57$ , Mdn = 4.00	$M = 4.46$ , $SD = 1.40$ , Mdn = 4.75	$p = 0.028$ $U = 3108.50$
Points	$M = 4.39$ , $SD = 1.46$ , Mdn = 5.00	$M = 4.52$ , $SD = 1.43$ , Mdn = 4.50	–
<b>Rewards</b>	$M = 5.16$ , $SD = 1.48$ , Mdn = 5.25	$M = 5.50$ , $SD = 1.39$ , Mdn = 5.75	–
Knowledge sharing	$M = 4.06$ , $SD = 1.52$ , Mdn = 4.25	$M = 4.26$ , $SD = 1.51$ , Mdn = 4.50	–
Unlockable content	$M = 4.70$ , $SD = 1.49$ , Mdn = 5.00	$M = 4.84$ , $SD = 1.53$ , Mdn = 5.00	–
Cheating	$M = 2.12$ , $SD = 1.16$ , Mdn = 2.00	$M = 2.35$ , $SD = 1.44$ , Mdn = 2.00	–
Social collaboration	$M = 4.23$ , $SD = 1.56$ , Mdn = 4.88	$M = 4.81$ , $SD = 1.61$ , Mdn = 5.25	$p = 0.009$ $U = 2963.50$
Social competition	$M = 4.09$ , $SD = 1.74$ , Mdn = 4.50	$M = 4.61$ , $SD = 1.76$ , Mdn = 4.75	$p = 0.048$ $U = 3180.50$

These results were published in our previous paper (Altmeyer et al. 2019)

### 5.2.1 SoC and gameful design elements

To investigate whether behavior change intentions have a moderating effect on the perceived persuasiveness of the gameful design elements, we performed a two-sided Mann–Whitney U test to analyze potential differences in the two groups (Low-SoC and High-SoC) for each gameful design element. Following the suggestions provided by Armstrong (2014), we did not adjust probability values for these tests, because we interpreted these tests independently (as unrelated group means) and because the results were used to inform the hypotheses in the laboratory experiment presented in Sect. 7.

The overview of the results can be found in Table 2. It can be seen that the perceived persuasiveness of the gameful design elements is different between the two groups. We found significant differences between the two groups for four gameful

**Table 3** Kendall's  $\tau$  and significance between the Hexad user types and the gameful design elements

	AC	DI	FS	PH	PL	SO
Virtual character	–	–	–	–	<b>.237**</b>	.114*
Custom goal	<b>.205**</b>	–	<b>.132*</b>	.119*	–	.106*
Personalized goal	<b>.211**</b>	–	–	.145**	–	–
Challenge	<b>.200**</b>	–	.145**	–	.177**	–
Badges	<b>.122*</b>	–	–	–	<b>.223**</b>	–
Points	<b>.201**</b>	–	.110*	.192**	<b>.169**</b>	.105*
Rewards	.114*	–	–	.152**	<b>.250**</b>	.109*
Knowledge sharing	.123*	–	–	<b>.234**</b>	–	.175**
Unlockable content	.140**	–	<b>.143**</b>	–	.163**	–
Cheating	–	<b>.157**</b>	–	–	–	–
Social collaboration	.147**	–	.153**	.145**	.216**	<b>.314**</b>
Social competition	.105*	–	–	–	<b>.370**</b>	<b>.204**</b>

Bold entries represent expected correlations (as stated in Table 1). \* $p < .05$ , \*\* $p < .01$ . These results were published in our previous paper (Altmeyer et al. 2019)

design elements. For instance, Badges and Challenges were perceived as significantly more persuasive in the High-SoC than in the Low-Soc group. Also, Social Competition and Social Collaboration were perceived as significantly more persuasive in the High-SoC group. In sum, we establish result **R1: Behavior change intentions have a moderating effect on the perceived persuasiveness of gameful design elements in the physical activity context.**

This main result is potentially explainable by goal-setting theory, stating that goals are most effective when users are committed to them (Tondello et al. 2018a; Locke and Latham 2002). This is unlikely for users in the Low-SoC group, since their motivation to increase their physical activity levels is not yet internalized and thus commitment is lower. Specifically regarding Badges and Challenges, participants in the Low-SoC group might have considered themselves as not to be able to reach the established goals (Fogg 2002). A potential reason for the significant difference between the groups regarding social gameful design elements (Social Competition and Social Collaboration) might be related to the fear to not be able to keep up with other users (Fogg 2002). This might have detrimentally affected users' perceived persuasiveness in the Low-SoC group. In sum, these findings show that the SoC is a relevant factor that should be considered in personalizing gameful systems in the physical activity context.

### 5.2.2 Hexad user types and gameful design elements

To analyze the impact of Hexad user types on the perceived persuasiveness of gameful design elements, we followed the approach of previous research using the Hexad model (Tondello et al. 2016; Orji et al. 2018; Kotsopoulos et al. 2018) and analyzed

correlations between Hexad user-type scores and the perceived persuasiveness of each gameful design element.

The overview of these findings is shown in Table 3. It can be seen that 16 positive correlations between Hexad user types and gameful design elements out of 17 expected correlations (see Table 1) were found. This replicates previous findings (Tondello et al. 2016; Orji et al. 2018; Kotsopoulos et al. 2018) and supports the usefulness of the Hexad user-type model in the physical activity context. The positive correlation between the gameful design element “Virtual Character” and the “Achiever” user type is the only correlation that was expected, but could not be found, given our data. Based on this, we establish **R2: The Hexad user type has a moderating effect on the perceived persuasiveness of gameful design elements in the physical activity context.** In addition to expected correlations, some unexpected correlations were found. This could be a result of considering a different context and using storyboards instead of textual descriptions, compared to Tondello et al. (2016). It is also in line with previous research (Orji et al. 2018; Tondello et al. 2016). A more detailed discussion of the online study can be found in Sect. 8.

## 6 Endless universe: design and implementation of a personalized gameful application to encourage physical activity

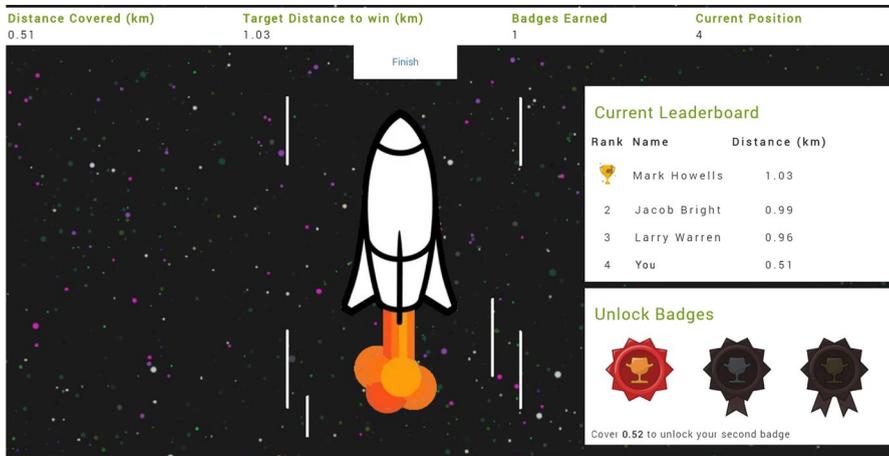
The *actual effects* of personalizing gameful applications based on behavior change intentions and Hexad user types on task performance and user experience cannot be investigated without allowing users to interact and experience the gameful design elements in a real system. Therefore, we implemented Endless Universe, a gameful application that builds upon the results of the online study to investigate the effects of personalization on these aspects.

### 6.1 Design and concept

Endless Universe ties the distance covered on a treadmill to the progress within several gameful design elements. To investigate which effects personalization has on measures related to the users’ performance and experience, we decided to use the findings from the storyboards-based online study presented before to tailor Endless Universe to a specific user group.

#### 6.1.1 Theme

We decided to use outer space as the main theme of the gameful application. This decision is based on previous research using gameful applications encouraging physical activity, which demonstrated that this theme is well perceived within the physical activity context (Saksono et al. 2015; Doyle et al. 2011a, b; Finkelstein et al. 2010; Buttussi et al. 2007; Cuzzort and Starner 2008). The core mechanic in the gameful application is a spaceship exploring an endless universe. Hereby, the real-time distance covered by the user on the treadmill has a direct influence on the



**Fig. 3** Screenshot of the Endless Universe application. The distance covered, target distance, number of badges unlocked and current position on the leaderboard are shown on the top. The leaderboard and badges are shown on the right side of the screen

speed of the spaceship moving forward in the space exploration. The spaceship is shown prominently in the middle of the screen, and a moving illusion is created by animating the background of the scene (i.e., stars and particles are moving faster or slower). The distance covered by the user is shown permanently in the application. When starting the application for the first time, an introduction is given to the users, explaining that they belong to an alien species which is competing to explore the universe with their spaceships. Figure 3 shows a screenshot of the application.

### 6.1.2 Goal setting

Endless Universe establishes a target distance to cover, which is shown next to the distance covered in the main screen of the application. This target distance is personalized to the user, i.e., based on a users' fitness level. This was done to make sure that the target distance is reachable to all users and thus comparable. This is in line with previous research within this context (Lin et al. 2006; Miller and Mynatt 2014). More specifically, this target distance was 10% higher than the previously covered distance. The gameful design elements, which are described next, operate on this target distance.

### 6.1.3 Gameful design elements

The findings of the storyboards-based online study presented above show that behavior change intentions and Hexad user types are relevant factors for personalizing gameful applications encouraging physical activity. Based on these findings, we derived a set of gameful design elements to investigate the effects of personalization. As such, we decided to use the gameful design elements *Badges*, *Challenges* and *Social Competition*. These gameful design elements were shown to be

perceived as significantly more persuasive in the High-SoC than in the Low-SoC group (**R1**, see Table 2) and to be positively correlated with the Achiever, Player and Socialiser Hexad user types (**R2**, see Table 3). Therefore, we expected that Endless Universe should be suitable for users belonging to the High-SoC group or scoring particularly high on Achiever, Socialiser or Player. The realization of the gameful design elements is described in the following:

### *Badges*

There are three different badges in the gameful application. To account for interpersonal performance differences, the thresholds to unlock badges were established relatively to the target distance. The first badge is unlocked when reaching 20% of the target distance and is visualized through a bronze trophy. The second badge, a silver trophy, is unlocked when reaching 50% of the target distance. Finally, the golden badge is unlocked when reaching 100% of the target distance. This progression concept follows the recommendations related to progression stairs in games by Werbach and Hunter (2012). The badges were shown on the right side of the screen and darkened until they were unlocked. The remaining distance until unlocking the next badge was shown permanently below the badges. Based on **R1** and **R2**, this gameful design element should be perceived particularly well by users belonging to the *High-SoC* group and users scoring high on the *Achiever* or *Player* factors of the Hexad.

### *Challenges*

The ultimate challenge of Endless Universe is to reach the target distance. This is explained to the user as part of the onboarding procedure before starting the gameful application. When reaching the target distance and thus mastering the main challenge of the application, a so-called explorer of the day trophy is unlocked and shown to the user. This gameful design element should be perceived particularly well by users belonging to the *High-SoC* group (**R1**) and users having a high *Achiever* score (**R2**).

### *Social Competition*

We used a leaderboard to introduce social competition to the gameful application, positioned on the right-hand side of the screen. In this leaderboard, fictitious users were shown, similar to previous gamification studies (Mekler et al. 2017). This was done to ensure the comparability across participants, i.e., that all participants had the same chance to rise in ranks, and to avoid introducing a confounding variable (Von Ahn and Dabbish 2008). Similar to *Badges*, there were three other fictitious users who covered distances that were calculated in relation to the target distance described above. The fictitious user on the first rank covered the target distance, the fictitious user on the second rank covered 5% less than the target distance and the fictitious user on the third rank covered 8% less than

the target distance. This follows the same progression scheme as was used for *Badges* and thus follows recommendations established by Werbach and Hunter (2012).

## 6.2 Implementation

The user interface part of Endless Universe was implemented as a web application and capturing the distance covered on the treadmill was realized by using an Arduino Uno board and a QRE1113 infrared reflectance sensor is comprised of an infrared emitting LED and an infrared sensitive phototransistor. The hardware and user interface are explained in the following.

### 6.2.1 Hardware to capture the covered distance on the treadmill

Since the covered distance is a direct input to the gameful application, we implemented a system to track the distance covered on the treadmill. We placed reflecting light tape on the belt of the treadmill in equal, pre-defined distances and used an infrared reflectance sensor to detect the tape. We used an Arduino Uno, which was connected to a PC via USB to send an event to the main application running on the PC whenever a tape was detected.

### 6.2.2 User interface

The number of events that were triggered when the reflecting tape on the belt was detected by the Arduino was sent via USB to a NodeJS Express webserver running on a PC in a real time. The webserver calculated the distance covered based on the number of detections, i.e., the total distance could be derived with a maximum discrepancy of 3.1 meters (the tape was placed every 3.1 meters). Besides calculating the covered distance, the webserver is responsible for the game logic, i.e., deriving the current rank of the user on the leaderboard, checking whether a badge should be unlocked and whether the main challenge was completed. This information is populated to the frontend using bidirectional websockets. The frontend itself was realized using HTML, CSS and JavaScript. Three.js was used for the visualization of the space, the rocket and to create the moving illusion with various speeds. Moreover, Bootstrap was used to make sure that the application adapts to various screen sizes, and jQuery was used to manipulate the DOM of the web application whenever updated data from the webserver have been sent.

## 7 Laboratory study: effects of personalization

To investigate whether the findings from the online study (**R1**, **R2**), which were based on the perception of storyboards, lead to effects on a user's performance or experience when actually interacting with a gameful application, we conducted a laboratory study. In this laboratory study, participants were running on a treadmill and thus interacted with Endless Universe. In the following, the procedure, method and the results of this study are presented.

### 7.1 Procedure and method

The study followed a within-subjects design with two conditions. When recruiting participants, we used the same validated questionnaire as in the online study to assess the SoC within the context of physical activity (Marcus et al. 2008), to make sure that an equal number of Low- and High-SoC participants was recruited. In the baseline condition, participants were running on a treadmill without getting any kind of feedback. (The display of the treadmill was covered using black foil.) In the intervention condition, Endless Universe was deployed on a 10-inch tablet device, which was placed where the display of the treadmill is located, to ensure that participants can easily see the gameful application. The study started with the baseline phase to avoid detrimental effects when removing gameful design elements (Hamari and Sarsa 2014) and to establish the target distance in the intervention phase (to make sure that the target distance is reachable to all users (Lin et al. 2006; Miller and Mynatt 2014). After giving informed consent, participants were asked to fill out a survey. In this survey, demographic data were collected. Next, the Hexad user type was assessed using the validated questionnaire by Tondello et al. (2018), followed by a validated questionnaire to assess the SoC within the context of physical activity (Marcus et al. 2008).

After completing this survey, participants were asked to run on the treadmill for 10 minutes in a speed that they felt comfortable with. They were told to stop running when feeling uncomfortable. Drinks were provided.

After running for 10 minutes, participants were asked to complete a second survey. In this survey, the validated version of the Positive and Negative Affect Schedule ("PANAS") (Watson et al. 1988) was administered in order to assess affective experiences while running. Next, participants were asked to fill out the 22-item task evaluation questionnaire of the Intrinsic Motivation Inventory ("IMI") (McAuley et al. 1989; Ryan 1982) to assess intrinsic motivation and enjoyment of the running activity. Finally, Borg's Rating of Perceived Exertion ("RPE") (Borg 1970) was administered to assess how exhausting participants perceived the activity. In this scale, users choose a number between 6 ("no exertion") and 20 ("maximum exertion possible") to describe their perceived exertion. Finally, a date for the intervention

phase was scheduled. We made sure that there is a break of at least one full day between the baseline and intervention phase.

The intervention phase followed exactly the same procedure. The only difference was that Endless Universe was in place while running. The task was exactly the same, i.e., participants were asked to run on the treadmill for 10 minutes in a speed that they felt comfortable with. The target distance was established based on the covered distance in the baseline phase, as described in Sect. 6.1.2. After running for 10 minutes, the same questionnaires as in the baseline (PANAS, IMI, RPE) were administered.

Participants were compensated by a 10 Euro amazon gift card. The study has been reviewed and received ethics clearance through an institutional Research Ethics Committee (#19-12-3).<sup>3</sup>

## 7.2 Hypotheses

Based on the findings of the storyboards-based pre-study and previous research, we expected to find evidence for the following hypotheses:

- H1** One-size-fits-all gamification affects performance and experience
- H1a** The covered distance is higher when using Endless Universe
- H1b** Users perceive running as more enjoyable using Endless Universe
- H1c** Users have stronger affective experiences with Endless Universe
- H2** SoC affects performance and experience with Endless Universe
- H2a** The improvement in distance is higher for High-SoC users
- H2b** High-SoC users perceive Endless Universe as more enjoyable
- H2c** High-SoC users have stronger affective experiences
- H3** Hexad types affect performance and experience with Endless Universe
- H3a** The improvement in distance is higher for AC, PL, SO
- H3b** AC, PL, SO perceive Endless Universe as more enjoyable
- H3c** AC, PL, SO have stronger affective experiences

**H1** is motivated by previous work, showing that gameful applications can increase physical activity and can have positive effects on the user experience when doing sports (Aldenaini et al. 2020; Koivisto and Hamari 2019). Consequently, **H1** can be seen as a replication of previous work and is important to demonstrate the overall effectiveness and validity of Endless Universe. **H1** is analyzed by conducting paired samples t-tests or Wilcoxon signed-rank tests (when the assumptions of the t-test were not met). **H2** stems from findings of the storyboards-based online study presented in Sect. 5. In this study, we found that the perceived persuasiveness of Social Competition, Badges and Challenges is significantly higher among High-SoC users. Since we are using these gameful design elements in Endless Universe, we expect that the increased perceived persuasiveness should be reflected in

---

<sup>3</sup> <https://erb.cs.uni-saarland.de/>, last accessed 2021/01/10 20:20:33.

**Table 4** Dependent variables of the laboratory study for the baseline and intervention condition and results of paired samples t-tests/Wilcoxon signed-rank tests (“Diff. sig.”) comparing them

	Baseline <i>N</i> = 20	Intervention <i>N</i> = 20	Diff. sig.
Distance covered [km]	<i>M</i> = 0.96, <i>SD</i> = 0.32, Mdn = 0.97	<i>M</i> = 1.13, <i>SD</i> = 0.36, Mdn = 1.12	<i>p</i> = 0.003 <i>Z</i> = 24.00
RPE [scale from 6 to 20]	<i>M</i> = 9.35, <i>SD</i> = 2.11, Mdn = 9.00	<i>M</i> = 11.10, <i>SD</i> = 2.95, Mdn = 11.50	<i>p</i> = 0.027 <i>t</i> = -2.40
IMI enjoyment [scale from 1 to 7]	<i>M</i> = 4.88, <i>SD</i> = 1.97, Mdn = 5.50	<i>M</i> = 5.43, <i>SD</i> = 1.42, Mdn = 5.67	-
IMI competence [scale from 1 to 7]	<i>M</i> = 4.42, <i>SD</i> = 1.89, Mdn = 4.84	<i>M</i> = 5.38, <i>SD</i> = 1.35, Mdn = 5.84	<i>p</i> = 0.008 <i>t</i> = -2.97
IMI pressure [scale from 1 to 7] Mdn = 1.33	<i>M</i> = 1.73, <i>SD</i> = 1.04, Mdn = 6.50	<i>M</i> = 6.07, <i>SD</i> = 1.09, <i>t</i> = -10.40	<i>p</i> < 0.001
IMI choice [scale from 1 to 7]	<i>M</i> = 6.03, <i>SD</i> = 1.27, Mdn = 6.67	<i>M</i> = 2.40, <i>SD</i> = 1.56, Mdn = 2.00	<i>p</i> < 0.001 <i>t</i> = 7.42
PANAS pos. [scale from 1 to 5]	<i>M</i> = 3.03, <i>SD</i> = 0.56, Mdn = 2.90	<i>M</i> = 3.40, <i>SD</i> = 0.83, Mdn = 3.45	-
PANAS neg. [scale from 1-5]	<i>M</i> = 2.91, <i>SD</i> = 0.14, Mdn = 3.00	<i>M</i> = 2.90, <i>SD</i> = 0.0.17, Mdn = 2.90	-

an increased actual performance and experience. **H2** is analyzed by splitting participants in Low- and High-SoC groups and conducting independent-samples t tests or Mann–Whitney U tests (when assumptions of the t test were not met). Similarly, **H3** bases on our findings from the online study, which revealed significant correlations between the Socialiser, Achiever, Player and the aforementioned gameful design elements. Also, previous research has demonstrated similar correlations for these gameful design elements in different contexts (Tondello et al. 2016; Orji et al. 2018; Kotsopoulos et al. 2018; Hallifax et al. 2019). To analyze **H3**, we calculated bivariate correlation coefficients. Similar to the online study, we used Kendall’s  $\tau$ , since it is well suited for nonparametric data (Howell 2002). Also, research has recommended using Kendall’s  $\tau$  when the sample size is rather low (Bishara and Hittner 2012). Since we established one-directional hypotheses beforehand (**H3a**, **H3b**, **H3c**) and to further increase the power of the correlation analysis, we used one-sided tests. Again, when interpreting the correlation coefficients, it should be considered that Kendall’s  $\tau$  is lower than Pearson’s *r* for the same effect sizes (see Sect. 5.1).

### 7.3 Results

We recruited 20 participants. Of those participants, 11 self-reported their gender as male and 9 as female. Most participants (50%) were aged 25–31 years, followed by 18–24 (45%) and 32–38 (5%). The number of participants across the Low- and

High-SoC groups was equal (10 participants in each group). Regarding the average scores of the Hexad user types questionnaire, Achievers ( $M = 24.80, SD = 2.35$ ) and Philanthropists ( $M = 24.10, SD = 3.35$ ) showed the highest and second-highest average scores, followed by Players ( $M = 24.10, SD = 2.92$ ) and Free-Spirits ( $M = 23.40, SD = 3.12$ ). Socialisers ( $M = 22.90, SD = 3.80$ ) and Disruptors ( $M = 17.40, SD = 3.95$ ) followed with lower average scores.

### 7.3.1 Effects of “One-Size-Fits-All” gamification

First, we investigated whether Endless Universe has an effect on the performance of users, i.e., whether it motivated participants to cover more distance than in the baseline. This is important to replicate previous research, which showed the effectiveness of one-size-fits-all gamification in this domain (Altmeyer et al. 2018b; Chen and Pu 2014; Koivisto and Hamari 2019).

Table 4 shows the means, standard deviations, medians and significant differences for all dependent variables of the study for the baseline and intervention phase. We found a significant difference in the covered distance between the baseline and intervention condition ( $Z = 24.00, p = 0.003$ ). Based on this, we establish result **R3: Participants covered a significantly higher distance when using Endless Universe**. Next, we analyzed whether RPE differs across the conditions. Again, we found a significant difference between the intervention and baseline phase in perceived exertion ( $t = -2.40, p = 0.027$ ). Thus, **R4: Perceived Exertion is higher when using Endless Universe** confirms that the increased distance (**R3**) is also reflected in the subjectively higher feeling of exertion. Regarding enjoyment and user experience, we compared the factors of the IMI and PANAS. Here, we found a significant difference for the competence ( $t = -2.97, p = 0.008$ ), pressure ( $t = -10.40, p < .001$ ), and choice ( $t = 7.42, p < .001$ ) factors of the IMI. No significant effects were found for the enjoyment factor ( $p = 0.20$ ). Thus, we establish **R5: Perceived competence and pressure is higher when using Endless Universe** and **R6: Perceived choice is lower when using Endless Universe**. Regarding affective experience, no significant effects were found for the positive ( $p = 0.08$ ) nor the negative affect factor ( $p = 0.62$ ).

### 7.3.2 Effects of SoC-personalization

Similar to the online study, we split participants in Low- and High-SoC groups and compared these two groups to check for significant effects. To ensure the comparability of the improvement of performance, we did not consider the absolute distance but calculated the relative improvement (i.e., we divided the distance covered in the intervention phase by the distance covered in the baseline phase). Table 5 provides an overview of descriptive data and significant differences. It can be seen that we could not find a significant effect in distance improvement between the Low- and High-SoC groups ( $p = 0.07$ ), and no significant effect was found for the perceived exertion between the groups ( $p = 0.24$ ). In addition, none of the factors

**Table 5** Dependent variables of the laboratory study in the Low- and High-SoC groups and results of independent t tests/Mann–Whitney U tests (“Diff. sig.”) comparing them

	Low-SoC <i>N</i> = 10	High-SoC <i>N</i> = 10	Diff. sig.
Distance improvement [intervention/baseline]	<i>M</i> = 1.37, <i>SD</i> = 0.31, Mdn = 1.23	<i>M</i> = 1.11, <i>SD</i> = 0.31, Mdn = 1.15	–
RPE [scale from 6–20]	<i>M</i> = 11.90, <i>SD</i> = 2.89, Mdn = 12.00	<i>M</i> = 10.30, <i>SD</i> = 2.95, Mdn = 11.00	–
IMI enjoyment [scale from 1–7]	<i>M</i> = 5.13, <i>SD</i> = 1.77, Mdn = 5.33	<i>M</i> = 5.73, <i>SD</i> = 0.97, Mdn = 5.83	–
IMI competence [scale from 1–7]	<i>M</i> = 5.07, <i>SD</i> = 1.62, Mdn = 5.17	<i>M</i> = 5.70, <i>SD</i> = 1.01, Mdn = 6.00	–
IMI pressure [scale from 1 to 7]	<i>M</i> = 5.80, <i>SD</i> = 1.25, Mdn = 6.00	<i>M</i> = 6.33, <i>SD</i> = 0.87, Mdn = 6.67	–
IMI choice [scale from 1 to 7]	<i>M</i> = 2.80, <i>SD</i> = 1.63, Mdn = 3.33	<i>M</i> = 2.00, <i>SD</i> = 1.23, Mdn = 1.67	–
PANAS pos. [scale from 1 to 5]	<i>M</i> = 3.02, <i>SD</i> = 0.55, Mdn = 2.75	<i>M</i> = 3.77, <i>SD</i> = 0.92, Mdn = 4.10	<i>p</i> = 0.040 <i>t</i> = 2.21
PANAS neg. [scale from 1 to 5]	<i>M</i> = 2.80, <i>SD</i> = 0.13, Mdn = 2.80	<i>M</i> = 3.00, <i>SD</i> = 0.15, Mdn = 3.00	<i>p</i> = 0.005 <i>t</i> = 3.16

**Table 6** Kendall’s  $\tau$  and significance between the Hexad user types and the dependent variables in the laboratory study

	AC	DI	FS	PH	PL	SO
Distance improvement [intervention/baseline]	–	– .387**	–	–	–	–
RPE [scale from 6 to 20]	–	–	–	–	–	–
IMI Enjoyment [scale from 1 to 7]	–	–	–	–	–	–
IMI competence [scale from 1 to 7]	–	–	–	–	–	.304*
IMI pressure [scale from 1 to 7]	–	–	.467**	–	–	–
IMI choice [scale from 1–7]	–	–	–	–	–	–
PANAS pos. [scale from 1 to 5]	–	–	–	–	–	–
PANAS neg. [scale from 1 to 5]	–	–	–	–	–	–

\**p* < .05, \*\**p* < .01

of the IMI revealed a significant difference (enjoyment:  $p = 0.36$ ; competence:  $p = 0.31$ ; pressure:  $p = 0.28$ ; choice:  $p = 0.23$ ). However, we found a significant effect for affective experiences, i.e., a significant effect was found for both positive ( $t = 2.21, p = 0.040$ ) and negative affect ( $t = 3.16, p = 0.005$ ). Both positive and negative affects were significantly higher in the High-SoC group. Consequently, we establish **R7: Participants in the High-SoC group had stronger affective experiences.**

### 7.3.3 Effects of hexad personalization

The results of the correlation analysis can be seen in Table 6. When analyzing the significant correlations between the dependent variables of the laboratory study and the AC, PL, SO Hexad user types, we found that the score in the Socialiser factor of the Hexad is positively correlated to the perceived competence of the IMI when interacting with Endless Universe, having a medium effect size. This suggests that Socialisers perceived the feedback of the gameful design elements as particularly confirming and leads to **R8: Endless Universe positively affected the perceived competence of Socialisers.** We also found correlations for Hexad user types besides AC, PL and SO. For these remaining Hexad user types, we expected to find either no conclusive correlations or expected that negative effects on the user experience or performance would be found. Since we did not have specific a priori formulated assumptions for these user types, we used two-tailed tests for them. We found a negative, medium-sized correlation between the distance improvement and the disruptor. This suggests that disruptors were not encouraged to increase their performance by Endless Universe and leads to **R9: The performance of Disruptors was negatively affected by Endless Universe.** Also, we found a medium-to-strong positive correlation between the perceived pressure and free spirits. This means that **R10: Perceived pressure was particularly high for Free Spirits when using Endless Universe.**

## 8 Discussion and limitations

In the course of the two main studies of this paper, we investigated Hexad user types and behavioral intentions as factors to personalize gameful applications in the context of physical activity. First, we investigated the potential of these factors by creating storyboards illustrating twelve commonly used gameful design elements in the fitness context. After ensuring that the storyboards are comprehensive and explain the intended gameful design elements in a qualitative pre-study ( $N = 8$ ), we conducted an online study assessing the perceived persuasiveness of each gameful design element. Our findings support the potential of these personalization factors. Next, we implemented a gameful application aiming to motivate users to cover a higher distance on a treadmill to investigate whether the theoretical findings of the storyboards-based study lead to effects on performance, enjoyment or affective experiences when allowing users to interact with a real implementation of gameful design elements. In this section, we

discuss the findings of the online study using storyboards, the laboratory study using the gameful application and the contributions of our paper.

### 8.1 Storyboards-based online study

In the online study, we investigated the effect of behavior change intentions and Hexad user types on the perceived persuasiveness of twelve commonly used gameful design elements in the physical activity context using storyboards. We contribute two main findings: First, we found multiple significant differences between both groups in the perceived persuasiveness of gameful design elements, supporting the potential and relevance of behavior change intentions as a factor to personalize gameful applications in the physical activity domain (**R1**). The second contribution of the online study lies in supporting the validity of the Hexad model in the physical activity context. Confirming previous findings (Tondello et al. 2016; Orji et al. 2018), we found 16 out of 17 expected correlations between gameful design elements and Hexad user types. Thus, our findings validate previous results (Tondello et al. 2016; Marczewski 2015) in another context and illustrate the usefulness of Hexad user types as a static factor to explain user preferences in this domain (**R2**). On a more abstract level, these findings show that considering contextual motivation (operationalized through SoC; increasing SoC reflecting more intrinsic motivation (Mullan and Markland 1997) might complement static factors such as the Hexad model and should be investigated further in future work. As a limitation, it should be noted that the storyboards, although evaluated for their suitability, are a matter of interpretation. This is particularly relevant for how the gameful design elements were illustrated and described. Related to this, the main limitation of the online study is the utilization of storyboards and assessing *perceived* persuasion. While this approach is common in personalization research targeting gameful systems (Orji et al. 2013, 2018; Halko and Kientz 2010; Hallifax et al. 2019; Altmeyer et al. 2018a; Orji et al. 2014), it does not allow to assess actual effects when giving participants the chance to experience a gameful application and interact with its gameful design elements. To bridge this limitation, we implemented a gameful application encouraging physical activity and investigated its effects on performance, affective experiences and enjoyment.

### 8.2 Laboratory study

In the laboratory study, we used the aforementioned gameful application and investigated its effectiveness and the effects of behavioral intentions and Hexad user types. We used the findings from the online study to decide which gameful design elements to use. Consequently, we ended up using Badges, Challenges and Social Competition. These elements were shown to be perceived as significantly more persuasive for user in the High-SoC group in the online study. Also, expected correlations were found between the perceived persuasiveness of these three elements and the Hexad user types such as Socialiser, Achiever and Player. Thus, by using these

gameful design elements, we expected to see positive effects on the aforementioned dependent variables for High-SoC users and users scoring particularly high on the Socialiser-, Achiever-, or Player- factors of the Hexad.

As a first step of our analysis, we investigated whether the gameful elements used in "Endless Universe" are effective (**H1**). We found that "Endless Universe" led to a significant increase in covered distance on the treadmill (**R3**) and also to a subjectively higher exertion (**R4**), thus supporting **H1a: The covered distance is higher when using "Endless Universe"**. This finding is important as it replicates previous research (Aldenaini et al. 2020; Koivisto and Hamari 2019) and thus demonstrates the validity of the gameful application itself. We also analyzed whether there is a difference in factors of the IMI. We found that perceived competence and perceived pressure are significantly higher when using "Endless Universe" (**R5**) and that perceived choice is significantly lower (**R6**). The increased perceived competence is considered as a positive predictor of intrinsic motivation and thus contributes positively to the enjoyment of "Endless Universe" (Ryan et al. 2006). On the other hand, perceived pressure is considered as a negative predictor of intrinsic motivation (Wilde et al. 2009). However, the increase in perceived pressure might also be related to a higher immersion, an enhanced focus on the task and thus a higher sense of flow (Harms et al. 2015; Csikszentmihalyi 1997). Therefore, the significant increase in perceived pressure might be perceived both negatively and positively and should be studied in future work. The fact that perceived choice is significantly lower when using "Endless Universe" might be related to the introduction of gameful design elements, which establish certain goals and norms which might establish more guidance and thus lead to less choice. Taking **R5** and **R6** together, we consider **H1b: Users perceive running as more enjoyable using "Endless Universe"** as partially supported. Since no significant effects were found regarding positive or negative affect, **H1c: Users have stronger affective experiences with "Endless Universe"** is not supported. These mixed results regarding user experience (**H1b**, **H1c**) might be related to interpersonal differences in the perception of gameful design elements, which have been shown by previous research (Tondello et al. 2016; Orji et al. 2018) and as part of the online study (**R1**,**R2**).

Therefore, as a next step, we analyzed whether such interpersonal differences could be explained by considering the behavioral intention and Hexad user type of participants. Regarding behavioral intentions (**H2**), we did not find any significant effects between Low- and High-SoC users regarding distance improvement or perceived exertion. Thus, **H2a: The improvement in distance is higher for High-SoC users** is not supported, given our data. A potential reason could be related to observer effects, i.e., the effect that participants act more ethically, more conscientiously or more efficiently when being observed (Monahan and Fisher 2010). During the experiment, one researcher was in the same room as the participant. This might have affected Low-SoC users more than High-SoC users to improve their performance, since Low-SoC users might have wanted to avoid drawing attention to the fact that they were performing worse than others. Consequently, they might have powered more in the baseline, but could not improve in the intervention. Regarding **H2b: High-SoC users perceive "Endless Universe" as more enjoyable**, we found no significant differences on the respective IMI factors (enjoyment, competence).

Thus, this hypothesis cannot be supported. However, it should be noted that the sample size to compare the Low- and High-SoC users was rather small (10 participants per group), which means that the chance of not finding small- to medium-sized effects is relatively high. Therefore, we acknowledge that the absence of significant effects (**H2a**, **H2b**) should not be seen as supporting evidence for the respective null hypotheses. Descriptively, both factors were considerably higher in the High-SoC group, which might suggest that a significant difference could have been found with more participants in each group and that the size of the actual effect was too small to be detected with a total N of 20. Finally, we found a significant increase in both positive and negative affects among High-SoC users (**R7**). This supports **H2c: High-SoC users have stronger affective experiences**.

The fact that positive affect was significantly higher when using "Endless Universe" supports that tailoring a gameful application to the SoC of users positively affects the user experience. Given that also negative affect was significantly higher when using "Endless Universe," these results need to be interpreted more carefully. There is a lot of criticism of considering positive and negative affect as polar opposites (Russell and Carroll 1999). Research has found strong positive correlations between the latent factors of positive and negative affect. Also, the instrument that we used, PANAS, actually does not measure opposite affective experiences (as the names of the latent variables might suggest) (Russell and Carroll 1999). In fact, the items of positive affect were chosen to represent a latent variable (named positive affect), which is defined as activation plus pleasantness. The negative items were chosen to represent a latent construct (named negative affect) defined as activation plus unpleasantness (Watson et al. 1988; Russell and Carroll 1999). This shows that these two latent constructs are not opposite on activation, which ultimately means that they are not opposite. We also found supporting evidence of this effect in our data. When analyzing a potential correlation between positive and negative affects, which should be strongly negative, assuming a bipolarity of both latent variables, we found that there exists an insignificant positive correlation between positive and negative affects (Kendall's  $\tau = 0.25$ ,  $p=0.17$ ). This supports the assumption that activation was the deciding cause for the increase in negative affect, instead of unpleasantness. This assumption is further supported by research showing that negative affect can lead to a positive user experience, especially within gameful systems (Bopp et al. 2016). Thus, we conclude that the increase in negative affect seems to be related to higher arousal and activation. Considering a significant increase in positive affect, this allows to interpret the results related to affective experience in a way that supports the assumption of a better user experience when using "Endless Universe."

Regarding Hexad user types, we found no evidence for **H3a: The improvement in distance is higher for AC, PL, SO**. Considering that correlations between gameful design elements and Hexad user types using self-reported measures were rather weak (Tondello et al. 2016; Orji et al. 2018), the absence of significant correlations between the improvement in distance and these Hexad user types might be related to the low sample size and the resulting low test power. Future work should consider a higher number of participants in order to be able to detect small- to medium-sized correlations. However, it should be noted that we found a negative correlation

between the Disruptor and distance improvement (**R9**), suggesting that Hexad user types seem to have an actual effect on performance. Furthermore, we found that perceived competence was positively correlated to the Socialiser user type (**R8**) and that perceived pressure was negatively correlated to the Free Spirit user type (**R10**). **R8** can be interpreted as partially supporting evidence for **H3b: AC, PL, SO perceive “Endless Universe” as more enjoyable**. For Players and Achievers, no significant correlations were found, meaning that **H3b** is only supported for the Socialiser. However, taking also **R10** into account, the importance of Hexad user types as a factor moderating the user experience in a gameful fitness application is strengthened and should be investigated further in upcoming interventions. Lastly, we did not find significant correlations regarding affective experiences; thus, **H3c: AC, PL, SO have stronger affective experiences** is not supported, given our data. Potentially, this could indicate that tailoring for Hexad user types affects measures related to motivation and the perception of gameful design elements more than the measures related to emotional responses evoked by those gameful design elements. However, this needs to be investigated in future work. Also, it should be noted that we used concrete implementations of gameful design elements, implying that certain design decisions needed to be made, which in turn might have affected the perception of the gameful design elements.

Finally, regarding the question of whether gameful fitness systems should be personalized using behavior change intentions or Hexad user types, the short answer based on our findings is “most probably yes.” No evidence was found for personalization affecting immediate performance-related measures (**H2a, H3a**). However, we found significant positive effects on the user experience of participants (**H2c, H3b**). This indicates that personalization using behavior change intentions or Hexad user types might affect the performance or behavior of users in the long run, i.e., the improved user experience might lead to improved retention rates and participants might be more motivated to keep increasing their physical activity. Consequently, beneficial effects on the performance and behavior of users are expected when conducting studies over a longer time span. This is an important direction that should be followed in future work.

## 9 Conclusion and future work

We investigated behavior change intentions and Hexad user types as factors to personalize gameful fitness systems. First, we created storyboards explaining twelve commonly used gameful design elements in the context of encouraging walking. These storyboards were made publicly available for replication purposes and to allow their usage for future studies. After showing that these storyboards explain the intended gameful design elements in a qualitative pre-study, we used them in an online study to explore whether behavioral intentions and Hexad user types moderate the perceived persuasiveness of them. The findings of this study supported the importance of both factors for personalization, since we found significant differences between Low- and High-SoC users for several gameful design elements as well as replicated previously found correlations between Hexad user

types and gameful design elements in the physical activity context. Next, we used these findings to conceptualize and implement a personalized, gameful system encouraging physical activity on a treadmill. This system was used in a laboratory study to analyze whether personalization based on the findings of the storyboards-based online study has any effect on the users' performance, enjoyment and affective experiences while running. This is important, as it tackles the problem that storyboards-based studies (which have been mostly used in past research) do not allow users to interact and experience gameful design elements. Our laboratory study showed that personalization based on behavior change intentions and Hexad user types does not seem to affect the immediate performance. However, significant effects were found on the user experience, i.e., on motivational aspects and on affective experiences. This improved user experience suggests that the behavior and performance of users might be positively affected in the long run, when personalizing gameful fitness applications.

Therefore, future work should investigate effects of personalization in user studies over a longer time span. This would allow to investigate whether the increased user experience leads to an increased performance over time. As an alternative, giving users the chance to decide whether they would want to use the system regularly would have provided more insights on potential effects on the behavior of users and could be studied in future work, too. Also, in-the-wild studies should be conducted to alleviate potential observer effects and study the impact of personalization in a more natural setting. This would shed light on whether the absence of effects when tailoring for the SoC is related to an observer bias in the laboratory setting. It is also important to investigate the impact of personalization on the user experience further to better understand the reasons for the effects that were found in this article, e.g., if the increase in the pressure factor of the IMI when personalizing the gameful application is related to an increased immersion. Also, more participants should be recruited, to be able to find effects with smaller effect sizes (which were reported in previous research). This would decrease the chance of type II errors, i.e., stating that there is no effect when a true effect is to be found, especially regarding potential correlations between Hexad user types and performance-related measures. Additionally, future work should investigate whether our findings can be replicated in different health-related contexts or using different gameful applications, to analyze the external validity of our results. We also recommend to investigate different ways of operationalizing context-related motivation, besides behavior change intentions, as this factor is neglected in personalization research for gameful applications. Finally, it seems worthwhile to study different measures such as flow and immersion to better understand the effects of personalization on the user experience. Related to this, more objective variables such as psychophysiological measures could be taken into account to better understand the various user experience-related findings of our studies.

**Acknowledgements** This work is partially supported by the German Federal Ministry of Education and Research (16SV8364).

**Funding** Open Access funding enabled and organized by Projekt DEAL.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Aldenaini, N., Alqahtani, F., Orji, R., Srinivas, S.: Trends in persuasive technologies for physical activity and sedentary behavior: a systematic review. *Front. Artif. Intell. J. Hum. Learn. Behav. Change* (2020). <https://doi.org/10.3389/frai.2020.00007>
- Altmeyer, M., Lessel, P.: The importance of social relations for well-being change in old age - do game preferences change as well? In: *Proceedings of the Positive Gaming: Workshop on Gamification and Games for Wellbeing* Amsterdam, The Netherlands (2017)
- Altmeyer, M., Lessel, P., Krüger, A.: Investigating gamification for seniors aged 75 +. In: *Proceedings of the 2018 Designing Interactive Systems Conference (DIS '18)*, ACM, New York, New York, USA, pp. 453–458. <https://doi.org/10.1145/3196709.3196799>, <http://dl.acm.org/citation.cfm?doid=3196709.3196799> (2018a)
- Altmeyer, M., Lessel, P., Sander, T., Krüger, A.: Extending a gamified mobile app with a public display to encourage walking. In: *Proceedings of the 22nd International Academic Mindtrek Conference on—Mindtrek '18*, pp. 20–29. <https://doi.org/10.1145/3275116.3275135>, <http://dl.acm.org/citation.cfm?doid=3275116.3275135> (2018b)
- Altmeyer, M., Lessel, P., Muller, L., Krüger, A.: Combining behavior change intentions and user types to select suitable gamification elements for persuasive fitness systems. In: *International Conference on Persuasive Technology*, Springer (2019)
- Armstrong, R.A.: When to use the bonferroni correction. *Ophthal. Physiol. Opt. J. Br. College Ophthal. Opt.* **34**(5), 502–508 (2014). <https://doi.org/10.1111/opo.12131>
- Birk, M.V., Friehs, M.A., Mandryk, R.L.: Age-based preferences and player experience: a crowd-sourced cross-sectional study. In: *Proceedings of the Annual Symposium on Computer–Human Interaction in Play—CHI PLAY '17*, pp. 157–170. <https://doi.org/10.1145/3116595.3116608>, <http://dl.acm.org/citation.cfm?doid=3116595.3116608> (2017)
- Bishara, A.J., Hittner, J.B.: Testing the significance of a correlation with nonnormal data: comparison of pearson, spearman, transformation, and resampling approaches. *Psychol. Methods* **17**(3), 399–417 (2012). <https://doi.org/10.1037/a0028087>
- Bopp, J.A., Mekler, E.D., Opwis, K.: Negative emotion, positive experience? In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*, pp. 2996–3006. <https://doi.org/10.1145/2858036.2858227>, <http://dl.acm.org/citation.cfm?doid=2858036.2858227> (2016)
- Borg, G.: Perceived exertion as an indicator of somatic stress. *Scand. J. Rehab. Med.* (1970)
- Bravata, D., Smith-Spangler, C., Sundaram, V., Gienger, A., Lin, N., Lewis, R., Stave, C., Olkin, I., Sirard, J.: Using pedometers to increase physical activity and improve health: a systematic review. *JAMA* **298**, 19 (2007). <https://doi.org/10.1001/jama.298.19.2296>
- Busch, M., Mattheiss, E., Hochleitner, W., Hochleitner, C., Lankes, M., Fröhlich, P., Orji, R., Tscheligi, M.: Using player type models for personalized game design: an empirical investigation. *Interact. Design Archit.* **28**(1), 145–163 (2016)
- Busch, M., Mattheiss, E., Orji, R., Fröhlich, P., Lankes, M., Tscheligi, M.: Player type models: towards empirical validation. In: *Conference on Human Factors in Computing Systems—Extended Abstracts Proceedings 07-12-May-(October 2017)*, pp. 1835–1841. <https://doi.org/10.1145/2851581.2892399> (2016b)

- Buttussi, F., Chittaro, L., Ranon, R., Verona, A.: Adaptation of graphics and gameplay in fitness games by exploiting motion and physiological sensors. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 4569 LNCS, pp. 85–96. [https://doi.org/10.1007/978-3-540-73214-3\\_8](https://doi.org/10.1007/978-3-540-73214-3_8) (2007)
- Cercos R, Mueller FF (2013) Watch your Steps: Designing a Semi-Public Display to Promote Physical Activity. Proceedings of The 9th Australasian Conference on Interactive Entertainment: Matters of Life and Death p 2, <https://doi.org/10.1145/2513002.2513016>
- Cham, S., Algashami, A., McAlaney, J., Stefanidis, A., Phalp, K., Ali, R.: Goal setting for persuasive information systems: five reference checklists. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 11433 LNCS, pp. 237–253. [https://doi.org/10.1007/978-3-030-17287-9\\_20](https://doi.org/10.1007/978-3-030-17287-9_20) (2019)
- Chen, Y., Pu, P.: HealthyTogether: exploring social incentives for mobile fitness applications. In: Proceedings of the Second International Symposium of Chinese CHI on—Chinese CHI '14 (April 2014), pp. 25–34. <https://doi.org/10.1145/2592235.2592240> (2014)
- Consolvo, S., Everitt, K., Smith, I., Landay, J.A.: Design requirements for technologies that encourage physical activity. In: Proceedings of the SIGCHI conference on Human Factors in Computing Systems—CHI '06, p. 457. <https://doi.org/10.1145/1124772.1124840>, <http://portal.acm.org/citation.cfm?doid=1124772.1124840.9809069v1> (2006)
- Consolvo, S., Klasnja, P., Mcdonald, D.W., Avrahami, D., Froehlich, J., Legrand, L., Libby, R., Mosher, K., Landay, J.A.: Flowers or a robot army? Encouraging awareness & activity with personal, mobile displays. In: Proceedings of the 10th International Conference on Ubiquitous Computing, pp. 54–63. <https://doi.org/10.1145/1409635.1409644> (2008a)
- Consolvo, S., Mcdonald, D.W., Toscos, T., Chen, M.Y., Froehlich, J., Harrison, B., Klasnja, P., Lamarca, A., Legrand, L., Libby, R., Smith, I., Landay, J.A.: Activity sensing in the wild: a field trial of Ubi-Fit garden. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1797–1806. <https://doi.org/10.1145/1357054.1357335> (2008b)
- Csikszentmihalyi, M.: Finding Flow: The Psychology of Engagement with Everyday Life. Basic Books, New York (1997)
- Cuzzort, S., Starner, T.: AstroWheelie: a wheelchair based exercise game. In: Proceedings—International Symposium on Wearable Computers, ISWC, pp. 113–114. <https://doi.org/10.1109/ISWC.2008.4911599> (2008)
- Deterding, S., Dixon, D., Khaled, R., Nacke, L.: From game design elements to gamefulness: defining gamification. In: Proceedings of the 15th International Academic MindTrek Conference ACM, 2011, pp. 9–15. <https://doi.org/10.1145/2181037.2181040> (2011)
- Doyle, J., Kelly, D., Caulfield, B.: Design considerations in therapeutic exergaming. In: 2011 5th International Conference on Pervasive Computing Technologies for Healthcare, PervasiveHealth 2011, pp. 389–393. <https://doi.org/10.4108/icst.pervasivehealth.2011.246115> (2011a)
- Doyle, J., Kelly, D., Patterson, M., Caulfield, B.: The effects of visual feedback in therapeutic exergaming on motor task accuracy. In: International Conference on Virtual Rehabilitation 2011. <https://doi.org/10.1017/CBO9781107415324.004>, [arXiv:1011.1669v3](https://arxiv.org/abs/1011.1669v3) (2011b)
- Drozd, F., Lehto, T., Oinas-Kukkonen, H.: Exploring perceived persuasiveness of a behavior change support system: a structural model. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 7284 LNCS, pp. 157–168. [https://doi.org/10.1007/978-3-642-31037-9\\_14](https://doi.org/10.1007/978-3-642-31037-9_14) (2012)
- Epstein, D., Ping, A., Fogarty, J., Munson, S.A.: A lived informatics model of personal informatics. In: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing, pp. 731–742. <https://doi.org/10.1145/2750858.2804250>, <http://dx.doi.org/10.1145/2750858.2804250> (2015)
- Finkelstein, S.L., Nickel, A., Barnes, T., Suma, E.A.: Astrojumper: designing a virtual reality exergame to motivate children with autism to exercise. In: Proceedings—IEEE Virtual Reality, pp. 267–268. <https://doi.org/10.1109/VR.2010.5444770> (2010)
- Fogg, B.J.: Persuasive Technology: Using Computers to Change What We Think and Do. Morgan Kaufmann Publishers 10(1145/764008), 763957 (2002)
- Gui, X., Chen, Y., Caldeira, C., Xiao, D., Chen, Y.: When fitness meets social networks: investigating fitness tracking and social practices on WeRun. In: Conference on Human Factors in Computing Systems—Proceedings 2017-May, pp. 1647–1659. <https://doi.org/10.1145/3025453.3025654> (2017)
- Halko, S., Kientz, J.A.: Personality and persuasive technology: an exploratory study on health-promoting mobile applications. In: Persuasive Technology, 5th International Conference, PERSUASIVE 2010

- (May), p. 299. <https://doi.org/10.1007/978-3-642-13226-1>, <https://www.mendeley.com/viewer/?fileId=d=507818b9-b728-6f89-722a-cb184ad5ab0e&documentId=80290fb9-00fd-3e4e-9ed6-a930bd5e5f30> (2010)
- Hallifax, S., Serna, A., Marty, J.C., Lavoué, G., Lavoué, E.: Factors to consider for tailored gamification. In: Proceedings of the Annual Symposium on Computer–Human Interaction in Play—CHI PLAY '19 (2019)
- Hamari, J., Sarsa, H.: Does gamification work?: A literature review of empirical studies on gamification. In: Hawaii International Conference on System Sciences, pp. 3025–3034. <https://doi.org/10.1109/HICSS.2014.377> (2014)
- Hamari, J., Koivisto, J., Pakkanen, T.: Do persuasive technologies persuade?: a review of empirical studies. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 8462 LNCS, pp. 118–136. [https://doi.org/10.1007/978-3-319-07127-5\\_11](https://doi.org/10.1007/978-3-319-07127-5_11), 9780201398298 (2014)
- Harms J, Biegler, S., Wimmer, C., Kappel, K., Grechenig, T.: Gamification of online surveys: design process, case study, and evaluation. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 9296(Section 4), pp. 219–236. [https://doi.org/10.1007/978-3-319-22701-6\\_16](https://doi.org/10.1007/978-3-319-22701-6_16) (2015)
- Howell, D.: Statistical methods for psychology. Duxbury (2020). <https://doi.org/10.2307/2348956>
- Jia, Y., Xu, B., Karanam, Y., Voids, S.: Personality-targeted gamification: a survey study on personality traits and motivational affordances. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems—CHI '16, pp. 2001–2013. <https://doi.org/10.1145/2858036.2858515>, <http://dl.acm.org/citation.cfm?doid=2858036.2858515> (2016)
- Jia, Y., Liu, Y., Yu, X., Voids, S.: Designing leaderboards for gamification: perceived differences based on user ranking, application domain, and personality traits. In: Conference on Human Factors in Computing Systems—Proceedings 2017-May(May), pp. 1949–1960. <https://doi.org/10.1145/3025453.3025826> (2017)
- Kappen, D.L., Nacke, L.E., Gerling, K.M., Tsotsos, L.E.: Design strategies for gamified physical activity applications for older adults. In: Hawaii International Conference on System Sciences, pp. 1309–1318. <https://doi.org/10.1109/HICSS.2016.166> (2016)
- King, D.E., Mainous, A.G., Carnemolla, M., Everett, C.J.: Adherence to healthy lifestyle habits in US adults, 1988–2006. *Am. J. Med.* **122**(6), 528–534 (2009). <https://doi.org/10.1016/j.amjme.2008.11.013>
- Klein, M.C., Manzoor, A., Mollie, J.S.: Active2Gether: a personalized m-health intervention to encourage physical activity. *Sensors (Switzerland)* **17**(6), 1–16 (2017). <https://doi.org/10.3390/s17061436>
- Koivisto, J., Hamari, J.: Gamification of physical activity: a systematic literature review of comparison studies. *Int. GamiFIN Conf.* **2359**, 106–117 (2019)
- Kotsopoulos, D., Bardaki, C., Lounis, S., Pramataris, K.: Employee profiles and preferences towards IoT-enabled gamification for energy conservation. *Int. J. Serious Games* **5**(2), 65–85 (2018)
- Li, I., Dey, A., Forlizzi, J.: A stage-based model of personal informatics systems. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, p. 557. <https://doi.org/10.1145/1753326.1753409>, <http://portal.acm.org/citation.cfm?doid=1753326.1753409> (2010)
- Lin, J.J., Mamykina, L., Lindtner, S., Delajoux, G., Strub, H.B.: Fish'n'Steps: encouraging physical activity with an interactive computer game. In: UbiComp 2006: Ubiquitous Computing, pp. 261–278. [https://doi.org/10.1007/11853565\\_16](https://doi.org/10.1007/11853565_16), [http://link.springer.com/libproxy.usc.edu/chapter/10.1007/11853565\\_16%5Cnfiles/3571/Lin\\_et~al.-2006- Fish'n?Steps Encouraging Physical Activity with a.pdf%5Cnfiles/3588/10.html](http://link.springer.com/libproxy.usc.edu/chapter/10.1007/11853565_16%5Cnfiles/3571/Lin_et~al.-2006- Fish'n?Steps Encouraging Physical Activity with a.pdf%5Cnfiles/3588/10.html), arXiv:1011.1669v3 (2006)
- Locke, E.A., Latham, G.P.: Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *Am. Psychol.* **57**(9), 705–717 (2002). <https://doi.org/10.1037/0003-066X.57.9.705>
- Marcus, B.H., Selby, V.C., Niaura, R.S., Rossi, J.S., Lavalley, D., Williams, J.M., Jones, M.V.: Self-efficacy and the stages of exercise behavior change. In: *Research Quarterly for Exercise and Sport*, pp. 93–109. <https://doi.org/10.1080/02701367.1992.10607557> (2008)
- Marczewski, A.: Even ninja monkeys like to play: gamification. CreateSpace Independent Publishing Platform, Game Thinking and Motivational Design (2015)
- McAuley, E.D., Duncan, T., Tammen, V.V.: Psychometric properties of the intrinsic motivation inventory in a competitive sport setting: a confirmatory factor analysis. *Res. Q. Exercise Sport* **60**(1), 48–58 (1989). <https://doi.org/10.1080/02701367.1989.10607413>
- McCrae, R.R., John, O.P.: An introduction to the five-factor model and its applications. *J. Pers.* **60**(2), 175–215 (1992). <https://doi.org/10.1111/j.1467-6494.1992.tb00970.x>

- McHugh, M.L.: Interrater reliability: the kappa statistic. *Biochemia Medica* 276–282 (2012)
- Mekler, E.D., Brühlmann, F., Tuch, A.N., Opwis, K.: Towards understanding the effects of individual gamification elements on intrinsic motivation and performance. *Comput. Hum. Behav.* **71**, 525–534 (2017). <https://doi.org/10.1016/j.chb.2015.08.048>
- Meyer, J., Beck, E., Von Holdt, K., Gansefort, D., Brand, T., Zeeb, H., Boll, S.: Actistairs: design and acceptance of a technology-based intervention to advocate stair-climbing in public spaces. In: *HealthMedia 2018—Proceedings of the 3rd International Workshop on Multimedia for Personal Health and Health Care*, co-located with MM 2018, pp. 59–66. <https://doi.org/10.1145/3264997.96.3264997> (2018)
- Middelweerd, A., Mollée, J., Klein, M., Manzoor, A., Brug, J., te Velde, S.J.: The use and effects of an app-based physical activity intervention “Active2Gether” in young adults: quasi-experimental trial. *JMIR Formative Research* **4**(1), e12538 (2020). <https://doi.org/10.2196/12538>
- Miller, A.D., Mynatt, E.D.: StepStream: a school-based pervasive social fitness system for everyday adolescent health. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (January), pp. 1–10. <https://doi.org/10.1145/2556288.2557190> (2014)
- Monahan, T., Fisher, J.A.: Benefits of observer effects: lessons from the field. *Qual. Res.* (2010). <https://doi.org/10.1177/1468794110362874>
- Mora, A., Tondello, G.F., Nacke, L.E., Arnedo-Moreno, J.: Effect of personalized gameful design on student engagement. In: *IEEE Global Engineering Education Conference, EDUCON 2018-April*, pp. 1925–1933. <https://doi.org/10.1109/EDUCON.2018.8363471> (2018)(2018)
- Mullan, E., Markland, D.: Variations in self-determination across the stages of change for exercise in adults. *Motiv. Emot.* **21**(4), 349–362 (1997). <https://doi.org/10.1023/A:1024436423492>
- Nacke, L.E., Bateman, C., Mandryk, R.L.: BrainHex: a neurobiological gamer typology survey. *Entertain. Comput.* **5**(1), 55–62 (2014). <https://doi.org/10.1016/j.entcom.2013.06.002>
- Nakajima, T., Lehdonvirta, V.: Designing motivation using persuasive ambient mirrors. *Pers. Ubiquitous Comput.* **17**(1), 107–126 (2013). <https://doi.org/10.1007/s00779-011-0469-y>
- Niess, J., Woźniak, P.W.: Supporting meaningful personal fitness: the tracker goal Evolution Model. In: *CHI Conference on Human Factors in Computing Systems Proceedings (CHI 2018) 2018-April*, pp. 1–12. <https://doi.org/10.1145/3173574.3173745> (2018)
- Orji, R., Mandryk, R.L., Vassileva, J., Gerling, K.M.: Tailoring persuasive health games to gamer type. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 2467–2476 (2013)
- Orji, R., Vassileva, J., Mandryk, R.L.: Modeling the efficacy of persuasive strategies for different gamer types in serious games for health. In: *User Modeling and User-Adapted Interaction*, pp. 453–498. <https://doi.org/10.1007/s11257-014-9149-8> (2014)
- Orji, R., Mandryk, R.L., Vassileva, J.: Gender, age, and responsiveness to cialdini’s persuasion strategies. In: *International Conference on Persuasive Technology 9072(June)*. <https://doi.org/10.1007/978-3-319-20306-5> (2015)
- Orji, R., Nacke, L.E., Di Marco, C.: Towards personality-driven persuasive health games and gamified systems. In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems—CHI ’17*, pp. 1015–1027. <https://doi.org/10.1145/3025453.3025577>, <http://dl.acm.org/citation.cfm?doid=3025453.3025577> (2017)
- Orji, R., Tondello, G.F., Nacke, L.E.: Personalizing persuasive strategies in gameful systems to gamification user types. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems—CHI ’18*. <https://doi.org/10.1145/3173574.3174009> (2018)
- Oyibo, K., Vassileva, J.: Investigation of the moderating effect of culture on users’ susceptibility to persuasive features in fitness applications. *Information* (2019). <https://doi.org/10.3390/info10110344>
- Oyibo, K., Orji, R., Vassileva, J.: Investigation of the persuasiveness of social influence in persuasive technology and the effect of age and gender. In: *Persuasive Technology (PPT’17) Workshop* (2017)
- Prochaska, J.O., Velicer, W.F.: The transtheoretical change model of health behavior. *Am. J. Health Promot.* **12**(1), 38–48 (1997). <https://doi.org/10.4278/0890-1171-12.1.38>
- Rajaratnam, S.M.W., Arendt, J.: Health in a 24-h society. *Lancet* **358**, 999–1005 (2001). [https://doi.org/10.1016/S0140-6736\(01\)06108-6](https://doi.org/10.1016/S0140-6736(01)06108-6)
- Russell, J.A., Carroll, J.M.: On the bipolarity of positive and negative affect. *Psychol. Bull.* **125**(1), 3–30 (1999)
- Ryan, R.M.: Control and information in the intrapersonal sphere: an extension of cognitive evaluation theory. *J. Pers. Soc. Psychol.* **43**(3), 450–461 (1982)

- Ryan, R.M., Deci, E.L.: Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *Am. Psychol.* **55**(1), 68–78 (2000). <https://doi.org/10.1037/0003-066X.55.1.68>
- Ryan, R.M., Rigby, C.S., Przybylski, A.: The motivational pull of video games: a self-determination theory approach. *Motivation and Emotion* **30**(4):347–363, <https://doi.org/10.1007/s11031-006-9051-8>, [arXiv:1011.1669v3](https://arxiv.org/abs/1011.1669v3) (2006)
- Saksono, H., Ranade, A., Kamarthi, G., Castaneda-Sceppa, C., Hoffman, J.A., Wirth, C., Parker, A.G.: Spaceship launch: designing a collaborative exergame for families. In: *CSCW 2015—Proceedings of the 2015 ACM International Conference on Computer-Supported Cooperative Work and Social Computing*, pp. 1776–1787. <https://doi.org/10.1145/2675133.2675159> (2015)
- Seaborn, K., Fels, D.: Gamification in theory and action: a survey. *Int. J. Hum. Comput. Stud.* **74**, 14–31 (2015). <https://doi.org/10.1016/j.ijhcs.2014.09.006>
- Tondello, G., Premasukh, H., Nacke, L.: A theory of gamification principles through goal-setting theory. In: *Proceedings of the 51st Hawaii International Conference on System Sciences (HICSS)* (January). [https://doi.org/10.1057/9780230594647\\_8](https://doi.org/10.1057/9780230594647_8), <http://hdl.handle.net/10125/50027> (2018a)
- Tondello, G.F.: Dynamic personalization of gameful interactive systems. PhD thesis, University of Waterloo. <http://hdl.handle.net/10012/14807> (2019)
- Tondello, G.F., Wehbe, R.R., Diamond, L., Busch, M., Marczewski, A., Nacke, L.E.: The gamification user types hexad scale. In: *The ACM SIGCHI Annual Symposium on Computer–Human Interaction in Play—CHI PLAY '16*. <https://doi.org/10.1145/2967934.2968082> (2016)
- Tondello, G.F., Mora, A., Nacke, L.E.: Elements of gameful design emerging from user preferences. In: *Proceedings of the Annual Symposium on Computer–Human Interaction in Play—CHI PLAY '17*, pp. 129–142. <https://doi.org/10.1145/3116595.3116627>, <http://dl.acm.org/citation.cfm?doid=3116595.3116627> (2017)
- Tondello, G.F., Mora, A., Marczewski, A., Nacke, L.E.: Empirical validation of the gamification user types hexad scale in english and Spanish. *Int. J. Hum. Comput. Stud.* (2018b). <https://doi.org/10.1016/j.ijhcs.2018.10.002>
- Truong, K.N., Hayes, G.R., Abowd, G.: Storyboarding: an empirical determination of best practices and effective guidelines. In: *Proceedings of the 6th ACM Conference on Designing Interactive Systems*, ACM, pp. 12–21. <https://doi.org/10.1145/1142405.1142410>, <http://portal.acm.org/citation.cfm?doid=1142405.1142410> (2006)
- Tyack, A., Mekler, E. : Self-determination theory in HCI games research: current uses and open questions. In: *CHI 2020—Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/3313831.3376723> (2020)
- Von Ahn, L., Dabbish, L.: Designing games with a purpose. *Commun. ACM* **51**(8), 58–67 (2008). <https://doi.org/10.1145/1378704.1378719>
- Walker, D.: Converting Kendall’s Tau for correlational or meta-analytic analyses. *J. Mod. Appl. Stat. Methods* **2**(2), 525–530 (2003)
- Watson, D., Clark, L.A., Tellegen, A.: Development and validation of brief measures of positive and negative affect: the PANAS scales. *J. Pers. Soc. Psycholog.* **54** (6), 1063–1070 (1988). <https://doi.org/10.1037/0022-3514.54.6.1063>, [arXiv:1011.1669v3](https://arxiv.org/abs/1011.1669v3)
- Werbach, K., Hunter, D.: *For the Win: How Game Thinking can Revolutionize your Business*. Wharton Digital Press (2012)
- Wilde, M., Bätz, K., Kovaleva, A., Urhahne, D.: Testing a short scale of intrinsic motivation. *Zeitschrift für Didaktik der Naturwissenschaften* **15**(31–45), 2404161 (2009)
- Xiao, J.J., O’Neill, B., Prochaska, J.M., Kerbel, C.M., Brennan, P., Bristow, B.J.: A consumer education programme based on the transtheoretical model of change. *Int. J. Consum. Stud.* **28**(1), 55–65 (2004). <https://doi.org/10.1111/j.1470-6431.2004.00334.x>
- Zuckerman, O., Gal-Oz, A.: Deconstructing gamification: evaluating the effectiveness of continuous measurement, virtual rewards, and social comparison for promoting physical activity. *Pers. Ubiquitous Comput.* **18**(7), 1705–1719 (2014). <https://doi.org/10.1007/s00779-014-0783-2>

**Maximilian Altmeyer** is a researcher at the German Research Center for Artificial Intelligence and a PhD student at Saarland University. In his research, he investigates the effects of gamification on the motivation of users in behavior change contexts, whether motivation can be enhanced through personalized gameful design, which factors are relevant to consider for personalization and which effects personalization has on users in gameful applications. Max published several papers at top conferences such as CHI and CHI Play.

**Pascal Lessele** is a postdoctoral researcher at the German Research Center for Artificial Intelligence. His research interests are situated in gaming-related contexts such as gamification and game live-streams. In these contexts, he investigates specifically whether offering more autonomy for users is reasonable and which benefits can be achieved through this. Understanding ways to customize and personalize these experiences is another important research direction for him, for example, here, he does research on “bottom-up” gamification, i.e., enabling users to decide on their gamification at the system’s runtime. Pascal published several papers at top conferences such as CHI and CHI Play.

**Subhashini Jantwal** is a former graduate student at Saarland University, now working as a user experience designer. She received her Bachelor’s degree in Computer Science Engineering from Uttarakhand Technical University. After developing a keen interest in human–computer interaction, she pursued her Master’s in Media Informatics from Saarland University with a focus on design and implementation of gameful software solutions. As a part of her Master’s thesis, she worked under the Media Informatics chair at the German Research Center for AI, at Saarbrücken. Her interests include behavioral sciences, computer science and media design pertaining to designing innovative human–machine interactions.

**Linda Muller** is a former undergraduate student at Saarland University, where she received her Bachelor’s degree in Media Informatics. Her thesis focused on personalized gameful design in health-related contexts and the role of Hexad user types and behavior change intentions. Her research interests include human–computer interaction, games and gameful design.

**Florian Daiber** is a postdoctoral researcher at the Ubiquitous Media Technology Lab (UMTL) at the German Research Center for Artificial Intelligence (DFKI) in Saarbrücken, Germany. His main research is in the field of human–computer interaction, 3D user interfaces and ubiquitous sports technologies. Currently, Florian is mainly involved in projects on 3D interaction in mixed realities and wearable technologies for sports and health. Florian is co-organizing the UbiComp Workshop on Ubiquitous Computing in the Mountains, the Workshop on Understanding human–computer interaction in Outdoor Recreation, the workshop on everyday proxy objects for virtual reality and the workshop on cross-reality (XR) interaction. He served on various program committees in the field of human–computer interaction, e.g., ACM ETRA, IEEE VR, and ACM Symposium on Computer–Human Interaction in Play (CHI PLAY).

**Antonio Krüger** is a CEO and scientific director of the German Research Center for Artificial Intelligence GmbH (DFKI) and head of the department “Cognitive Assistants” at DFKI. He is a full professor for Computer Science at Saarland University (since 2009), Head of the Ubiquitous Media Technology Lab and scientific director of the Innovative Retail Laboratory (IRL) at DFKI. From 2004 to 2009, he was a professor of computer science and geoinformatics at the University of Münster and acted as the managing director of the institute for geoinformatics. He studied computer science and economics at Saarland University and finished his PhD in 1999 as a member of the Saarbrücken graduate school of Cognitive Science. Antonio has published more than 200 scientific articles and papers in internationally recognized journals and conferences and is member of several steering committees, editorial boards and scientific advisory committees.