International Journal of Human-Computer Studies 127 (2019) 95–111

Contents lists available at ScienceDirect

International Journal of Human-Computer Studies

journal homepage: www.elsevier.com/locate/ijhcs

Empirical validation of the Gamification User Types Hexad scale in English and Spanish

Gustavo F. Tondello[a,*], Alberto Mora[b], Andrzej Marczewski[c], Lennart E. Nacke[d]

[a] HCI Games Group, Games Institute, and Cheriton School of Computer Science, University of Waterloo, 200 University Avenue West, Waterloo, ON N2L 3G1, Canada
[b] Estudis d'Informàtica, Multimedia i Telecomunicació, Universitat Oberta de Catalunya, Barcelona, Spain
[c] Gamified UK, New Haw, Surrey, UK
[d] HCI Games Group, Games Institute, Department of Drama and Speech Communication, Cheriton School of Computer Science, Department of Systems Design Engineering, and Stratford School of Interaction Design and Business, University of Waterloo, Waterloo, ON, Canada

ARTICLE INFO

Keywords:
Gamification
Gameful design
User Types
Hexad

ABSTRACT

Gamification, the use of game elements in non-game systems, is now established as a relevant research field in human-computer interaction (HCI). Several empirical studies have shown that gameful interventions can increase engagement and generate desired behavioral outcomes in HCI applications. However, some inconclusive results indicate that we need a fuller understanding of the mechanisms and effects of gamification. The Gamification User Types Hexad scale allows us to parse different user motivations in participants' interactions with gameful applications, which are measured using a self-report questionnaire. Each user type represents a style of interaction with gameful applications, for example, if the interactions are more focused on achievements, socialization, or rewards. Thus, by scoring an individual in each one of the user types of the Hexad model, we can establish a profile of user preferences for gameful interactions. However, we still lack a substantial empirical validation of this scale. Therefore, we set out to validate the factor structure of the scale, in both English and Spanish, by conducting three studies, which also investigated the distribution of the Hexad's user types in the sample. Our findings support the structural validity of the scale, as well as suggesting opportunities for improvement. Furthermore, we demonstrate that some user types are more common than others and that gender and age correlate with a person's user types. Our work contributes to HCI research by further validating the utility of the Gamification User Types Hexad scale, potentially affording researchers a deeper understanding of the mechanisms and effects of gameful interventions.

1. Introduction

'Gamification' is the use of game design elements in non-game contexts (Deterding et al., 2011). It is often used to create gameful applications or systems (i.e., applications that use game design elements) aimed at encouraging specific behaviors and altering behavior patterns. These applications are found in a broad range of domains, such as health and well-being, education, training, online communities, crowdsourcing, sustainability, customer loyalty, marketing, as well as in the enhancement of staff morale, motivation, and productivity (Raftopoulos et al., 2015; Seaborn and Fels, 2014). Furthermore, the literature also uses the term 'gameful design' as an alternative to that of 'gamification' (Deterding, 2015). According to Deterding et al. (2011), both terms frame the same phenomenon but differ in their intentional properties: gamification has the intention of using game elements in non-leisure contexts, while gameful design aims to create gameful experiences. However, the use of game elements (*gamification*) usually leads to gameful experiences. One of the best ways to create gameful experiences (*gameful design*) is by using game elements on the other hand. Hence, the results are similar in practice and thus in the present paper we use the two terms interchangeably.

Studies have shown that gamification can lead to positive behavioral changes (Hamari et al., 2014; Johnson et al., 2016; Seaborn and Fels, 2014); however, we currently do not fully understand the mechanisms behind these behavioral effects. For example, Hamari et al. (2014) identified confounding factors such as the context of the application or service being gamified and the 'qualities' of the users. Regarding the qualities, or characteristics, of the users, researchers have become increasingly interested in understanding individuals' varying motivations and the effects of gameful interventions on different people. This has inevitably brought into view the goal of personalized gameful systems adapted to individual users.

* Corresponding author.
E-mail addresses: gustavo@tondello.com (G.F. Tondello), amoraca@uoc.edu (A. Mora), rzej@gamified.uk (A. Marczewski), lennart.nacke@acm.org (L.E. Nacke).

https://doi.org/10.1016/j.ijhcs.2018.10.002
Received 2 October 2017; Received in revised form 20 July 2018; Accepted 12 October 2018
Available online 13 October 2018
1071-5819/

Nevertheless, we lack tools in multiple languages to assess users' motivations and preferred types of interactions with gameful systems, which limits our ability to design personalized systems. The Gamification User Types Hexad scale (Tondello et al., 2016) allows us to parse different user motivations in users' interactions with gameful applications using a self-report questionnaire. However, we still lack a substantial empirical validation of this scale in various languages. To address this shortcoming, we set out to validate the factor structure of the scale, in two major languages, English and Spanish, by conducting three large-scale studies. Additionally, we also investigated the distribution of the Hexad's user types in our sample.

The efficacy of individuated personalization has been demonstrated in the contexts of user interface design (Arazy et al., 2015; Nov and Arazy, 2013), persuasive technology (Kaptein et al., 2015, 2012), games (Bakkes et al., 2012; Connolly et al., 2012; Orji et al., 2014, 2013), and recently also in gamification (Orji et al., 2017). Consequently, we believe that gameful systems are more effective when personally adapted to each user. Gameful systems are effective when they help users achieve their goals, which often involve knowledge acquisition, supporting their changes in attitude or behavior, or engaging their interest in specific topics (Busch et al., 2015).

A promising approach to the personalization of gameful applications is to consider how user motivation is affected by their personality traits or the category (or categories) of 'user type(s)' they represent (Dixon, 2011; Ferro et al., 2013; Jia et al., 2016; Orji et al., 2017; Tondello et al., 2016). Empirical studies have shown that a user's personality traits (Goldberg, 1993) can accurately predict their level of enjoyment of several widely used game design elements, such as levels, points, leaderboards, avatars, quests, or challenges (Jia et al., 2016; Tondello et al., 2016). Player types themselves have also been shown to be partially correlated to personality traits (Nacke et al., 2014; Tondello et al., 2016). Nevertheless, a decade of games user research has demonstrated that personality traits provide only a partial explanation of differing motivations and playing style in games. This has fostered the notion that specifically tailored player typologies can predict playing styles more accurately than the more general taxonomies of personality traits such as the five factor model (Hamari and Tuunanen, 2014; Tondello et al., 2016; Yee, 2016; Yee et al., 2012). Therefore, creating a standardized and validated scale to assess user types according to their interactions with gameful systems is a promising approach with real potential utility in personalizing such systems.

To address this need, Marczewski (2015a) developed the Gamification User Types Hexad framework, based on research into human motivation, player types, and practical design experience. The User Types Hexad categorizes different styles of interaction with gameful applications according to six distinct types: 'Philanthropists', 'Socialisers', 'Free Spirits', 'Achievers', 'Players', and 'Disruptors', each of which we introduce in more detail in the following section. Marczewski (2015b) also suggested different game design elements that may support corresponding user types. Extending his work, Tondello et al. (2016) developed and validated a standard 24-item scale for scoring an individual according to each of the six user types. Their initial validation study with 133 university students demonstrated the viability of the scale through reliability and factor analyses. They also demonstrated the potential of the Hexad framework to personalize gameful systems by analyzing the correlation of each of the user types with 32 design elements commonly employed in gameful design and by showing positive correlations between the Hexad user types and the corresponding game design elements.

Nonetheless, despite the promising results, Tondello et al.'s work was conducted with a small sample and was limited to students of a single University and a single language. Therefore, we conducted three large-scale empirical validation studies. The first study used data from our online survey of 668 participants and interrogated personalized gameful design. The second study also collected data via an online survey, in which we distributed a tool that allowed individuals to find out their own Hexad user types scores; this survey collected data from 1681 participants. Using data from both studies, we investigated the reliability and internal consistency of the scale in both English and Spanish (see the Appendix for the detail of the scales used). The results demonstrate the scale's structural validity in both languages; however, there remains scope for improving a few subscales. Therefore, we conducted a third study with 152 participants with the goal of investigating potential improvements to the Achiever and Free Spirit subscales that could solve some of the issues identified in the first two studies. Additionally, we present an account of the user types distribution in the sample and demonstrate that gender and age are correlated to the participant's user types scores.

Our work contributes to the field of human-computer interaction (HCI) by providing empirical evidence of the structural validity of the Gamification User Types Hexad scale, enabling researchers to use it confidently in their work to improve our understanding of the mechanisms and effects of gameful interventions. Furthermore, we present evidence of the prevalence in our sample of some user types over others, as well as evidence of correlations of gender and age with the user types scores. This information is of great value in informing future gamification research and practice, allowing designers to create systems that are either more engaging for a broader segment of the population, or that target specific users, according to the specific business needs.

This paper first presents a review of the literature on player and user typologies (Section 2), which serves as a theoretical underpinning of this work. Next, we present the methods and results of the three large-scale studies aimed at validating the Hexad user types survey in English and Spanish (Sections 3–5). Finally, we summarize and discuss the results (Section 6) and conclude by analyzing the implications of these findings in HCI research on gamification and gameful design (Section 7).

2. Related work about player typologies and user typologies

Researchers in games and HCI have been studying different motivations and playing styles for over a decade and representing them as player typologies. One of the oldest and most frequently cited player type models is Bartle's (Bartle, 1996). Bartle studied what players desired from multi-user dungeons (MUDs) through a discussion between dozens of senior players. He identified four player types based on two axes that express the player's desire to interact with or act on the virtual world or on other players: 'Achievers' (acting on the world), 'Explorers' (interacting with the world), 'Socialisers' (interacting with other players), and 'Killers' (acting on other players). Bartle later extended it by adding a third dimension (Bartle, 2005): implicit or explicit (i.e., whether the player actions are automatic and unconscious or considered and planned). Thus, each of the four original types was divided into two sub-types. The 'implicit' sub-types are, respectively, 'Opportunists', 'Hackers', 'Friends', and 'Griefers'. The 'explicit' sub-types are (also respectively) 'Planners', 'Scientists', 'Networkers', and 'Politicians'.

Several other researchers followed this first attempt by Bartle to classify player preferences, attempting to create more accurate models based on empirical data. Based on a factor analysis of questions inspired by the original Bartle's player types, Yee (Yee, 2006; Yee et al., 2012) identified three main components of player motivation with ten subcomponents: 'achievement' (advancement, mechanics, competition), 'social' (socializing, relationship, teamwork), and 'immersion' (discovery, role-playing, customization, escapism). More recently, Yee (2015) expanded on this work by conducting a factor analysis with a large number of participants and developed a 'Gamer Motivation Profile' comprising 12 dimensions grouped in six clusters: 'Action' (Destruction and Excitement), 'Social' (Competition and Community), 'Mastery' (Challenge and Strategy), 'Achievement' (Competition and Power), 'Immersion' (Fantasy and Story), and 'Creativity' (Design and Discovery). Although this study aimed to define player motivations in

relation to a large variety of games, and is empirically supported by factor analysis, they have not presented a publicly available standardized assessment tool.

With similar goals, the first Demographic Game Design model (DGD1) (Bateman and Boon, 2006) tried to identify a broader range of player types by adapting the Myers-Briggs Type Indicator (MBTI) (Myers, 1962) to games. It proposed the player styles 'Conqueror', 'Manager', 'Wanderer', and 'Participant'. The second Demographic Game Design model (DGD2) (Bateman et al., 2011) explored what was termed the 'hard-core to casual' dimension, and interrogated different skill sets as well as players' preferences for single- and multiplayer gameplay. These two models served as the basis for the BrainHex player typology. BrainHex (Bateman et al., 2011; Nacke et al., 2014) is a top-down player typology, which takes inspiration from neurobiological player satisfaction research (Bateman and Nacke, 2010), previous typology approaches, discussions of patterns of play, and the literature on game emotions. It features seven archetypes denoting distinct player motivations. The seven BrainHex archetypes are: 'Achiever' (motivated by completion), 'Conqueror' (motivated by challenge), 'Daredevil' (motivated by excitement and risk), 'Mastermind' (motivated by strategic reasoning), 'Seeker' (motivated by exploration and curiosity), 'Socialiser' (motivated by social interactions), and 'Survivor' (motivated by frightening experiences). BrainHex supplements existing research with a more diverse array of player types and it has been initially investigated regarding its psychometric properties (Busch et al., 2016). It has been used in a number of recent studies in HCI (Birk et al., 2015; Orji et al., 2014; Zeigler-Hill and Monica, 2015). However, initial assays at empirical validation have shown that BrainHex does require further improvement, as demonstrated by its significantly low reliability scores (Busch et al., 2016).

Hamari and Tuunanen (2014) conducted a systematic review of these and other player type models to investigate their commonalities. The authors note that MMOs and online games are more frequently covered than other genres in several of these studies, and thus that this compromises the generalizability of these models. Furthermore, they compared all the analyzed models and suggested that they could be synthesized in five key dimensions pertaining to motivations of play: 'Achievement', 'Exploration', 'Sociability', 'Domination', and 'Immersion'.

While these models are often used in personalizing gameful systems, they were built specifically for game design. Therefore, their usefulness for gameful design in non-game applications or systems is limited. More recent papers have proposed new models specifically built to identify user preferences in gameful systems. In this context, Ferro et al. (2013) studied several models of personality and player types, aiming to find the similarities between them as well as to relate them to different game design elements. Their work grouped personality traits, player types, and game elements in five player categories: 'Dominant', 'Objectivist', 'Humanist', 'Inquisitive', and 'Creative'. However, their work was purely theoretical and lacks empirical validation.

Barata et al. (2017, 2014) studied data regarding student performance and gaming preferences from a gamified university level engineering course and identified four student types related to different gaming preferences: 'Achievers', 'Regular Students', 'Half-hearted Students', and 'Underachievers'. The authors suggest this framework may be used in future gamified education projects to tailor the course to the different characteristics of the students. Thus, their model is promising, but it is focused on a single (pedagogical) application domain. Differentially, the Gamification User Types Hexad (Marczewski, 2015a) covers a broad range of gameful systems. Therefore, we consider the Hexad model as having greater potential suitability in personalizing gameful systems, and thus warrants further research.

2.1. The Gamification User Types Hexad

Marczewski (2015a) proposed six user types that differ in the degree to which they can be motivated by either intrinsic (e.g., self-realization) or extrinsic (e.g., rewards) motivational factors when interacting with gameful systems. The user types are personifications of people's intrinsic and extrinsic motivations, as defined by self-determination theory (SDT). Within HCI research, the principles of SDT (Deci and Ryan, 1985; Ryan et al., 2006; Ryan and Deci, 2000a, 2000b) are often used to explain behavior motivation in interaction with digital technologies. SDT suggests that individual motivation to engage in a task can be located within different grades of internalization, ranging from wholly external to wholly internal motivation. In a simplified model, motivation can be *intrinsic* (i.e., afforded by the individual's perception of a task as enjoyable by itself), or *extrinsic* (i.e., afforded by factors outside of the task, such as expected outcomes that may result from completing the task).

SDT posits that a task is more likely to be intrinsically enjoyable when it supports three basic human psychological needs: *competence*, the feeling of having the skills needed to accomplish the task at hand; *autonomy*, the perception of being in control of a situation; and *relatedness*, the feeling of involvement with others. Additional work in the field notes the importance of these three pillars and indicates that they can make a strong and positive contribution to a person's mental health (Ryan et al., 2016). Accordingly, the Hexad model derives three of its intrinsically motivated types from SDT's psychological needs: 'Achievers' (*competence*), 'Free Spirits' (*autonomy*), and 'Socialisers' (*relatedness*).

Furthermore, there is evidence that *meaning* (purpose) facilitates internalization, increasing the motivation to carry out uninteresting but important activities (Deci et al., 1994; Grant, 2008), and leads to increased happiness and life satisfaction (Huta and Waterman, 2014; Peterson et al., 2005). This evidence informs the Hexad model's 'Philanthropist' user type.

Contrarily, the 'Player' user type is derived from SDT's notion of extrinsic motivation, i.e., it describes users who are mainly motivated to interact with a system in pursuit of external outcomes, such as rewards.

Below, we list the user types and further detail their identifying characteristics (Marczewski, 2015a; Tondello et al., 2016).

Philanthropists are motivated by *purpose*. They are altruistic and willing to give without expecting a reward.

Socialisers are motivated by *relatedness*. They want to interact with others and create social connections.

Achievers are motivated by *competence*. They seek to progress within a system by completing tasks, or prove themselves by tackling difficult challenges.

Free Spirits are motivated by *autonomy*, meaning freedom to express themselves and act without external control. They like to create and explore within a system.

Players are motivated by *extrinsic rewards*. They will do whatever to earn a reward within a system, independently of the type of the activity.

Disruptors are motivated by the triggering of *change*. They tend to disrupt the system either directly or through others to force negative or positive changes. They like to test the system's boundaries and try to push further. This type is not derived from SDT, but from empirical observation of this behavior within online systems (Marczewski, 2015c). Although disruption can sometimes be negative (e.g., that caused by 'Cheaters' or 'Griefers'), this is not always the case, because disruptors can also work to improve the system.

According to Tondello et al. (2016), the user types slightly overlap because some of their underlying motivations are related. Achievers and Players are both motivated by achievement but differ in their focus: Those in the Player category focus on extrinsic rewards while Achievers focus on competence. Philanthropists and Socialisers are both motivated to interact with other players. However, they differ because a Socialiser's interest resides solely in the interaction with other players, while Philanthropists are motivated in their interactions to help others (altruism). Finally, Free Spirits and Disruptors are both motivated by autonomy and creativity. However, Free Spirits stay within the system

limits without a desire to change them, while Disruptors seek to expand beyond these boundaries to change the system.

While the Hexad model was proposed a priori based on SDT, as explained above, Tondello et al. (2016) later developed a standard survey scale to score an individual's inclination towards each one of the Hexad user types. They followed a three-step procedure to create the scale: an expert workshop to generate a pool of suggested items for each user type, an expert rating (with a different group of experts) to verify the content validity of the suggested items and select the five best items for each user type, and an initial empirical validation study. The study collected data from 133 undergraduate and graduate students from the University of Waterloo, Canada, to test the scale's internal reliability and conduct a factor analysis; as well as follow-up data from 40 of the original participants to test the scale's test-retest stability. During the study, the authors decided to remove the least reliable item from each subscale, proposing a final 24-item (four items per subscale) standard survey for the Hexad User Types. Their results showed that all subscales had an internal reliability $\alpha \geq 0.698$ and a test-retest stability $r \geq 0.631$, except for the Player subscale, which had $r = 0.357$. Regarding the distribution of the user types in the sample, Philanthropists, Free Spirits, and Achievers showed the highest averages, closely followed by Socialisers and Players, whereas Disruptors showed the lowest average among all types.

Furthermore, Tondello et al. (2016) also demonstrated the usefulness of the Hexad user types to personalize gameful applications by investigating the correlations of each user type with 32 game design elements commonly used in gamification. They presented a table that significantly correlates several design elements with each user type and suggested that this information could be potentially useful in personalization. To this end, a designer could assess the target user cohort's user type profiles (or an individual user's profile) by employing the proposed survey scale; next, they could focus the gameful application's design efforts on those game design elements that are more likely to be enjoyable for the predominant user types in the profile, according to the correlation table.

It is important to note that player typologies have often been criticized for discussing types as discontinuous psychological factors, instead of presenting and measuring the traits in the form of a continuous scale (Hamari and Tuunanen, 2014). However, this is not the case with the Hexad model, which measures each user type score on a continuous scale and presents the results as a collection of six scores, corresponding to each type. In this way, and similar to other typologies, the Hexad user types should be understood as an archetypical categorization, where the types represent users for whom certain motivations are stronger than others (Hamari and Tuunanen, 2014). For example, a user who scores higher in the Free Spirit category and could be, thus, labeled singularly as being a 'Free Spirit', will be more motivated to pursue autonomous interactions with a gameful system, although the other motivations should still be present in a weaker degree.

The three studies presented here repeated the validation techniques employed by Tondello et al. (internal reliability analysis and factor analysis) using the same 24-item Hexad User Types Scale, but with larger and broader datasets, in addition to also carrying out a confirmatory factor analysis. Therefore, we provide empirical evidence that increases our confidence in the structural validity of the scale as a protocol to measure an individual's relatedness to each one of the six user types. In the following sections, we first present the results of each study and then discuss them all together.

3. First study

In the first study, we analyzed data collected during September and October 2016 from an online survey on personalized gamification.

3.1. Procedure

The survey was deployed as an online instrument using the LimeSurvey software. Participants were asked to complete a 15-minute survey composed of questions focused on their preferences while using digital gameful applications. The survey consisted of five sections with a total of 67 questions grouped as follows: demographics (age, gender, country, and native language); gaming habits; Hexad user types; examples of games participants knew; and participant's experiences with different game design elements. This survey was part of a larger study on personalized gamification, which will be reported elsewhere. For the present study, we focused only on the sections related to demographics and the Hexad user types. The Hexad user types section employed the 24 items suggested by Tondello et al. (2016) (see the Appendix).

The survey could be completed anonymously and allowed participants to skip any of the proposed questions or abandon the survey at any time. Prior to the decision to participate, participants were presented with an online informed consent form. In appreciation of the effort and time invested by respondents, they could participate in a draw, which only required the submission of a valid e-mail address after completion of the survey.

The survey could be completed by participants in English, Spanish, Catalan, or Portuguese. Two independent native speakers separately translated all the statements and descriptions into each language from the original version (which was in English for the Hexad user types survey; in Spanish for the remainder of the survey). Finally, each translated version was compared and assessed by an independent third native speaker during the design cycle before the survey activation in a continuous, discursive, improvement process.

3.2. Participants

We recruited participants by e-mail (in both academic and non-academic environments), as well as via social networks (Facebook, LinkedIn, Twitter, and Reddit), game events (Barcelona World Games), and Learning Management Systems from the participating institutions (Universitat Oberta de Catalunya, Universidad de La Laguna, and University of Waterloo). The study was approved by the ethics committees of the participating institutions. Participants were required to be at least 18 years old to participate and were not offered direct remuneration, but they were offered an opportunity to enter a draw to win one of two €50 prizes.

The total number of participants who answered the survey was 925. However, we discarded 257 participants who did not answer all the questions related to the Hexad user types survey, a necessary condition to allow accurate evaluation. Of the remaining 668 responses, the languages used to answer the survey were distributed as follows: Spanish (53.9%), English (29.3%), Catalan (11.4%), and Portuguese (5.4%). After looking at the language distribution, we concluded that we did not have a large enough sample to validate the Catalan and Portuguese translations of the scale. Therefore, we decided to discard these responses and validate only the English and Spanish versions.

Thus, the final dataset contained 556 responses: 360 in Spanish and 196 in English. The participants were 323 men, 224 women, and 9 did not inform. Participants' ages ranged from 18 to 65 years ($M = 30.37$, $SD = 10.07$) and were skewed towards younger participants (with 60% of participants under 30), possibly due to a dissemination focused on higher education institutions and the topic of the survey (gamification) being more appealing to a younger audience. The participants' native languages were distributed as follows: Spanish (62.4%), English (22.1%), and other (15.5%). Most participants answered the survey in their native language. The majority of participants whose native language was not available answered the survey in English. Participants were from 46 different countries, but with an irregular distribution, with a higher number of respondents from those countries where the survey was better advertised (see Table 1).

Table 1
Participant distribution per country of residence.

Country	Frequency	Percent
Argentina	19	3.4%
Canada	92	16.5%
China	8	1.4%
Colombia	14	2.5%
Germany	10	1.8%
Mexico	30	5.4%
Spain	281	50.5%
United Kingdom	9	1.6%
United States of America	21	3.8%
Venezuela	8	1.4%
Other (< 1% each)	62	11.3%
N/A	2	0.4%

3.3. Results

We analyzed the dataset by conducting the following procedures: internal reliability analysis, correlation between user types, and factor analysis. Because our aim was to assess the validity of the model, a confirmatory factor analysis (CFA) represented a more appropriate procedure than an exploratory factor analysis (EFA) (Levine, 2005). However, this raised the issue of comparability, since Tondello et al. (2016) have only reported an EFA. Therefore, to provide both a means of comparison with their prior work and a more reliable validity assessment, we conducted both an EFA and a CFA for this study. Additionally, we also provide a description of the user types' distribution and an analysis of the correlations of gender and age with the user types scores in this distribution.

3.3.1. Internal reliability and correlations

Table 2 presents the internal reliability analyses (Cronbach's α) for each subscale corresponding to each of the Hexad user types in the survey. We present both the overall scores (considering the whole sample) and the scores per survey language, to evaluate if the translations or cultural factors could have influenced the survey's reliability. Overall, the reliability scores are acceptable ($\alpha > 0.70$), except for those relating to the Free Spirit category in the English-language version of the survey, which are slightly below this level (0.629).

Table 3 presents the bivariate correlation coefficients and significance levels between each Hexad type and all the others. We employed Kendall's τ instead of the more common Pearson's r because the user type scores were non-parametric. As in previous work, we found some partial overlapping between the user types, but some of the observed significant correlations differ from those previously reported by Tondello et al. (2016).

3.3.2. Exploratory factor analysis

To enable a comparison with Tondello et al.'s results and to provide a richer set of evidences of the scale's structural validity, we first provide results from an exploratory factor analysis. The Kolmogorov–Smirnov test showed that the distributions of the Likert responses for all variables were significantly not normal, and several variables had skewness and/or kurtosis values above 1.0. Therefore, we followed the recommendation for conducting the factor analysis using polychoric correlations instead of the more traditional Pearson's correlations (Muthén and Kaplan, 1985). The correlation matrices were adequate for the analysis, with a KMO (Kaiser-Meyer-Olkin test) = 0.746 for the English sample and KMO = 0.844 for the Spanish sample; and Bartlett's test of sphericity was significant for both samples ($\chi^2_{(276)} = 1782.1$, $p < .001$ for the English sample; $\chi^2_{(276)} = 3771.9$, $p < .001$ for the Spanish sample). We used the software FACTOR 10.8.03 (Lorenzo-Seva and Ferrando, 2013) employing the Unweighted Least Squares method for factor extraction and a normalized direct oblimin rotation (because we expected the factors to partly overlap). Since our intention was to validate the existing Hexad model, we forced an analysis with six factors.

We present the results separately for the English and Spanish scales in Tables 4 and 5. In the EFA overall, the factor loads are higher for the combinations of item and factor that we were expecting, except for items P4, F2, and F3 in Spanish, which do not seem to be a good fit for the Philanthropist / Free Spirit factor as intended. Moreover, there is some partial overlapping between factors (represented by the items that score on more than one factor), which was expected since we found

Table 2
Internal reliability scores for each Hexad user type (overall and per language).

User Types	α (overall)	α (en)	α (sp)
Philanthropist	0.799	0.748	0.814
Socialiser	0.823	0.825	0.826
Free Spirit	0.699	0.629	0.727
Achiever	0.787	0.730	0.808
Player	0.864	0.843	0.874
Disruptor	0.759	0.788	0.746

Table 3
Bivariate correlation coefficients (Kendall's τ) and significance between each Hexad user type and all others.

User Types	Philanthropist	Socialiser	Free Spirit	Achiever	Player
Socialiser	0.386**				
Free Spirit	0.304**	0.209**			
Achiever	0.208**	0.129**	0.281**		
Player	−0.045	0.065*	0.030	0.103**	
Disruptor	0.021	0.020	0.189**	0.084**	0.097**

* $p < .05$.
** $p < .01$.

Table 4
Rotated factor loads for each of the Hexad survey items in English ($N = 196$).

User Types	Item	Rotated factor loads					
		1 (A)	2 (D)	3 (F)	4 (P)	5 (R)	6 (S)
Philanthropist (P)	**P1**				**0.667**		
	P2			0.312	**0.616**		
	P3				**0.605**		
	P4				**0.703**		
Socializer (S)	**S1**						**0.670**
	S2						**0.728**
	S3						**0.583**
	S4						**0.853**
Free Spirit (F)	**F1**			**0.554**			
	F2			**0.542**			
	F3			**0.507**			
	F4			**0.503**			
Achiever (A)	**A1**	**0.419**		0.337			
	A2	**0.787**					
	A3	**0.748**					
	A4	**0.586**					
Player (R)	**R1**					**0.690**	
	R2					**0.854**	
	R3					**0.753**	
	R4					**0.842**	
Disruptor (D)	**D1**		**0.723**				
	D2		**0.764**				
	D3		**0.791**				
	D4		**0.566**				
Eigenvalues		5.04	3.48	2.69	2.14	1.50	1.16
% of variance		20.98	14.51	11.21	9.91	6.26	4.82

Note. Exploratory factor analysis based on the polychoric correlations between items with an Unweighted Least Squares method and a normalized direct oblimin rotation. For improved readability, only the factor loads ≥ 0.30 are shown. The coefficients in bold type correspond to the item loads in the factor where they were expected to load higher.

Table 5
Rotated factor loads for each of the Hexad survey items in Spanish (N = 360).

User Types	Item	Rotated factor loads					
		1 (A)	2 (D)	3 (F)	4 (S)	5 (P)	6 (R)
Philanthropist (P)	**P1**					**0.453**	
	P2					**0.412**	
	P3					**0.620**	
	P4			0.321	0.309		
Socializer (S)	**S1**				**0.779**		
	S2				**0.773**		
	S3				**0.701**		
	S4				**0.755**		
Free Spirit (F)	**F1**			**0.662**			
	F2					0.400	
	F3					0.345	
	F4			**0.847**			
Achiever (A)	**A1**	**0.634**					
	A2	**0.815**					
	A3	**0.813**					
	A4	**0.669**					
Player (R)	**R1**						**0.869**
	R2						**0.959**
	R3						**0.711**
	R4						**0.842**
Disruptor (D)	**D1**		**0.729**				
	D2		**0.547**				
	D3		**0.809**				
	D4		**0.608**				
Eigenvalues		6.46	3.52	2.27	2.15	1.18	1.00
% of variance		26.91	14.63	9.47	8.97	4.95	4.16

Note. Exploratory factor analysis based on the polychoric correlations between items with an Unweighted Least Squares method and a normalized direct oblimin rotation. For improved readability, only the factor loads ≥ 0.30 are shown. The coefficients in bold type correspond to the item loads in the factor where they were expected to load higher.

significant correlations between the user types. However, this demonstrates that the survey items might not be capable of uniquely measuring each user type. This overlapping was more prominent in the Spanish survey.

The Goodness of Fit Index (GFI) = 0.984 for the English sample and GFI = 0.993 for the Spanish sample. Moreover, the root mean square of residuals (RMSR) = 0.0404 for the English sample, with an expected mean value of RMSR for an acceptable model ≤ 0.0716 as calculated by FACTOR; and the RMSR = 0.0295 for the Spanish sample, with an expected mean value of RMSR for an acceptable model calculated by FACTOR ≤ 0.0528. Therefore, both indices support the goodness of fit of the model to the data.

3.3.3. Confirmatory factor analysis

To evaluate further the goodness of the Hexad survey scale's fit to the theoretical model, we conducted a confirmatory factor analysis, using structural equation modeling in IBM SPSS Amos 24 (2016) with a maximum likelihood method. The six Hexad user types were modeled as latent variables, the 24 survey items were modeled as observed variables, and the four items associated with each user type were modeled as reflections of the respective latent variable (see Fig. 1). We only used the measurement model for the goals of our study. All parameters were left free to be estimated. Following Kline's suggestion, we report the results of the chi-squared test (χ^2) and the root mean square error of approximation (RMSEA) to evaluate the goodness of fit of the model (Kline, 2010). Table 6 further details the standardized (β) and unstandardized (B) regression weights, as well as the standard errors (SE) and critical ratios (CR) for each of the scale's items.

In the scale in English, the chi-squared test did not support the evidence for a good model fit ($\chi^2_{237} = 498.861, p < .001$). However, the test is known to inflate the statistical values for large sample sizes; therefore, the RMSEA should be a more reliable measure of fit for our study (Schmitt, 2011). The calculated RMSEA = 0.075 (90% CI = [.066, 0.084]), which is above the recommended cut for a well-fitted model (0.06 according to Schmitt (2011)). Therefore, the CFA results demonstrate that the measurement model is close to an acceptable fit, but that it has room for improvement. Particularly, the individual regression weights per item showed that items F2, F3, and F4 were a weaker fit to the Free Spirit subscale.

For the Spanish scale, the chi-squared test also failed to support the evidence for a good model fit ($\chi^2_{237} = 559.865, p < .001$). However, the calculated RMSEA = 0.062 (90% CI = [.055, 0.068]) is on the borderline of the recommended cut for a well-fitted model (0.06). Therefore, the CFA results demonstrate that the measure model is very close to an acceptable fit, but could still have some improvements. Particularly, items S3, D2, and D4 seem to be the weakest fits for their subscales per the individual regression weights.

3.3.4. Distribution

Table 7 reports the average scores and standard deviation for each Hexad user type in the sample. For the distribution analyses, we combined English and Spanish responses in a single dataset because our goal was to analyze the Hexad user types more broadly. Thus, the differential languages used in the survey were not relevant to this analysis because they served just to enable users with different native languages to participate. As in previous results, Philanthropists, Free Spirits, and Achievers showed the highest average scores, followed by Socialisers and Players, with Disruptors showing a significantly lower average.

In breaking down this distribution by gender, there seems to be a significant gendered difference between men's and women's scores on the user types Socialiser and Disruptor at $p < .05$, as well as those on Philanthropist and Achiever at $p < .10$, demonstrated by the t test. However, the mean differences are small: less than one point in average, from the 28 available for each subscale (see Table 8). Table 9 demonstrates that women tend to score slightly higher in philanthropism, socialization, autonomy, and achievement (although autonomy was not significant), whereas men tend to score slightly higher in disruption.

Looking at age, results demonstrate significant correlations between age and all user types except Disruptor (see Table 10). It seems intrinsic motivations (based on philanthropism, socialization, autonomy, and achievement) increase with age, whereas extrinsic motivations (based on rewards) decrease with age, although the effect sizes are small ($r \leq 0.2$).

4. Second study

In the second study, we analyzed data collected from July to December 2016 from an online survey advertised as a tool to let users test their own Hexad user type. Although it was conducted and reported separately, it occurred roughly during the same period of the first study.

4.1. Procedure

The survey was deployed on a public website (Gamified UK) using a specifically developed script. Participants were invited to take the Hexad user types survey (which also employed the 24 items suggested by Tondello et al. (2016)) to test their own Hexad user type. In addition, they could optionally inform their gender and age range. An e-mail address was asked for, to avoid duplicate answers; however, the addresses were recorded separately from the dataset to maintain anonymity.

After each completed survey, the website calculated the scores for each user type and presented the user with a chart of the results. Furthermore, all anonymous results were openly provided on the same website, reporting only the compounded scores for each user type, but not the participants' disaggregated answers.

This survey could be completed in English, Spanish, Portuguese, Italian, German, French, Turkish, or Russian. The English version was

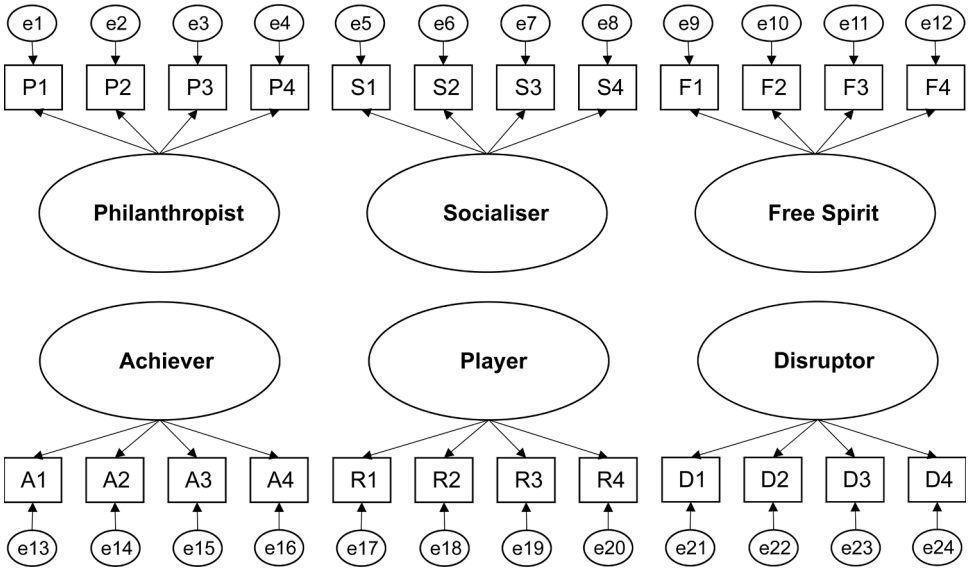

Fig. 1. Path model used for structural equation modeling.

originally suggested by Tondello and colleagues. The German version was also provided by the same researchers, as it was produced during the practitioners' workshop that created the survey items (which was conducted in German). The Spanish and Portuguese versions were the same as those used in the first study. The remaining versions were provided by voluntary translators and, therefore, were produced with less rigour.

4.2. Participants

We recruited participants through social media (mainly Twitter, Facebook, and gamification blogs). As previously stated, the survey was advertised as a user type test that allowed users to know their own Hexad user types profile. There were no restrictions on participation (except for the e-mail address check to avoid duplicate responses) and participants did not receive any compensation for participation.

The total number of participants who answered the survey was 1681. Table 11 shows the distribution of languages used by participants. English, Spanish, and Russian were the predominant languages used, whereas the remaining languages did not include enough

Table 7
Average scores and standard deviation for each Hexad user type (N = 556).

User Types	Mean score	S.D.
Philanthropist	23.52	3.82
Socialiser	21.26	4.46
Free Spirit	23.62	3.49
Achiever	23.53	3.64
Player	20.98	5.20
Disruptor	16.05	5.03

Table 6
Regression weights for each of the Hexad survey items.

User Types	Item	English (N = 196)				Spanish (N = 360)			
		β	B	SE	CR	β	B	SE	CR
Philanthropist (P)	**P1**	0.702	1.000			0.826	1.000		
	P2	0.675	1.156	0.153	7.546	0.655	0.933	0.073	12.717
	P3	0.628	0.977	0.136	7.158	0.714	0.993	0.070	14.096
	P4	0.612	1.051	0.150	7.010	0.723	0.990	0.069	14.315
Socialiser (S)	**S1**	0.804	1.000			0.752	1.000		
	S2	0.727	0.850	0.076	9.843	0.823	1.120	0.076	14.761
	S3	0.668	0.874	0.097	9.023	0.589	0.847	0.080	10.584
	S4	0.750	0.981	0.097	10.136	0.797	1.064	0.074	14.382
Free Spirit (F)	**F1**	0.630	1.000			0.657	1.000		
	F2	0.561	0.994	0.176	5.638	0.638	0.972	0.099	9.800
	F3	0.475	0.857	0.171	5.006	0.644	1.052	0.107	9.873
	F4	0.538	0.823	0.150	5.478	0.599	1.023	0.110	9.328
Achiever (A)	**A1**	0.607	1.000			0.749	1.000		
	A2	0.686	1.155	0.174	6.626	0.764	1.195	0.093	12.850
	A3	0.611	1.140	0.184	6.208	0.691	1.249	0.106	11.830
	A4	0.654	1.076	0.166	6.468	0.678	1.046	0.090	11.617
Player (R)	**R1**	0.693	1.000			0.837	1.000		
	R2	0.835	1.148	0.117	9.845	0.918	1.060	0.052	20.220
	R3	0.720	0.898	0.102	8.827	0.659	0.730	0.054	13.494
	R4	0.796	1.050	0.110	9.572	0.766	0.810	0.049	16.519
Disruptor (D)	**D1**	0.595	1.000			0.662	1.000		
	D2	0.719	0.996	0.136	7.339	0.563	0.641	0.074	8.715
	D3	0.829	1.337	0.173	7.725	0.816	1.167	0.112	10.425
	D4	0.665	1.040	0.149	6.988	0.585	0.804	0.089	8.989

Note. Confirmatory factor analysis conducted with a maximum likelihood structural equation modeling.

Table 8
Independent samples *t* test between user types and gender (*N* = 547).

User Types	*t*	*df*	*p*	mean diff.	95% CI of mean diff. lower	upper
Philanthropist	1.872	545	.062	0.622	−0.031	1.275
Socialiser	2.216	545	.027	0.850	0.097	1.604
Free Spirit	1.028	545	.304	0.314	−0.286	0.914
Achiever	1.800	545	.072	0.569	−0.052	1.192
Player	−0.533	545	.594	−0.241	−1.128	0.647
Disruptor	−2.093	545	.037	−0.912	−1.769	−0.056

Table 9
Average scores and standard deviation for each Hexad user type by gender (N = 547).

User Types	Male (*N* = 323) Mean score	S.D.	Female (*N* = 224) Mean score	S.D.
Philanthropist	23.27	3.87	23.89	3.75
Socialiser	20.98	4.49	21.83	4.30
Free Spirit	23.50	3.47	23.81	3.57
Achiever	23.31	3.68	23.88	3.58
Player	21.07	5.20	20.83	5.19
Disruptor	16.39	4.92	15.48	5.15

Table 10
Bivariate correlation analysis (Pearson's *r*) between user types and age (*N* = 556).

User Types	*r*	95% CI Lower	Upper
Philanthropist	0.204**	0.123	0.282
Socialiser	0.112**	0.029	0.193
Free Spirit	0.119**	0.036	0.200
Achiever	0.178**	0.096	0.257
Player	−0.110**	−0.191	−0.027
Disruptor	0.045	−0.038	0.128

** p < .01.

Table 11
Distribution of language used to answer the survey (*N* = 1681).

Language	Frequency	Percent
English	1073	63.8%
German	26	1.5%
Spanish	255	15.2%
French	19	1.1%
Italian	7	0.4%
Portuguese	5	0.3%
Russian	220	13.1%
Turkish	76	4.5%

participants to enable robust analysis. However, notwithstanding the size of the Russian-language user responses, we decided to analyze only the English and Spanish versions for this study because our first survey had no Russian-language version and we thus lacked an equivalent Russian-language dataset.

Therefore, the final dataset contained 1328 participants: 1073 in English and 255 in Spanish. There were 426 men, 375 women, 10 who reported as being of other genders, and 517 who did not specify a gender. Participants' ages were collected in ranges as detailed in Table 12. The survey did not ask about the participant's home country or native language.

4.3. Results

To enable further comparisons, we analyzed the dataset using the same procedures as in the first study: internal reliability analysis, correlation between user types, exploratory factor analysis, confirmatory factor analysis, description of the user types' distribution, and analysis of the correlations of gender and age with the user type scores in the distribution.

Table 12
Distribution of participants ages in the final dataset (*N* = 1328).

Age range	Frequency	Percent
17 or younger	31	2.3%
18–20	80	6.0%
21–29	230	17.3%
30–39	231	17.4%
40–49	139	10.5%
50–59	83	6.3%
60 or older	22	1.7%
N/A	512	38.6%

Table 13
Internal reliability scores for each Hexad user type (overall and per language).

User Types	α (overall)	α (en)	α (sp)
Philanthropist	0.774	0.774	0.774
Socialiser	0.827	0.828	0.820
Free Spirit	0.642	0.660	0.543
Achiever	0.610	0.616	0.594
Player	0.727	0.716	0.758
Disruptor	0.687	0.699	0.640

4.3.1. Internal reliability and correlations

Table 13 presents the internal reliability analyses (Cronbach's α) for each subscale corresponding to each of the Hexad user types in the survey. As in the first study, we have also split the sample per the survey's language to evaluate whether the translations or cultural factors could have influenced reliability. Overall, the reliability scores are acceptable ($\alpha > 0.70$) for the Philanthropist, Socialiser, and Player subscales, but slightly lower for the Free Spirit, Achiever, and Disruptor user types. The results were similar for both languages; however, the Free Spirit score was slightly lower in the Spanish language version of the survey.

Table 14 presents the bivariate correlation coefficients and significance levels between each Hexad types and all others. Although the scores vary, the position of significant correlations in this table are similar to those in the first study reported in this paper.

4.3.2. Exploratory factor analysis

Once more, we conducted an exploratory factor analysis employing the same method as before to enable comparisons between studies. The Kolmogorov–Smirnov test also showed that the distributions of the Likert responses for all variables were significantly not normal for this sample, and there was also significant skewness and kurtosis for some variables. Therefore, we employed the same method as before: we used the software FACTOR 10.8.03 (Lorenzo-Seva and Ferrando, 2013) with the Unweighted Least Squares method for factor extraction and a

Table 14
Bivariate correlation coefficients (Kendall's τ) and significance between each Hexad user type and all others.

User Type	Philanthropist	Socialiser	Free Spirit	Achiever	Player
Socialiser	0.382**				
Free Spirit	0.126**	0.034			
Achiever	0.207**	0.124**	0.213**		
Player	0.008	0.150**	0.076**	0.173**	
Disruptor	−0.011	−0.037	0.286**	0.044*	0.032

* p < .05.
** p < .01.

Table 15
Rotated factor loads for each of the Hexad survey items in English (N = 1073).

User Types	Item	Rotated factor loads 1 (D)	2 (R)	3 (F)	4 (P)	5 (S)	6 (A)
Philanthropist (P)	**P1**				**0.825**		
	P2				**0.571**		
	P3				**0.481**		
	P4				**0.687**		
Socializer (S)	**S1**					**0.632**	
	S2					**0.699**	
	S3					**0.543**	
	S4					**0.749**	
Free Spirit (F)	**F1**			**0.691**			
	F2			**0.488**			
	F3			**0.475**			
	F4			**0.587**			
Achiever (A)	**A1**						**0.571**
	A2						**0.367**
	A3						**0.372**
	A4						**0.794**
Player (R)	**R1**		**0.692**				
	R2		**0.565**				
	R3		**0.659**				
	R4		**0.482**				
Disruptor (D)	**D1**	**0.596**					
	D2	**0.814**					
	D3	**0.459**					
	D4	**0.656**					
Eigenvalues		4.93	3.34	2.29	1.75	1.11	1.05
% of variance		20.56	13.91	9.55	7.31	4.64	4.40

Note. Exploratory factor analysis based on the polychoric correlations between items with an Unweighted Least Squares method and a normalized direct oblimin rotation. For improved readability, only the factor loads ≥ 0.30 are shown. The coefficients in bold type correspond to the item loads in the factor where they were expected to load higher.

Table 16
Rotated factor loads for each of the Hexad survey items in Spanish (N = 255).

User Types	Item	Rotated factor loads 1 (P)	2 (F)	3 (R)	4 (D)	5 (A)	6 (S)
Philanthropist (P)	**P1**	**0.772**					
	P2	**0.472**					
	P3	**0.572**					
	P4	**0.692**					
Socializer (S)	**S1**						**0.739**
	S2						**0.791**
	S3						**0.696**
	S4						**0.736**
Free Spirit (F)	**F1**		**0.481**				
	F2				0.394		
	F3				0.460	0.359	
	F4		**0.751**				
Achiever (A)	**A1**					**0.716**	
	A2			0.307	−00.429		
	A3					**0.452**	
	A4					**0.649**	
Player (R)	**R1**			**0.771**			
	R2			**0.798**			
	R3			**0.576**			
	R4			**0.691**			
Disruptor (D)	**D1**				**0.633**		
	D2				**0.384**		
	D3				**0.508**		
	D4				**0.525**		
Eigenvalues		4.66	3.40	2.59	1.89	1.38	1.02
% of variance		19.41	14.19	10.78	7.87	5.73	4.24

Note. Exploratory factor analysis based on the polychoric correlations between items with an Unweighted Least Squares method and a normalized direct oblimin rotation. For improved readability, only the factor loads ≥ 0.30 are shown. The coefficients in bold type correspond to the item loads in the factor where they were expected to load higher.

normalized direct oblimin rotation, forcing an analysis with six factors. The correlation matrices were adequate for the analysis, with a KMO = 0.830 for the English sample and KMO = 0.768 for the Spanish sample; and Bartlett's test of sphericity was significant for both samples ($\chi^2_{(276)} = 7159.6$, $p < .001$ for the English sample; $\chi^2_{(276)} = 1830.3$, $p < .001$ for the Spanish sample).

We present the results separately for English and Spanish in Tables 15 and 16. As in the previous study, the overall factor loads from the EFA are higher for the combined items and factors where expected. However, in English, items A2 and A3 seem to be a weaker fit to their factors. In Spanish, item D2 seems to be a weaker fit to its factor; additionally, items F2, F3, and A2 scored less than 0.30 in their respective factors, appearing to be a better fit with other user types. Moreover, the partial overlapping between factors (represented by the items that score on more than one factor) once again appeared as expected but demonstrates that the survey items are not capable of completely differentiating each user type.

The GFI = 0.993 for the English sample and GFI = 0.986 for the Spanish sample. Moreover, the RMSR = 0.0258 for the English sample, with an expected mean value of RMSR for an acceptable model ≤ 0.0304 as calculated by FACTOR; and the RMSR = 0.0368 for the Spanish sample, with an expected mean value of RMSR for an acceptable model calculated by FACTOR ≤ 0.0626. Again, both indices support the goodness of fit of the model to the data.

4.3.3. Confirmatory factor analysis

As with the companion survey, we conducted a confirmatory factor analysis to evaluate further the goodness of fit of the Hexad survey scale to the theoretical model. We used the same procedure as before: a CFA using structural equation modeling in IBM SPSS Amos 24 (2016) with a maximum likelihood method and all parameters free to be estimated. The six Hexad user types were modeled as latent variables, the 24 survey items were modeled as observed variables, and the four items associated with each user type were modeled as reflections of the respective latent variable (see Fig. 1). As with our analysis of the first survey, we report the results of the chi-squared test (χ^2) and the root mean square error of approximation (RMSEA) to evaluate the goodness of fit of the model (Kline, 2010). Table 17 details the standardized (β) and unstandardized (B) regression weights, as well as the standard errors (SE) and critical ratios (CR) for each of the scale's items.

For the English scale, the chi-squared test did not support the evidence for a good model fit ($\chi^2_{237} = 1076.803$, $p < .001$). However, the calculated RMSEA = 0.057 (90% CI = [.054, 0.061]) is just below the recommended cut for a well-fitted model (0.06 according to Schmitt (2011)). Since the RMSEA should be a better indicator of fit due to the large sample, the CFA results suggest that the measure model is a good one to represent the theoretical factors corresponding to the Hexad user types. However, since the statistic is too close to the borderline, improvements would still be welcome. Particularly, items A2, A3, and R3 appear to be weaker fits for their subscale.

Regarding the Spanish scale, the chi-squared test also failed to support the evidence for a good model fit ($\chi^2_{237} = 526.967$, $p < .001$). However, the calculated RMSEA = 0.069 (90% CI = [.061, 0.077]) is just above the recommended cut for a well fit model (0.06). Therefore, we conclude that the model is close to a good fit; however, improvements could be made. In particular, items F2, F3, A2, A3, and D2 seem to be the weakest fits to their subscales.

4.3.4. Distribution

Table 18 reports the average scores and standard deviation for each Hexad user type in the sample. As in the first study reported in this paper, Philanthropists, Free Spirits, and Achievers showed the highest average scores, although this time Free Spirits' scores were slightly higher than the other two. Once more, Socialisers and Players followed with somewhat lower scores and Disruptors showed a significantly lower average.

Table 17
Standardized regression weights for each of the Hexad survey items.

User Types	Item	English (N = 1073)				Spanish (N = 255)			
		β	B	SE	CR	β	B	SE	CR
Philanthropist (P)	**P1**	0.740	1.000			0.719	1.000		
	P2	0.694	1.037	0.052	19.865	0.577	0.851	0.105	8.098
	P3	0.602	0.871	0.050	17.530	0.694	1.011	0.106	9.538
	P4	0.686	1.041	0.053	19.695	0.731	1.123	0.113	9.929
Socialiser (S)	**S1**	0.749	1.000			0.741	1.000		
	S2	0.765	0.937	0.041	23.086	0.740	0.898	0.083	10.754
	S3	0.735	0.931	0.042	22.303	0.751	1.097	0.101	10.898
	S4	0.703	0.957	0.045	21.373	0.694	0.914	0.090	10.139
Free Spirit (F)	**F1**	0.625	1.000			0.577	1.000		
	F2	0.524	0.781	0.062	12.654	0.405	0.564	0.122	4.628
	F3	0.558	0.811	0.061	13.206	0.396	0.662	0.145	4.553
	F4	0.570	0.851	0.064	13.374	0.592	1.276	0.219	5.828
Achiever (A)	**A1**	0.693	1.000			0.634	1.000		
	A2	0.430	0.786	0.070	11.148	0.443	1.018	0.194	5.247
	A3	0.423	0.785	0.071	10.998	0.458	1.135	0.211	5.376
	A4	0.633	0.931	0.064	14.626	0.626	1.227	0.189	6.478
Player (R)	**R1**	0.609	1.000			0.739	1.000		
	R2	0.793	1.213	0.076	16.028	0.745	0.891	0.095	9.431
	R3	0.470	0.658	0.054	12.228	0.543	0.681	0.091	7.448
	R4	0.635	0.969	0.064	15.213	0.644	0.754	0.087	8.633
Disruptor (D)	**D1**	0.526	1.000			0.668	1.000		
	D2	0.591	0.877	0.069	12.653	0.416	0.425	0.085	5.028
	D3	0.728	1.311	0.096	13.606	0.596	0.824	0.129	6.383
	D4	0.590	1.024	0.081	12.644	0.547	0.756	0.124	6.114

Note. Confirmatory factor analysis conducted with a maximum likelihood structural equation modeling.

Table 18
Average scores and standard deviation for each Hexad user type (N = 1328).

User Types	Mean score	S.D.
Philanthropist	22.90	3.81
Socialiser	20.77	4.66
Free Spirit	23.16	3.21
Achiever	22.45	3.53
Player	20.21	4.33
Disruptor	17.23	4.78

Table 19
Independent samples t test between user types and gender (N = 801).

User Types	t	df	p	mean diff.	95% CI of mean diff. Lower	Upper
Philanthropist	5.622	799	.000	1.417	0.922	1.911
Socialiser	3.353	799	.001	1.078	0.447	1.709
Free Spirit	1.762	799	.079	0.359	−0.041	0.759
Achiever	1.454	799	.146	0.361	−0.126	0.848
Player	−1.397	799	.163	−0.425	−1.022	0.172
Disruptor	−2.406	799	.016	−0.806	−1.465	−0.148

Table 20
Average scores and standard deviation for each Hexad user type by gender (N = 801).

User types	Male (N = 426) Mean score	S.D.	Female (N = 375) Mean score	S.D.
Philanthropist	22.55	3.89	23.97	3.14
Socialiser	20.40	4.49	21.47	4.60
Free Spirit	23.21	3.04	23.57	2.68
Achiever	22.26	3.42	22.62	3.60
Player	20.35	4.24	19.92	4.36
Disruptor	17.64	4.60	16.83	4.88

In breaking down this distribution by gender, the t test showed a significant difference between men's and women's scores on the user types Philanthropist, Socialiser, and Disruptor at $p < .05$, and on Free Spirit at $p < .10$. We only considered the two main genders (male/female) in the analysis, as the number of participants who reported a different gender was not big enough to afford useful conclusions. Overall, the mean differences are small: up to 1.42 points in average from the 28 available for each subscale, with the correlation of gender with philanthropism a bit stronger than with the other types (see Table 19). Table 20 again demonstrates that women tend to score a bit higher in philanthropism, socialization, autonomy, and achievement (although this time the differential score in the 'achievement' category was not significant), whereas men scored a bit higher in 'disruption'.

Regarding age, we were not able to perform a correlation analysis as we did in the first study because data were collected categorically (in ranges) instead of in scale (exact values). Therefore, we employed an analysis of variance (ANOVA). We also observed that the variance was not homogeneous across groups; therefore, in addition to the ANOVA tests, we also conducted a non-parametric test (independent-samples Kruskal-Wallis; KW) to verify the results from the ANOVA. Both the ANOVA and the KW tests suggest that age is significantly correlated with participants' scores on the user types Philanthropist, Socialiser, Player, and Disruptor (see Table 21). The effect sizes (η^2) suggest moderate correlations. Additionally, since neither the ANOVA nor the KW tests measure effect order, we employed the Jonckheere-Terpstra (JP) test to evaluate if the significant effects were ordered. The results suggested that all significant correlations were in fact ordered. Table 22 details the average scores and standard deviations for each user type by age and allows us to interpret the effects. As in the previous study, the results suggest that intrinsic motivations (philanthropism and

Table 21
One-way ANOVA, Kruskal-Wallis, and Jonckheere-Terpstra tests between user types and age (N = 816).

User Types	F	df	p (ANOVA)	η^2	p (KW)	p (JP)
Philanthropist	10.871	6, 809	.000	0.075	.000	.000
Socialiser	3.441	6, 809	.002	0.025	.002	.004
Free Spirit	1.761	6, 809	.104	0.013	.131	.186
Achiever	0.976	6, 809	.440	0.007	.734	.477
Player	7.898	6, 809	.000	0.055	.000	.000
Disruptor	3.622	6, 809	.001	0.026	.001	.000

Table 22
Average scores and standard deviation for each Hexad user type by age range ($N = 816$).

	Philant.		Socialiser		Free Spirit		Achiever		Player		Disruptor	
Age	*M*	SD	*M*	SD	*M*	SD	*M*	SD	*M*	SD	*M*	SD
17 or younger	21.74	4.0	21.84	5.7	22.48	3.1	21.65	5.0	22.03	4.7	15.55	4.2
18–20	21.51	3.9	19.71	4.9	23.94	2.6	22.13	3.4	21.95	3.7	17.60	4.0
21–29	22.70	3.9	20.62	4.9	23.07	3.0	22.57	3.3	20.81	4.0	16.36	4.7
30–39	23.23	3.3	20.65	4.4	23.36	2.8	22.49	3.4	19.79	4.0	17.58	4.7
40–49	24.34	3.1	21.68	4.0	23.63	3.1	22.47	3.5	19.44	4.0	17.27	4.9
50–59	24.86	2.5	22.29	3.7	23.59	2.7	23.00	3.0	19.05	4.7	18.42	5.1
60 or older	24.32	3.0	21.41	3.8	23.77	2.9	21.68	3.9	17.59	5.6	18.86	4.6

Table 23
Distribution of participants ages in the final dataset ($N = 152$).

Age range	Frequency	Percent
17 or younger	16	10.5%
18–20	9	5.9%
21–29	31	20.4%
30–39	27	17.8%
40–49	20	13.2%
50–59	12	7.9%
60 or older	4	2.6%
N/A	33	21.7%

socialization) increase with age, whereas extrinsic motivations (rewards) decrease with age. In addition, 'disruption' also seems to increase with age.

5. Third study

For the third study, we modified some statements of the Achiever and Free Spirit subscales with the goal of improving the weaknesses identified in the first two studies. We then invited the participants from the second study who had manifested interest in participating of future studies to answer the modified survey. Data were collected from June to July 2017.

5.1. Procedure

We employed the same survey used in the second study, with the only difference being the substitution of two Achiever items and one Free Spirit item in the survey:

- Achiever item "It is important to me to always carry out my tasks completely" was replaced by "It is important to me to continuously improve my skills".
- Achiever item "It is difficult for me to let go of a problem before I have found a solution" was replaced by "I enjoy emerging victorious out of difficult circumstances".
- Free Spirit item "I like to try new things" was replaced by "Opportunities for self expression are important to me".

The rationale for the changes in the Achiever items was to better capture the participant's preference for skill improvement and overcoming difficult situations, which are theorized characteristics for this archetype. On the other hand, we removed the two items related with the willingness to finish a task without interruption, which seemed to work weakly as a representation of this archetype according to the findings from our first two studies. For the Free Spirit subscale, we removed the item that seemed weaker in the first two studies, perhaps because it was too short and ambiguous, and included a new item related to self-expression, which is one of the theorized preferences of the Free Spirit archetype.

This time, the survey could be completed in English or Spanish. The appendix provides the complete listing of the scale items.

5.2. Participants

We invited participants from the second study who had authorized us to contact them for future studies by e-mail. Again, participants did not receive any compensation for participation. This time, 152 participants answered the survey in English and 12 in Spanish. Unfortunately, we did not receive a sufficient number of responses in Spanish to allow us to conduct statistical analyses. Therefore, we focus our analyses in the dataset with 152 responses in English (68 men, 56 women, 1 who reported as being of other gender, and 27 who did not specify a gender). Participants' ages were collected in ranges as detailed in Table 23.

5.3. Results

To enable further comparisons, we analyzed the dataset using the same procedures as in the first two studies: internal reliability analysis, correlation between user types, exploratory factor analysis, confirmatory factor analysis, and description of the user types' distribution. We did not perform an analysis of the correlations of gender and age with the user type scores this time because this was not the goal of this study.

5.3.1. Internal reliability and correlations

Table 24 presents the internal reliability analyses (Cronbach's α) for each subscale corresponding to each of the Hexad user types in the survey. Overall, the reliability scores are acceptable ($\alpha > 0.70$) for all user types, except Free Spirit ($\alpha = 0.60$). Table 24 also presents the

Table 24
Internal reliability of each subscale and bivariate correlation coefficients (Kendall's τ) with significance between each Hexad user type and all others.

User type	α	Philanthropist	Socialiser	Free Spirit	Achiever	Player
Philanthropist	0.704					
Socialiser	0.788	0.394*				
Free Spirit	0.596	0.252*	0.115			
Achiever	0.711	0.293*	0.244*	0.391*		
Player	0.748	−0.083	0.257*	−0.003	0.114	
Disruptor	0.700	0.010	−0.097	0.326*	0.144	0.044

* $p < .01$.

Table 25
Rotated factor loads for each of the Hexad survey items in English (N = 152).

User Types	Item	Rotated factor loads					
		1 (P)	2 (R)	3 (S)	4 (F)	5 (D)	6 (A)
Philanthropist (P)	**P1**	**0.575**					
	P2	**0.602**					
	P3	**0.795**					
	P4	**0.584**					
Socializer (S)	**S1**			**0.565**			
	S2			**0.718**			
	S3			**0.674**			
	S4			**0.704**			
Free Spirit (F)	**F1**				**0.735**		
	F2					0.361	
	F3				**0.419**		
	F4				**0.504**		
Achiever (A)	**A1**						**0.817**
	A2						**0.727**
	A3						**0.457**
	A4						**0.556**
Player (R)	**R1**		**0.585**				
	R2		**0.520**				
	R3		**0.757**				
	R4		**0.706**				
Disruptor (D)	**D1**					**0.635**	
	D2					**0.860**	
	D3					**0.455**	
	D4					**0.631**	
Eigenvalues		4.90	3.14	2.64	1.72	1.28	1.19
% of variance		20.44	13.10	11.00	7.16	5.32	4.97

Note. Exploratory factor analysis based on the polychoric correlations between items with an Unweighted Least Squares method and a normalized direct oblimin rotation. For improved readability, only the factor loads ≥ 0.30 are shown. The coefficients in bold type correspond to the item loads in the factor where they were expected to load higher.

bivariate correlation coefficients and significance levels between each Hexad types and all others. Once more, the position of significant correlations in this table are similar to those in the first two studies reported in this paper; however, it is noteworthy that this time the Achiever scores were not significantly correlated with Player and Disruptor. Therefore, it seems that the modifications introduced in the scale could better differentiate participants between these user types.

5.3.2. Exploratory factor analysis

Similar to the first two studies, the Likert scale responses were non-parametric, so we employed the software FACTOR 10.8.03 (Lorenzo-Seva and Ferrando, 2013) with the polychoric correlations as input for the factor analysis, the Unweighted Least Squares method for factor extraction and a normalized direct oblimin rotation, forcing an analysis with six factors. The correlation matrix was adequate for the analysis, with a KMO = 0.714, and Bartlett's test of sphericity was significant ($\chi^2_{(276)} = 1113.1$, $p < .001$). We present the results in Table 25. The overall factor loads from the EFA are higher for the combined items and factors where expected; however, there is a relevant overlapping between some of the Philanthropist items with the Socialiser items. Additionally, the modifications introduced for the Achiever subscale seem to have improved it in comparison with the second study, but the modification to the Free Spirit subscale does not evidence a sufficient improvement: although the new item (F4) weighted well in the Free Spirit factor, item F2 did not contribute well to the factor.

Nonetheless, the Goodness of Fit Index (GFI) = 0.981 and the Root Mean Square of Residuals (RMSR) = 0.0430, with an expected mean value of RMSR for an acceptable model ≤ 0.0814 as calculated by FACTOR. Thus, both indices support the goodness of fit of the model to the data.

5.3.3. Confirmatory factor analysis

As before, we conducted a confirmatory factor analysis to evaluate further the goodness of fit of the Hexad survey scale to the theoretical model. We used the same procedure as before: a CFA using structural equation modeling in IBM SPSS Amos 24 (2016) with a maximum likelihood method and all parameters free to be estimated. The six Hexad user types were modeled as latent variables, the 24 survey items were modeled as observed variables, and the four items associated with each user type were modeled as reflections of the respective latent variable (see Fig. 1). Table 26 details the standardized (β) and unstandardized (B) regression weights, as well as the standard errors (SE) and critical ratios (CR) for each of the scale's items.

Table 26
Regression weights for each of the Hexad survey items in English (N = 152).

User Types	Item	β	B	SE	CR
Philanthropist (P)	**P1**	0.602	1.000		
	P2	0.634	1.212	0.224	5.413
	P3	0.575	1.083	0.212	5.104
	P4	0.628	1.239	0.230	5.385
Socialiser (S)	**S1**	0.634	1.000		
	S2	0.762	1.175	0.168	6.995
	S3	0.693	1.223	0.185	6.604
	S4	0.697	1.243	0.187	6.636
Free Spirit (F)	**F1**	0.651	1.000		
	F2	0.453	0.601	0.141	4.248
	F3	0.371	0.579	0.161	3.595
	F4	0.630	1.115	0.211	5.288
Achiever (A)	**A1**	0.681	1.000		
	A2	0.625	1.049	0.179	5.852
	A3	0.496	0.629	0.128	4.903
	A4	0.679	1.140	0.186	6.131
Player (R)	**R1**	0.724	1.000		
	R2	0.875	1.214	0.154	7.860
	R3	0.388	0.444	0.102	4.345
	R4	0.602	0.804	0.120	6.679
Disruptor (D)	**D1**	0.659	1.000		
	D2	0.547	0.946	0.223	4.250
	D3	0.787	1.805	0.380	4.755
	D4	0.655	1.479	0.321	4.603

Note. Confirmatory factor analysis conducted with a maximum likelihood structural equation modeling.

Like the first two studies, the chi-squared test did not support the evidence for a good model fit ($\chi^2_{237} = 372.480$, $p < .001$). However, the calculated RMSEA = 0.062 (90% CI = [.049, 0.073]) is on the borderline of the recommended cut for a well-fitted model (0.06 according to Schmitt (2011)). The estimated model fit is very similar to that achieved in the second study presented in this paper. This time, the items with the lower weight for their subscales were F2, F3, A3, and R3.

5.3.4. Distribution

Table 27 reports the average scores and standard deviation for each Hexad user type in the sample. Once more, Philanthropists, Free Spirits, and Achievers showed the highest average scores, with Achievers' scores being slightly higher than the other two this time. Also, like the first two studies, Socialisers and Players followed with somewhat lower scores and Disruptors showed a significantly lower average.

Table 27
Average scores and standard deviation for each Hexad user type (N = 152).

User Types	Mean score	S.D.
Philanthropist	23.68	2.93
Socialiser	20.98	4.37
Free Spirit	23.45	2.95
Achiever	24.26	3.00
Player	20.66	4.44
Disruptor	16.72	4.68

6. Discussion

This research analyzed data from three substantial survey studies aimed at evaluating the Gamification User Types Hexad scale proposed by Tondello et al. (2016) in two languages: English and Spanish (see the Appendix for the complete scales used). To that end, we carried out a reliability analysis as well as exploratory and confirmatory factor analyses on the three data sets. In addition, we examined the distribution of each user type in the sample and how participant's demographics (gender and age) relate to their scores. These are the main findings:

- Empirical evidence supports the structural validity of the scale in both English and Spanish. However, some improvements are desirable to improve the reliability of a few specific survey items, particularly those related to the Free Spirit and Achiever user types.
- Philanthropist and Socialiser user types seem to be moderately correlated.
- Philanthropist, Free Spirit, and Achiever are the prevalent user types. On the other hand, Disruptor is the least common user type.
- Results suggest that a person's user type is correlated with their gender and age. Women seem to score slightly higher than males on average in all the intrinsic motivations, whereas men seem to score slightly higher in disruption. Additionally, intrinsic motivations seem to slightly increase with age, whereas extrinsic motivations seem to decrease with age.

The following subsections discuss each one of these findings in more detail.

6.1. Scale validity

The reliability analysis from the first study showed that most subscales are internally consistent (see Table 2). A notable exception is the Free Spirit subscale, which showed slightly lower reliability scores in both languages. The analysis from the second study also showed that most subscales are internally consistent, with the Free Spirit and Achiever subscales showing slightly lower consistencies than desired (see Table 13). Overall, when compared to the prior work by Tondello and colleagues, the results are similar. Therefore, these results evidence that the internal consistency of the subscales is adequate, but that there is scope for improvements. However, we were careful to not rely only on Cronbach's alpha as the indicator of scale dimensionality and consistency because it is well known that high alpha values can instead be indicators of lengthy scales, parallel items, or narrow coverage of the constructs under consideration (Panayides, 2013). Thus, we also carried out exploratory and confirmatory factor analyses to further verify the scale's internal consistency with more robust methods.

It is noteworthy that the lower reliability scores of Free Spirit and Achiever items in Spanish did not occur in the first study. Unfortunately, we do not have data about the participant's native country in the second study, but we do know that most participants who answered the survey in the first study were from Spain. Therefore, the difference that appeared in the second study might be because participants were not as familiar with the language as the participants in the first study, or were from a different Spanish-speaking country, which might have different linguistic or cultural norms than the Spanish participants.

The exploratory factor analysis of our first survey showed that most items loaded higher in the factors they were expected to, except for a few Free Spirit items. These items do not seem to be a good representation of their factor, and were likely the reason the internal consistency of the Free Spirit subscale was a bit lower than the others. In our study of the second survey, the EFA similarly showed a good correspondence of higher items' loads with the expected factors, except for some Free Spirit and Achiever items. Additionally, there were a few issues with Philanthropist items in Spanish. Thus, besides confirming potential issues with the same Free Spirit items, the EFA of our second survey explains the lower consistency to be found in the Achiever subscale. As it happens, the prior work of Tondello et al. has also pointed to lower loads for items F2, F3, A2, and A3 in their respective factors. Therefore, we conclude that although the subscales are consistent overall, these four items should be improved to enhance the scale's reliability.

The confirmatory factor analysis from our study of the first survey suggested that the measurement model is close to a good fit with the theory, but improvements are desirable. In English, the CFA also points to potential improvements in the Free Spirit items. However, the standardized regression weights were more balanced in Spanish; thus, they did not help us identify which items needed improvement. In our analysis of the second survey, the CFA showed a slightly better fit between the measurement model and the theory in English—just within the acceptable threshold considering the calculated RMSEA. However, the same did not occur in Spanish, where the RMSEA remained close but slightly above the borderline. An analysis of the standardized regression weights suggests a need to improve Achiever items A2, A3, and R3 in English and F2, F3, A2, A3, and D2 in Spanish.

To investigate potential improvements in the scale regarding the Free Spirit and Achiever subscales, we then conducted the third study replacing one item of the Free Spirit subscale and one item of the Achiever. The data were only analyzed in English because the new dataset did not contain enough responses in Spanish ($N = 12$). The results of the EFA and CFA showed that the overall reliability and model fit remained similar in comparison with the first two studies. However, an inspection of the item weights in both factor analyses showed that the two newly introduced items F4 and A4 loaded well in their respective subscales. This might suggest that these replacements represent a step in the right direction and that these two subscales might be improved even further in the future with additional adjustments.

Looking at the correlations between user types, there are several significant ones. In the results of the first survey, the most relevant correlations (with $\tau > 0.20$) occurred between Philanthropist and Socialiser, Philanthropist and Free Spirit, and Free Spirit and Achiever (see Table 3). In the second study, they occurred between Philanthropist and Socialiser, and Free Spirit and Disruptor (see Table 14). In the third study, they occurred between Philanthropist and Socialiser, Free Spirit and Achiever, and Free Spirit and Disruptor (see Table 24). When comparing these results to prior work, we noted that Tondello and colleagues found several more significant correlations, between almost all combinations except for those with the Disruptor type. Since the two survey studies presented in this work analyzed much larger and diverse datasets, the results may be considered more dependable.

The correlation between the Philanthropist and Socialiser user types needs special attention because it showed consistently higher coefficients ($\tau \approx 0.40$) in all studies, suggesting a moderate correlation. The theoretical background suggests a partial overlap between these user types, since both are related to social interactions; however, there should be a difference in that Philanthropists should be more motivated by interactions in which they can help others, whereas Socialisers should be more motivated by the social interactions *per se*, even those that do not involve helping others. The results from all the studies suggest that this overlap might be even stronger than anticipated, meaning that a correlation between these two types does indeed seem to exist, i.e., one cannot be highly motivated by socialization without being at least moderately motivated by the will to help others, and vice versa.

The correlations between Achiever and Free Spirit scores that consistently appeared in all analyses also deserve attention because they were not predicted by the theory. Moreover, considering the lower consistency scores of some of the items in these subscales, consistently demonstrated by the factor analyses, we conclude that future improvement of these items should help us discriminate between these

two user types.

6.2. User types distribution

Across all the three studies, the Philanthropist, Free Spirit, and Achiever user types consistently scored higher on average than the other types. This suggests that these are generally the three strongest motivations for user interaction with gameful systems. This is consistent with self-determination theory, which posits that perceived autonomy and competence are innate psychological needs that individuals seek to satisfy to increase their happiness. Similarly, SDT suggests that the pursuit of meaning leads to easier internalization of necessary (but not intrinsically enjoyable) tasks and increased happiness. The Socialiser and Player user types consistently scored just a bit lower than the three strongest types across all studies, i.e., about 2–3 points (out of 28) lower in average. This also makes sense according to SDT, since 'relatedness' is the third psychological need that facilitates intrinsic motivation, and rewards are one of the common means of facilitating extrinsic motivation. On the other hand, the Disruptor user type consistently scored lower than all the other types, about 5–7 points lower than the highest scoring types. This clearly demonstrates that the motivation for change is less prevalent in the cohort than other motivation factors, although it is still relevant.

Regarding the correlations of demographic variables in the user types' scores, results from both studies suggest that both age and gender are correlated to an individual's user types profile. Women seem to score slightly (just under one point) higher than males on average in all the intrinsic motivations, whereas men seem to score slightly (also just under one point) higher in disruption. Additionally, all the user types showed some correlation with age, suggesting that the intrinsic motivations slightly increase with age (about 1–3 points from a person's 20s to their 60s), whereas extrinsic motivations decrease with age (also about 1–3 points). Disruption also seemed to increase with age, but the effect was only statistically significant in the results of the second survey. These results suggest that the motivations to interact with gameful systems are not stable through an individual's lifetime and vary over time, perhaps in a consistent way; however, the expected difference is small, so we should expect a small variation from one's basic motivations rather than a wholesale deviation. Therefore, as a guideline, designers can expect that older users will be slightly more intrinsically motivated than younger ones, particularly regarding the motivation of purpose, which showed a proportionally stronger correlation with age than the other user types, and slightly less motivated by extrinsic rewards.

6.3. Limitations

The goal of the three studies presented in this article was to validate the factor structure of the Gamification User Types Hexad scale with large samples. Although we collected large datasets in the surveys, the geographical distribution of participants in the first study was concentrated in the countries where the survey was more intensively disseminated, with a special concentration in Spain. The second survey was available in more languages and was more broadly disseminated on the internet, thus, we believe it might have attracted a more diverse sample. However, we did not collect information on the participants' country of origin or native language, so we cannot be certain. On the other hand, the third study collected data from a smaller number of participants. Therefore, future studies should aim to repeat the scale validation with an even more diverse participant sample, trying to collect data from participants from all over the world. Moreover, although we collected data in several languages, only English- and Spanish-language responses provided large enough cohorts to enable meaningful analysis, in both cases. Therefore, we concentrated our efforts on validating these two versions of the scale, leaving the additional translations to be validated in future work.

Furthermore, although the results showed that the scale is generally reliable, they also identified specific points for improvement, which we have highlighted, and which should be addressed in future work.

Finally, Tondello et al. also carried out additional analyses that we did not reproduce: test-retest stability, correlation of the Hexad user types with personality traits, and correlation of the Hexad user types with different game design elements. This is because our goal was focused on validating the factor structure of the scale. Moreover, the process followed to create the scale items described by Tondello et al. was meant to guarantee the construct validity of the scale because the items were generated by an expert panel and validated by a different expert panel. However, they did not report any measure of construct validity, and we did not further investigate it in this work. Consequently, future work should also repeat these analyses with larger samples to verify Tondello et al.'s findings, as well as employ adequate methods to assess the face, content, criterion, and construct validity of the scale.

7. Conclusion

In the present work, we conducted three large scale survey studies to validate the structure of the Gamification User Types Hexad scale in English and Spanish, and to investigate the distribution of each user type in the cohorts. We demonstrated that the scale structural validity is generally acceptable through reliability analysis and factor analysis. This means that the Hexad user types survey is suitable for use in future work investigating the effects of gamification or developing guidelines and methods for personalized gameful design. Based on the results presented in this paper, we recommend that future work use the modified scale we employed in our third study (see the Appendix for the complete scale). The scale can be used to assess participants' user types in future HCI research involving gamification or gameful design. This could be useful, for example, to verify if the effects of gameful interventions or methods are moderated by the user types. It can also be used by practitioners to design applications that are personalized to the preferences of individual users.

Nevertheless, the results also suggested that some improvements could be made to improve the Hexad scale's validity. Particularly, looking at the modified survey used in our third study, the following survey items should still be investigated and potentially improved to enhance the reliability of the Free Spirit and Achiever subscales and better discriminate (reduce the correlation) between them: F2 ('I often let my curiosity guide me.'), F3 ('Being independent is important to me.'), and A3 ('It is important to me to continuously improve my skills.'). Additionally, there were some additional items that only had issues in one of the studies and for one of the languages, thus suggesting that further studies should be conducted to verify our findings. Moreover, future work can also investigate the face, content, criterion, and construct validity of the scale.

Regarding the distribution of user types in our cohorts, the results suggest that Philanthropist, Free Spirit, and Achiever are on average the strongest motivations, closely followed by Socialiser and Player; conversely, the Disruptor user type consistently has lower average scores. The participants' user type scores were also significantly correlated to their genders and ages. Women scored slightly higher than men in all intrinsic motivations, whereas men scored slightly higher in disruption on average. Additionally, the influence of intrinsic motivators seems to increase as a person ages, whereas that of extrinsic motivations (rewards) seems to decrease with age.

Furthermore, the evidence suggests there is a stronger correlation between the Philanthropist and Socialiser types than the theory anticipated, suggesting the possibility of an improvement to the theory itself, i.e., it should acknowledge that a person who is highly motivated by philanthropism will probably also be motivated by socialization in some degree, and vice versa.

Our work provides a valuable contribution to HCI research in

gamification and gameful design by presenting highly robust empirical evidence on the structural validity of the Gamification User Types Hexad scale. This will allow researchers to use the scale in future studies to better understand the mechanisms and effects of gameful interventions, ultimately leading to a better comprehension of the psychological processes behind them and enabling the creation of better methods and guidelines to design effective and personalized gameful systems.

Funding

This work was supported by the CNPq, Brazil; NSERC [grant RGPIN-418622-2012, University of Waterloo]; SSHRC [grant 895-2011-1014, IMMERSe]; Mitacs [grant IT07255]; CFI [grant 35819]; Agència de Gestió d'Ajuts Universitaris i de Recerca, Generalitat de Catalunya [Industrial Doctorate programme 2014-DI-006]; and the Spanish Government [project TIN2013-45303-P].

Declaration of interest

Mr Marczewski created the original Hexad and hosted the survey for studies 2 and 3. The Hexad forms part of his design process when he consults on gamification. However, he does not directly sell the Hexad or use of it exclusively. He provided the raw data from the online version of the survey but was in no way involved in data analysis or conclusions of this paper. There are no other known conflicts of interest associated with this publication.

Appendix

Table 28
The Gamification User Types Hexad scale used in the first two studies.

User Types	#	English items	Spanish items
Philanthropist	**P1**	It makes me happy if I am able to help others.	Me hace feliz ser capaz de ayudar a los demás.
	P2	I like helping others to orient themselves in new situations.	Me gusta guiar a los demás en las situaciones nuevas.
	P3	I like sharing my knowledge.	Me gusta compartir mi conocimiento con los demás.
	P4	The wellbeing of others is important to me.	El bienestar de los demás es importante para mí.
Socialiser	**S1**	Interacting with others is important to me.	Interactuar con los demás es importante para mí.
	S2	I like being part of a team.	Me gusta formar parte de un equipo.
	S3	It is important to me to feel like I am part of a community.	Sentir que formo parte de una comunidad es importante para mí.
	S4	I enjoy group activities.	Disfruto con las actividades grupales.
Free Spirit	**F1**	It is important to me to follow my own path.	Seguir mi propio camino es importante para mí.
	F2	I often let my curiosity guide me.	A menudo me dejo guiar por la curiosidad.
	F3	I like to try new things.	Me gusta probar cosas nuevas.
	F4	Being independent is important to me.	Ser independiente es importante para mí.
Achiever	**A1**	I like defeating obstacles.	Me gusta superar las dificultades.
	A2	It is important to me to always carry out my tasks completely.	Realizar siempre por completo mis tareas es importante para mí.
	A3	It is difficult for me to let go of a problem before I have found a solution.	Me resulta difícil abandonar un problema antes de encontrarle una solución.
	A4	I like mastering difficult tasks.	Me gusta dominar tareas difíciles.
Player	**R1**	I like competitions where a prize can be won.	Me gustan las competiciones en las que se pueda ganar un premio.
	R2	Rewards are a great way to motivate me.	Los premios son una buena manera de motivarme.
	R3	Return of investment is important to me.	Recuperar lo invertido es importante para mí.
	R4	If the reward is sufficient I will put in the effort.	Si el premio es adecuado, voy a hacer un esfuerzo.
Disruptor	**D1**	I like to provoke.	Me gusta provocar.
	D2	I like to question the status quo.	Me gusta cuestionar el estado de las cosas.
	D3	I see myself as a rebel.	Me describo a mí mismo como un rebelde.
	D4	I dislike following rules.	No me gusta seguir las reglas.

Note. Each subscale is scored by adding together the value of the participant's responses to the four items that comprise the subscale.

Table 29
The Gamification User Types Hexad scale used in the third study.

User Types	#	English items	Spanish items
Philanthropist	**P1**	It makes me happy if I am able to help others.	Me hace feliz ser capaz de ayudar a los demás.
	P2	I like helping others to orient themselves in new situations.	Me gusta guiar a los demás en las situaciones nuevas.
	P3	I like sharing my knowledge.	Me gusta compartir mi conocimiento con los demás.
	P4	The wellbeing of others is important to me.	El bienestar de los demás es importante para mí.
Socialiser	**S1**	Interacting with others is important to me.	Interactuar con los demás es importante para mí.
	S2	I like being part of a team.	Me gusta formar parte de un equipo.
	S3	It is important to me to feel like I am part of a community.	Sentir que formo parte de una comunidad es importante para mí.
	S4	I enjoy group activities.	Disfruto con las actividades grupales.
Free Spirit	**F1**	It is important to me to follow my own path.	Seguir mi propio camino es importante para mí.
	F2	I often let my curiosity guide me.	A menudo me dejo guiar por la curiosidad.
	F3	Being independent is important to me.	Ser independiente es importante para mí.
	F4	Opportunities for self expression are important to me.	Tener la oportunidad de expresarme es importante para mí.
Achiever	**A1**	I like overcoming obstacles.	Me gusta superar las dificultades.
	A2	I like mastering difficult tasks.	Me gusta dominar tareas difíciles.
	A3	It is important to me to continuously improve my skills.	Mejorar continuamente mis habilidades es importante para mí.
	A4	I enjoy emerging victorious out of difficult circumstances.	Me gusta salir victorioso de las circunstancias difíciles.

(*continued on next page*)

Table 29 (*continued*)

User Types	#	English items	Spanish items
Player	**R1**	I like competitions where a prize can be won.	Me gustan las competiciones en las que se pueda ganar un premio.
	R2	Rewards are a great way to motivate me.	Los premios son una buena manera de motivarme.
	R3	Return of investment is important to me.	Recuperar lo invertido es importante para mí.
	R4	If the reward is sufficient I will put in the effort.	Si el premio es adecuado, voy a hacer un esfuerzo.
Disruptor	**D1**	I like to provoke.	Me gusta provocar.
	D2	I like to question the status quo.	Me gusta cuestionar el estado de las cosas.
	D3	I see myself as a rebel.	Me describo a mí mismo como un rebelde.
	D4	I dislike following rules.	No me gusta seguir las reglas.

Note. Each subscale is scored by adding together the value of the participant's responses to the four items that comprise the subscale.

References

Arazy, O., Nov, O., Kumar, N., 2015. Personalityzation: UI personalization, theoretical grounding in HCI and design research. AIS Trans. Hum. Comput. Interact. 7, 43–69.

Bakkes, S., Tan, C.T., Pisan, Y., 2012. Personalised gaming. In: Proceedings of The 8th Australasian Conference on Interactive Entertainment Playing the System - IE '12. ACM, pp. 1–10. https://doi.org/10.1145/2336727.2336731.

Barata, G., Gama, S., Jorge, J.A.P., Gonçalves, D.J.V., 2014. Relating gaming habits with student performance in a gamified learning experience,. In: Proceedings of the First ACM SIGCHI Annual Symposium on Computer-Human Interaction in Play - CHI PLAY '14. New York, NY, USA. ACM, pp. 17–25. https://doi.org/10.1145/2658537.2658692.

Barata, G., Gama, S., Jorge, J., Gonçalves, D., 2017. Studying student differentiation in gamified education: a long-term study. Comput. Hum. Behav. 71, 550–585. https://doi.org/10.1016/j.chb.2016.08.049.

Bartle, R., 2005. Virtual worlds: why people play. Massively Mult. Game Dev. 2.

Bartle, R., 1996. Hearts, Clubs, diamonds, spades: players who suit MUDs. J. MUD Res. 1.

Bateman, C., Lowenhaupt, R., Nacke, L.E., 2011. Player typology in theory and practice. In: Proceedings of DiGRA 2011.

Bateman, C., Nacke, L.E., 2010. The neurobiology of play. In: Proceedings of Futureplay 2010. Vancouver, BC, Canada. ACM, pp. 1–8. https://doi.org/10.1145/1920778.1920780.

Bateman, C.M., Boon, R., 2006. 21st century game design (Game Development Series). Charles river media.

Birk, M.V., Toker, D., Mandryk, R.L., Conati, C., 2015. Modeling motivation in a social network game using player-centric traits and personality traits. In: Proceedings of User Modeling, Adaptation and Personalization. Springer, pp. 18–30. https://doi.org/10.1007/978-3-319-20267-9_2.

Busch, M., Mattheiss, E., Orji, R., Fröhlich, P., Lankes, M., Tscheligi, M., 2016. Player type models – Towards empirical validation. In: Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems. ACM.. https://doi.org/10.1145/2851581.2892399.

Busch, M., Mattheiss, E., Orji, R., Marczewski, A., Hochleitner, W., Lankes, M., Nacke, L.E., Tscheligi, M., 2015. Personalization in serious and persuasive games and gamified interactions. In: Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play - CHI PLAY '15. ACM, pp. 811–816. https://doi.org/10.1145/2793107.2810260.

Connolly, T.M., Boyle, E.A., MacArthur, E., Hainey, T., Boyle, J.M., 2012. A systematic literature review of empirical evidence on computer games and serious games. Comput. Educ. 59, 661–686. https://doi.org/10.1016/j.compedu.2012.03.004.

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

Deci, E.L., Ryan, R.M., 1985. Intrinsic Motivation and Self-Determination in Human Behavior. Plenum, New York London.

Deterding, S., 2015. The Lens of intrinsic skill atoms: a method for gameful design. Hum. Comput. Interact. 30, 294–335. https://doi.org/10.1080/07370024.2014.993471.

Deterding, S., Dixon, D., Khaled, R., Nacke, L.E., 2011. From game design elements to gamefulness: defining "gamification. In: Proceedings of the 15th International Academic MindTrek Conference. Tampere, Finland. ACM, pp. 9–15. https://doi.org/10.1145/2181037.2181040.

Dixon, D., 2011. Player types and gamification. In: CHI 2011 Workshop Gamification: Using Game Design Elements in Non-Game Contexts, pp. 12–15.

Ferro, L.S., Walz, S.P., Greuter, S., 2013. Towards personalised, gamified systems: an investigation into game design, personality and player typologies. In: Proceedings of the 9th Australasian Conference on Interactive Entertainment: Matters of Life and Death - IE '13, pp. 1–6. https://doi.org/10.1145/2513002.2513024.

Goldberg, L.R., 1993. The structure of phenotypic personality traits. Am. Psychol. 48, 26. https://doi.org/10.1037/0003-066X.48.1.26.

Grant, A.M., 2008. The Significance of task significance: job performance effects, relational mechanisms, and boundary conditions. J. Appl. Psychol. 93, 108–124. https://doi.org/10.1037/0021-9010.93.1.108.

Hamari, J., Koivisto, J., Sarsa, H., 2014. Does gamification work? - a literature review of empirical studies on gamification. In: Proceedings of the Annual Hawaii International Conference on System Sciences, pp. 3025–3034. https://doi.org/10.1109/HICSS.2014.377.

Hamari, J., Tuunanen, J., 2014. Player types: a meta-synthesis. Trans. Digit. Games Res. 1.

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

Jia, Y., Xu, B., Karanam, Y., Voida, S., 2016. Personality-targeted gamification: a survey study on personality traits and motivational affordances. In: Proceedings of the 34th Annual ACM Conference on Human Factors in Computing Systems - CHI '16, https://doi.org/10.1145/2858036.2858515.

Johnson, D., Deterding, S., Kuhn, K.-A., Staneva, A., Stoyanov, S., Hides, L., 2016. Gamification for health and wellbeing: a systematic review of the literature. Internet Interv. 6, 89–106. https://doi.org/10.1016/j.invent.2016.10.002.

Kaptein, M., De Ruyter, B., Markopoulos, P., Aarts, E., 2012. Adaptive persuasive systems: a study of tailored persuasive text messages to reduce snacking. ACM Trans. Interact. Intell. Syst. 2, 1–25. https://doi.org/10.1145/2209310.2209313.

Kaptein, M., Markopoulos, P., De Ruyter, B., Aarts, E., 2015. Personalizing persuasive technologies: explicit and implicit personalization using persuasion profiles. Int. J. Hum. Comput. Stud. 77, 38–51. https://doi.org/10.1016/j.ijhcs.2015.01.004.

Kline, R.B., 2010. Principles and Practice of Structural Equation Modeling, 3rd ed. The Guilford Press.

Levine, T.R., 2005. Confirmatory factor analysis and scale validation in communication research. Commun. Res. Rep. 22, 335–338. https://doi.org/10.1080/00036810500317730.

Lorenzo-Seva, U., Ferrando, P.J., 2013. FACTOR 9.2: a comprehensive program for fitting exploratory and semiconfirmatory factor analysis and IRT models. Appl. Psychol. Meas. 37, 497–498.

Marczewski, A., 2015a. User types. Even Ninja Monkeys Like to Play: Gamification, Game Thinking & Motivational Design. CreateSpace Independent Publishing Platform, pp. 69–84.

Marczewski, A., 2015b. Gamification mechanics and elements. Even Ninja Monkeys Like to Play: Gamification, Game Thinking & Motivational Design. CreateSpace Independent Publishing Platform, pp. 165–177.

Marczewski, A., 2015c. Dark personalities and the user type hexad [WWW Document]. Gamified UK URL. http://www.gamified.uk/2015/10/26/dark-personalities-and-the-user-type-hexad/ accessed 3.27.16.

Muthén, B., Kaplan, D., 1985. A comparison of some methodologies for the factor analysis of non-normal Likert variables. Br. J. Math. Stat. Psychol. 38, 171–189.

Myers, I.B., 1962. The Myers-Briggs Type Indicator. Consulting Psychologists Press, Palo Alto, CA.

Nacke, L.E., Bateman, C., Mandryk, R.L., 2014. BrainHex: a neurobiological gamer typology survey. Entertain. Comput. 5, 55–62. https://doi.org/10.1016/j.entcom.2013.06.002.

Nov, O., Arazy, O., 2013. Personality-targeted design: theory, experimental procedure, and preliminary results. In: Proceedings of the 2013 Conference on Computer Supported Cooperative Work - CSCW '13. ACM, pp. 977–984. https://doi.org/10.1145/2441776.2441887.

Orji, R., Mandryk, R.L., Vassileva, J., Gerling, K.M., 2013. Tailoring persuasive health games to gamer type. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13, pp. 2467–2476. https://doi.org/10.1145/2470654.2481341.

Orji, R., Nacke, L.E., DiMarco, C., 2017. Towards personality-driven persuasive health games and gamified systems. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '17. ACM, pp. 1015–1027. https://doi.org/10.1145/3025453.3025577.

Orji, R., Vassileva, J., Mandryk, R.L., 2014. Modeling the efficacy of persuasive strategies for different gamer types in serious games for health. User Model. User Adapt. Interact. 24, 453–498. https://doi.org/10.1007/s11257-014-9149-8.

Panayides, P., 2013. Coefficient alpha: interpret with caution. Eur. J. Psychol. 9. https://doi.org/10.5964/ejop.v9i4.653.

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

Raftopoulos, M., Walz, S., Greuter, S., 2015. How enterprises play : towards a taxonomy for enterprise gamification. In: DiGRA 2015: Diversity of Play, pp. 1–17.

Ryan, R.M., Deci, E.L., 2000a. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am. Psychol. 55, 68–78. https://doi.org/10.1037/0003-066X.55.1.68.

Ryan, R.M., Deci, E.L., 2000b. Intrinsic and extrinsic motivations: classic definitions and new directions. Contemp. Educ. Psychol. 25, 54–67. https://doi.org/10.1006/ceps.1999.1020.

Ryan, R.M., Deci, E.L., Vansteenkiste, M., 2016. Autonomy and autonomy disturbances in self-development and psychopathology: research on motivation, attachment, and clinical process. In: Cicchetti, D. (Ed.), Developmental Psychopathology. Wiley, pp. 1–54. https://doi.org/10.1002/9781119125556.devpsy109.

Ryan, R.M., Rigby, C.S., Przybylski, A., 2006. The motivational pull of video games: a self-determination theory approach. Motiv. Emot. 30, 347–363. https://doi.org/10.1007/s11031-006-9051-8.

Schmitt, T.A., 2011. Current methodological considerations in exploratory and confirmatory factor analysis. J. Psychoeduc. Assess. 29, 304–321. https://doi.org/10.1177/0734282911406653.

Seaborn, K., Fels, D.I., 2014. Gamification in theory and action: a survey. Int. J. Hum. Comput. Stud. 74, 14–31. https://doi.org/10.1016/j.ijhcs.2014.09.006.

Tondello, G.F., Wehbe, R.R., Diamond, L., Busch, M., Marczewski, A., Nacke, L.E., 2016. The gamification user types Hexad scale. In: Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play - CHI PLAY '16. Austin, TX, USA. ACM, pp. 229–243. https://doi.org/10.1145/2967934.2968082.

Yee, N., 2016. Gaming motivations align with personality traits [WWW Document]. URL. http://quanticfoundry.com/2016/01/05/personality-correlates/.

Yee, N., 2015. Gamer motivation model overview and descriptions [WWW Document]. URL. http://quanticfoundry.com/2015/12/15/handy-reference/.

Yee, N., 2006. Motivations for play in online games. CyberPsychol. Behav. 9, 772–775. https://doi.org/10.1089/cpb.2006.9.772.

Yee, N., Ducheneaut, N., Nelson, L., 2012. Online gaming motivations scale: development and validation. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 2803–2806. https://doi.org/10.1145/2207676.2208681.

Zeigler-Hill, V., Monica, S., 2015. The HEXACO model of personality and video game preferences. Entertain. Comput. 11, 21–26. https://doi.org/10.1016/j.entcom.2015.08.001.