

# Tailoring Gamification for Adolescents: a Validation Study of Big Five and Hexad in Dutch

Jeroen Ooge

KU Leuven

Department of Computer Science  
Leuven, Belgium  
jeroen.ooge@kuleuven.be

Katrien Verbert

KU Leuven

Department of Computer Science  
Leuven, Belgium  
katrien.verbert@kuleuven.be

Robin De Croon

KU Leuven

Department of Computer Science  
Leuven, Belgium  
robin.decroon@kuleuven.be

Vero Vanden Abeele

KU Leuven

Department of Computer Science  
Leuven, Belgium  
vero.vandenabeele@kuleuven.be

## ABSTRACT

Gamification researchers deem adolescents a particularly interesting audience for tailored gamification. However, empirical validation of popular player typologies and personality trait models thus far has been limited to adults. As adolescents exhibit complex behaviours that differ from older adults, these models may need adaptation. To that end, we collected a unique data set of Big Five Inventory and Hexad questionnaire answers in Dutch from 402 adolescents. Confirmatory factor analysis showed that the Dutch forms of the BFI-10, BFI-44 and Hexad scales performed substandard when used with adolescents. Through exploratory factor analysis, we investigated underlying problems, and provide preliminary suggestions on how to improve measurements. In particular, we propose to simplify the Hexad model, and to reformulate specific items. With this study, we hope to contribute to the debate on how to improve the tailoring of interactive systems for adolescents.

## CCS CONCEPTS

• **Human-centered computing** → **HCI theory, concepts and models; Empirical studies in HCI.**

## KEYWORDS

gamification; adolescents; tailoring; personalisation; gameful design; Hexad; Big Five; scale validation

## ACM Reference Format:

Jeroen Ooge, Robin De Croon, Katrien Verbert, and Vero Vanden Abeele. 2020. Tailoring Gamification for Adolescents: a Validation Study of Big Five and Hexad in Dutch. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play (CHI PLAY '20), November 2–4, 2020, Virtual*

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).  
*CHI PLAY '20, November 2–4, 2020, Virtual Event, Canada*

© 2020 Association for Computing Machinery.  
ACM ISBN 978-1-4503-8074-4/20/11...\$15.00  
<https://doi.org/10.1145/3410404.3414267>

*Event, Canada.* ACM, New York, NY, USA, 13 pages. <https://doi.org/10.1145/3410404.3414267>

## 1 INTRODUCTION

In the past decade, we have witnessed the rise of gamification, *the intentional use of game elements for non-gaming tasks and in non-gaming contexts, to create a gameful experience* [e.g. 24, 39, 88]. Gamification was found an effective strategy for engaging users and promoting desirable behaviours in a wide range of contexts, for example business and marketing [39, 45], health and well-being [51, 77, 85], education [25, 92] and sustainability [52].

As the gamification field matures, researchers are increasingly questioning the effectiveness of one-size-fits-all approaches [73]. A growing number of studies suggests that gamification strategies are likely to result in higher engagement, and to improve desired outcomes, when adapted to end users [1, 55, 69, 77]. In the quest towards tailored gamification, demographic characteristics such as age and gender, and models of personality traits (e.g. the Big Five Inventory [35]) and player styles (e.g. BrainHex [71, 72], Hexad [98]) play a significant role [10, 56, 74]. Many studies have investigated user typologies and their associated scales to predict preferences for different gamification elements [3, 38, 46, 95]. Overall, studies yield mixed to tentatively positive outcomes [38].

In particular, a considerable number of gamification studies addresses *adolescents* [e.g. 12, 13, 19, 21, 34, 37, 57, 66, 76, 84, 90, 91] because this age group is considered especially receptive to gamification. As savvy users of interactive and socially connected technologies, adolescents are formed by today's media and ludification of culture [8]. Their potential to shape themselves and society's future [32] makes them an interesting target for gamification.

Interestingly, even though gamification researchers deem adolescents a particularly interesting audience, they have not yet investigated the validity of player typologies, personality trait models, and associated scales for them; empirical investigations and validations thus far have been limited to adults. In fact, not only gamification studies have insufficiently considered adolescents: teenagers (aged 13–19) present an underexplored space in the whole of Human-Computer Interaction (HCI) research [33, 83]. While HCI researchers often treat adolescents as *young adults*, adolescence is a distinct developmental stage marked by a strong need

for autonomy, becoming independent from parents, and building meaningful relationships with peers instead [61]. Moreover, puberty comes with rapid hormonal and neurological changes, which result in increased exploration and risk-taking as brains have not yet fully matured [6, 60, 75]. Since adolescents typically exhibit complex behaviours that differ from older adults [33], current personality trait and player models may need adaptation.

Hence, it is important to verify whether widely used models such as the Big Five Inventory (BFI) [35, 47, 50] and the Hexad framework [98] are suitable for tailoring gamified systems for adolescents. To that end, our study collected responses on Dutch forms of the BFI-44 and Hexad questionnaires from 402 teenagers between 13 and 19 years old, as part of a larger study on personalised gamification on an online mathematics platform. Through confirmatory and exploratory factor analyses, we investigated the validity of the BFI and Hexad models in the adolescent population. The results bring forward multiple issues that call for increasing carefulness when studies with adolescents rely on user scales that are validated for adults.

Our contributions to the HCI research field are threefold. First, we collect a unique data set of BFI-44 and Hexad questionnaire submissions in Dutch from 402 adolescents, which we publish online together with an expert-approved Dutch translation of the Hexad questionnaire<sup>1</sup>. Second, upon a careful confirmatory factor analysis, we show that the current forms of the BFI-10, BFI-44 and Hexad scales are inadequate when used with adolescents. Third, we execute an exploratory factor analysis to provide preliminary suggestions on how to improve the models. In sum, we hope these findings contribute to the discussion on how personality trait models and player typologies can be tailored for adolescents.

## 2 BACKGROUND

This section briefly reviews approaches towards tailored gamification, focusing on the Hexad and Big Five models, and introduces the challenges of HCI research with adolescents.

### 2.1 Tailored Gamification

Gamification strategies aim to motivate and engage people in non-entertainment contexts [24, 88]. Initially, studies centred around whether gamification works as a whole, and did not yet tailor design to users [73]. However, researchers quickly found that applying gamification techniques leads to mixed results, depending on the type of end user and the particular context [e.g. 29, 39, 68, 88]. This questioned the legitimacy of one-size-fits-all approaches.

The current wave in gamification research investigates why, how and when gamification methods are effective, taking the end user’s personal characteristics into account. Gamified applications are, for example, personalised based on psychological inventories such as Big Five [35] and Minnesota Multiphasic Personality [11], and user typologies specifically developed for game(ful) design, e.g. BrainHex [71, 72] and Hexad [98]. Research on how these personality traits and player types are associated with gamification strategies revealed interesting yet overall weak correlations [46, 76, 94, 96, 98]. An increasingly studied topic is the influence of gender and (adult) age [7] on the preference for gamification elements: women seem

**Table 1: Hexad gamification user types and their associated motivations based on [98].**

	User type	Motivated by
F	Free Spirit	autonomy as in SDT, i.e. feeling in control of a situation
Ac	Achiever	competence as in SDT, i.e. feeling skilled enough to accomplish tasks
S	Socialiser	relatedness as in SDT, i.e. being involved with others, social connections
Ph P	Philanthropist Player	purpose, i.e. accomplishing meaningful tasks extrinsic rewards, i.e. rewards that are independent of the activity
D	Disruptor	change, i.e. disrupting the system and testing its boundaries

to score slightly higher than men on intrinsic motivation types; and intrinsic motivations slightly increase with age [69], contrary to extrinsic motivations [95]. The latter finding suggests that age matters, hence also adolescents’ motivations to interact with gamified systems may differ from those of adults.

### 2.2 Hexad Model

Gamification research commonly types users with the Hexad framework, which draws on Self Determination Theory (SDT) [22] to define four intrinsically motivated user types (*Free Spirit*, *Achiever*, *Socialiser* and *Philanthropist*) and one extrinsically motivated user type (*Player*). One extra user type (*Disruptor*) is empirically derived from user behaviour in online systems [64]. Hexad comes with a 7-point Likert scale [98], which consists of the 24 items (six times four) in Table 11. The individual scores on the six different types can be used to tailor gamified applications since each user type is associated with preferences towards specific kinds of gamification elements, as shown in Table 1.

Recently, Hexad was found to outperform BrainHex and the Big Five in identifying user preferences for gamification elements [38], and empirical validation of the Hexad framework with adults supports the scale’s reliability and internal consistency [2, 95, 98]. This suggests that Hexad is a suitable instrument to tailor gamification strategies to user characteristics [82]. However, studies validating Hexad have not yet recruited participants under the age of 18.

### 2.3 Big Five Model

The Big Five, also known as the Five Factor Model, is a taxonomy of five personality traits (*Openness*, *Conscientiousness*, *Extraversion*, *Agreeableness* and *Neuroticism*), which each combine many personality characteristics [50], as summarised in Table 2. The descriptive model arose after decades of research [35, 47, 50, 89] and can be assessed by a wide range of questionnaires, including NEO PI-R [17], FIPI and TIPI [36], the Big Five Inventory (BFI) with 44 items (BFI-44) [49, 50], and a shortened version of the latter with ten items (BFI-10) [81]. Table 12 contains the items of BFI-10 and BFI-44.

HCI researchers interested in personality increasingly incorporate the Big Five in their research to allow comparison of their findings with others [23, 38, 46, 98]. Most HCI studies administer

<sup>1</sup><https://github.com/JeroenOoge/tailoring-gamification-adolescents>

**Table 2: Big Five personality traits based on [46, 50, 76].**

	Trait	Description
O	Openness	devise novel ideas; hold unconventional values; willingly question authority; be curious, imaginative, creative
C	Conscientiousness	actively plan, organise and carry out tasks; be self-disciplined, goal-oriented, dependable
E	Extraversion	seek out new opportunities and excitement; be outgoing, expressive, ambitious
A	Agreeableness	help others and expect help in return; be considerate, cooperative, tolerant, helpful, friendly, caring
N	Neuroticism	be fearful, sad, embarrassed, distrustful; have difficulty managing stress; be nervous, sensitive, emotionally unstable

compact scales such as BFI-10 [e.g. 76, 98]. Whereas this choice is understandable from a pragmatic perspective, BFI-10 leads to substantial losses in captured variance, retest reliability and convergent validity compared to BFI-44 (especially for Agreeableness) and should therefore only be used when participant time is extremely limited [81]. Alternative Big Five personality scales have been developed specifically for adolescents [48, 63], but these do not seem popular in current HCI research. Moreover, we are unaware of any validations of the BFI-10 and BFI-44 scales with adolescents.

## 2.4 Tailoring for Adolescents

Given their unique developmental stage, adolescents may not fully comply with personality and player style models for adults. No longer children, not yet adults [61], teenagers rapidly undergo profound psychological and physical changes [62, 79] at the onset of puberty and sexuality. Their *raging hormones* trigger emerging adulthood, including the need for a personal identity, parental distancing and closer interaction with peer groups, and increased sensation-seeking and risk-taking behaviour [5]. It is therefore plausible that adolescents score higher on factors measuring autonomy or relatedness, and factors associated with exploratory or risky behaviour.

Moreover, adolescents do not necessarily interpret or process questionnaires in the same way as adults. As they extensively explore their selves (teenagers are often perceived as self-obsessed [5]), surveys that poll for personality or type may sharpen their self-consciousness and bias their self-reporting. Additionally, adolescents are typically preoccupied with peer relationships and highly susceptible to peer influence, which may stimulate answering that is desirable in the eyes of peers. This socially desirable answering can be further reinforced when questionnaires are filled out in the presence of others, for example in a school context. Finally, while adolescents already think in a more logical, abstract way than children, and have a more elaborate vocabulary, their cognitive abilities are still different from those of adults. This may complicate cognitive processing and understanding of scale items that were validated with adults only.

In sum, models and instruments developed for adults may not be appropriate for adolescents [33]. There is a need for incorporating *TeenCI* [83] in HCI to better understand teenagers and develop suitable methods for them. We believe this need is particularly urgent in the gamification domain because adolescents are a popular target for gamified services and apps. Yet, we are unaware of studies that address the validity of models currently used to tailor for teenagers. Our study tries to fill this gap by investigating the widely used Big Five Inventory and Hexad models.

## 3 STUDY DESIGN AND METHODS

We empirically studied the internal reliability and validity of the Big Five Inventory and Hexad scales with adolescents (aged 13–19).

### 3.1 Instruments

For the Big Five Inventory, we used a validated Dutch translation [23] of the BFI-44 questionnaire. For Hexad, lacking such a validated version, we translated the original scale [98] to Dutch ourselves. Three fluent English language experts independently back-translated the items and reported no meaningful differences with the original ones, thereby confirming the adequacy of our translation in Table 11. For both questionnaires, the items were randomly ordered, as instructed in the original scales, and a control question was inserted halfway.

Before administering the questionnaires, two think-aloud studies, each involving five adolescents, verified the study procedure. Some participants did not understand ‘status quo’ and ‘provoke’ in Hexad’s items D1 and D2. To mitigate confusion, yet respect the validated items, we added tooltips to the online questionnaire, explaining ‘status quo’ as ‘the state of affairs; the things as they are’ and ‘to provoke’ as ‘to challenge’. The questionnaires’ implementation is available online.

### 3.2 Data Collection

We contacted the board of eleven secondary schools in Belgium (Flanders) via email. Six schools expressed interest, and thirteen teachers invited their students from years 3–6 (equivalent to US high school) to fill out the BFI-44 and Hexad questionnaires online before accessing a gamified mathematics platform.

An information leaflet describing the research objectives was handed out to all students beforehand. Students had the opportunity to refuse cooperation, and all teachers declared in an informed consent not to coerce anyone in participating. Additional measurements protected the privacy of all underage participants: parents agreed in an informed consent with the processing of their teenage children’s data, and the ethical committee of KU Leuven approved the study (reference number G-2019 04 1618).

### 3.3 Data Analysis

Internal reliability was examined with Cronbach’s  $\alpha$ . Data validity was assessed through factor analysis [15], which aims to describe *indicators* (observed variables) in terms of fewer latent *factors* (unobserved variables). *Exploratory factor analysis* (EFA) is suitable for exploring observed data when little or no theoretical knowledge is available about potential latent factors. When such prior knowledge does exist, *confirmatory factor analysis* (CFA) is used

**Table 3: Fit indices and their cut-off levels as used for CFA model evaluation in this study.**

Index family	Index	Description	Cut-off	Cut-off
Absolute fit	$\chi^2/df$	Chi-squared test statistic divided by the degrees of freedom	< 3	< 5
	SRMR	Standardised root mean square residual, values $\in [0, 1]$ where 0 is perfect fit	$\leq 0.08$	$\leq 0.10$
	RMSEA	Root mean square error of approximation, values $\in [0, \infty)$ where 0 is perfect fit	$\leq 0.08^*$	$\leq 0.10^*$
	GFI	Goodness of fit index, values $\in [0, 1]$ where 1 is perfect fit	$\geq 0.95$	$\geq 0.90$
	AGFI	Adjusted goodness of fit index, interpretation similar to GFI	$\geq 0.95$	$\geq 0.90$
Comparative fit	CFI	Comparative fit index, values $\in [0, 1]$ where 1 is perfect fit	$\geq 0.95$	$\geq 0.90$
	TLI	Tucker-Lewis index, interpretation similar to CFI but values $\notin [0, 1]$ are possible	$\geq 0.95$	$\geq 0.90$

Note: Green cut-offs indicate good fit, yellow cut-offs fair fit. \*Additional support for good fit if the upper limit of the 90%-confidence interval is below the cut-off.

for theory-testing, whereby one (or a small number of) plausible model(s) is fitted to the data. CFA is typically used in later phases of scale development, after the underlying structure has been established by theory and verified through EFAs [9]. Our CFA and EFA for the BFI-10, BFI-44 and Hexad models followed the procedural steps recommended in [9, 58] and were executed with the open source software R (3.6.2), using the packages lavaan (0.6-5) and psych (1.9.12.31). All R scripts are available online.

### 3.4 Confirmatory Factor Analysis

The CFA used *maximum likelihood* (ML) estimation, the most common approach in the literature [9, 16, 20, 27]. ML estimation assumes continuous indicators with a multivariate normal distribution, a sufficiently large sample, and independence between indicators and corresponding errors. These assumptions must be checked before running ML estimations because severe violations distort parameter estimates [9]. We verified that  $|\text{skewness}(i)| \leq 2$  and  $|\text{kurtosis}(i)| \leq 2$  for all scale items  $i$  because ordinal Likert-type data such as our BFI and Hexad data can then be treated as continuous [28, 30, 70, 99]. To detect extremely non-normal items, we visually inspected QQ-plots and histograms [100]. Finally, we checked that the mean absolute value of the skewnesses was smaller than 1, and that items had absolute skewness and kurtosis below 1 (more stringent bound than for continuity) because then ML estimation is not dissuaded and little distortion is to be expected [70].

In CFA, models are evaluated in terms of various fit indices, which reflect how well the proposed model reproduces the observed correlation matrix. While these indices and their corresponding acceptable cut-off values are still hotly debated in the CFA literature, consensus has grown that researchers should take into account fit indices from different families, which each reflect a specific model aspect [9, 58, 67]. Specifically, *absolute fit indices* indicate the degree of discrepancy between the observed correlation matrix and the one predicted by the model, thus assessing model fit on a global level, whereas *comparative fit indices* compare the specified solution to a simple baseline model. Table 3 summarises the fit indices and cut-off values used in this study, recommended by [42, 44, 58, 65].

### 3.5 Exploratory Factor Analysis

As a preliminary step for EFA, the Kaiser-Meyer-Oblin (KMO) test [54] and Bartlett's test of sphericity ensured the factorability [4]

of the BFI and Hexad data. The EFA used ML estimation for factor extraction, requiring multivariate normality as in the CFA.

Determining the number of factors is critical in EFA [40]. First, in accordance with existing theory on the Big Five and Hexad models, we investigated solutions with five and six factors for the BFI and Hexad data respectively. Next, departing from existing theory, multiple criteria were consulted to empirically decide on a number of factors more suitable to explain the data's underlying structure [31, 40]: the Kaiser-Guttman criterion [53], Cattell's scree test [14] and parallel analysis [43] with 100 simulations.

Factor rotation [78] in EFA leads to more interpretive models and encompasses *orthogonal rotation* methods, which yield uncorrelated factors, and *oblique rotation* methods, which allow inter-factor correlation. As the latter lead to more accurate and realistic results [31, 67, 87], our EFA used oblique promax rotation, unless stated otherwise. Rotated EFA solutions were interpreted in terms of meaningfulness, communalities, total explained variance, and factor loadings and cross-loadings in the pattern matrix. In general,  $|\text{factor loading}| \geq 0.70$  is ideal because the corresponding factor then accounts for nearly 50% of the item's variance. An item's communality is the sum of its squared loadings, which measures the amount of variance explained by the common factors. Hence, items with communality under 0.40 fail to load significantly on any factor and are candidates for removal when optimising the model [18].

## 4 RESULTS

This section first presents the steps taken to screen the collected data, and provides descriptive statistics. Next, the results of the CFA show how the empirical data fitted to the theoretical BFI-10, BFI-44 and Hexad models, comparing different fit metrics to commonly accepted standards. As these results will prove to be substandard, the last part of this section presents the results of an EFA in which we tried to understand the underlying issues, and to identify latent factors that reconstruct the complexity of the observed data better than the existing BFI and Hexad models.

### 4.1 Data Screening and Descriptives

In total, 402 adolescents (aged 13–19) from years 3–6 of secondary school participated in the study between September 2018 and May 2019. Only 38 of them belonged to years 5–6 (US junior–senior, aged 17–19). To obtain a more focused sample, these were removed. Hence, the data screening solely considered 363 adolescents from

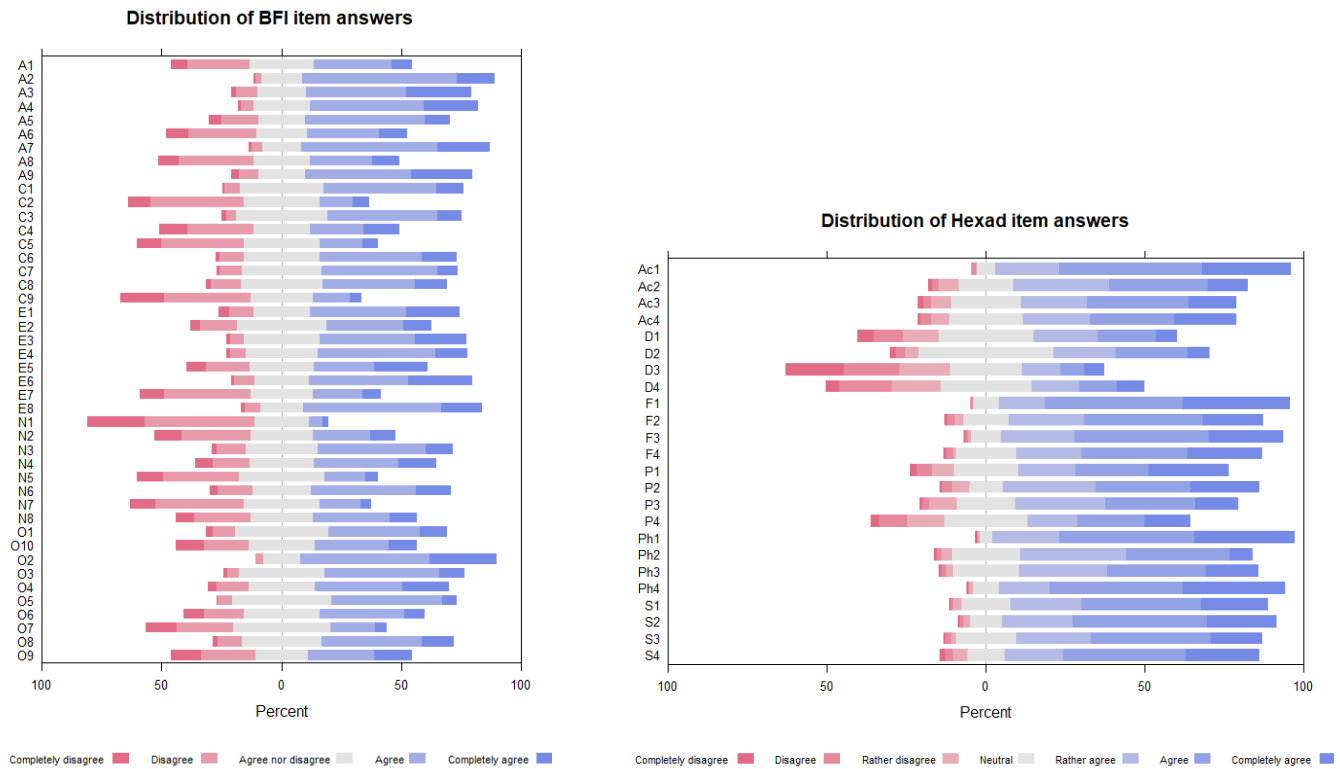


Figure 1: Diverging stacked bar charts for 297 BFI-44 (left) and 293 Hexad (right) questionnaire submissions.

Table 4: Descriptive statistics for the screened BFI-44 and Hexad data.

	O	C	E	A	N	F	Ac	S	Ph	P	D
mean	2.80	3.09	3.28	3.44	2.76	22.68	21.36	22.12	22.32	20.30	16.45
variance	0.29	0.36	0.44	0.28	0.38	8.71	11.42	14.20	10.05	16.37	18.21
skewness	-0.19	0.18	-0.33	-0.39	-0.04	-0.68	-0.24	-0.96	-0.74	-0.27	0.30
kurtosis	-0.01	-0.30	-0.04	0.54	-0.37	1.04	-0.19	1.45	1.70	0.10	-0.08
Cronbach's $\alpha$	0.74	0.80	0.83	0.68	0.74	0.56	0.61	0.81	0.74	0.66	0.64

Note: Values  $\alpha < 0.70$  are in red.

years 3–4 (US freshman–sophomore, aged 13–17, average = 14.96, SD = 0.75; excluding one participant who declared an age of five).

Data screening was performed on the raw data, reverse-scored whenever necessary. We considered participants attentive if they correctly answered the control question in the questionnaires, did not provide identical answers for all other items, and completed all items. These inclusion criteria were met by 297 BFI-44 submissions (66 removed) and 293 Hexad submissions (70 removed).

Next, we inspected data descriptives and internal consistency of the factors. Figure 1 visualises the Likert-type answers with diverging stacked bar charts [41]. The violin plots in Figure 2 show the mutual distribution of the Big Five personality traits and the Hexad gamification user types, and Table 4 reports descriptive statistics and the internal reliability (Cronbach's  $\alpha$ ). The Cronbach's  $\alpha$  levels show low internal consistency for the gamification user types: four

out of six types score below 0.70, which is considered the threshold for acceptable internal consistency [86], with Free Spirit scoring particularly low.

## 4.2 Confirmatory Factor Analysis

Before testing the hypothesised Big Five and Hexad models in a CFA, we checked the requirements for ML estimation. None of the items had  $|\text{skewness}| > 2$ , and only Ph1 and A2 had  $|\text{kurtosis}| > 2$ . In fact, only six Hexad items had  $|\text{skewness}| > 1$ ; five BFI-44 and seven Hexad items had  $|\text{kurtosis}| > 1$ . In all QQ-plots, the observed quantiles approximated the theoretical ones fairly well. Moreover, the mean absolute skewnesses were 0.38 and 0.68 for BFI-44 and Hexad respectively. Hence, our data were appropriate for CFA with ML estimation.

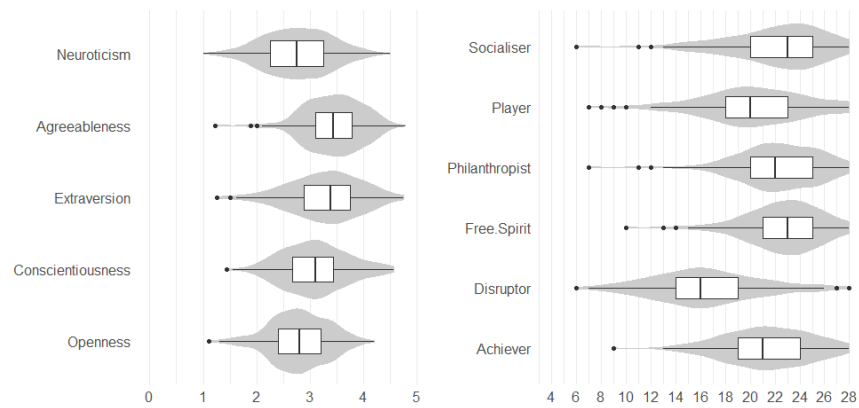


Figure 2: Violin plots for 297 BFI-44 (left) and 293 Hexad (right) questionnaire submissions.

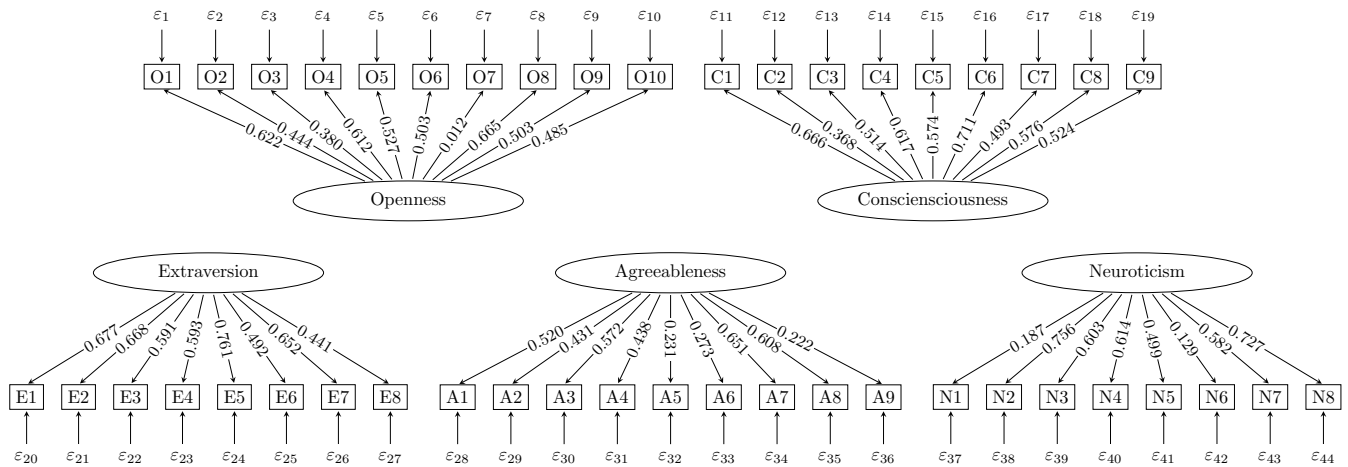


Figure 3: Path diagram for the BFI-44 model. Circles and rectangles represent factors and indicators respectively; arrows indicate direct effects. Factor covariations and error terms  $\epsilon_i$  are suppressed for clarity.

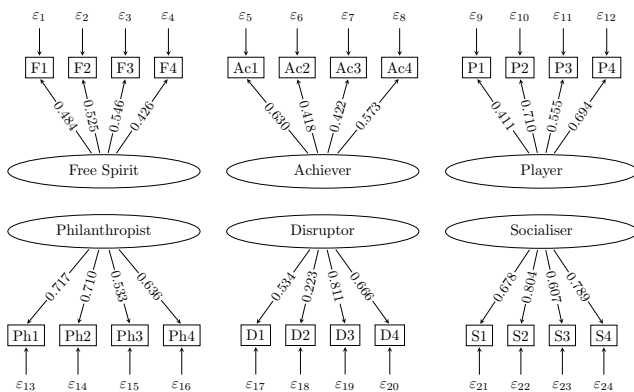


Figure 4: Path diagram for the Hexad model, with encodings similar to Figure 3.

A CFA with freely estimated factor loadings and error terms yielded low factor loadings: only four BFI-44 loadings and six Hexad loadings were larger than the desirable 0.7, as depicted in Figures 3 and 4. Table 5 contains the fit indices for the BFI-10, BFI-44 and Hexad models, and is color coded along the recommended cut-offs in Table 3. Only the BFI-10 model has an acceptable absolute fit: the GFI and AGFI indices for BFI-44 and Hexad are far below the required cut-offs. Moreover, the low comparative CFI and TLI indices show that none of the models fit remarkably better to the data than a baseline model where all covariances are set to zero.

### 4.3 Exploratory Factor Analysis

Since the models proved to poorly fit the data in the CFA, we executed an additional EFA to identify problematic items, and to analyse the factor structure underlying the data.

We initially verified the data factorability. Bartlett's sphericity test was significant ( $p < 0.001$ ) for the BFI-10 ( $\chi^2_{45} = 324.50$ ), BFI-44 ( $\chi^2_{946} = 4636.76$ ) and Hexad ( $\chi^2_{276} = 2143.23$ ) data. The

**Table 5: Fit indices for BFI-10, BFI-44 and Hexad.**

Index	BFI-10	BFI-44	Hexad
$\chi^2$ df, p	60.59 25, 0.00	2412.74 892, 0.00	679.12 237, 0.00
$\chi^2$ /df	2.42	2.70	2.87
SRMR	0.06	0.10	0.09
RMSEA 90%-CI	0.07 [0.05, 0.09]	0.08 [0.07, 0.08]	0.08 [0.07, 0.09]
GFI	0.96	0.68	0.84
AGFI	0.91	0.65	0.80
CFI	0.88	0.62	0.77
TLI	0.78	0.59	0.73

Note: Green, yellow and red cells respectively indicate good, fair and bad fit, according to the recommended cut-offs.  $\chi^2$  is only included for reference as relying on it is discouraged [e.g. 9, 30, 42]. CI = confidence interval

KMO scores for BFI-44 and Hexad were also ‘meritorious’ [54] (0.80 and 0.79 respectively) so we considered the data suitable for factorisation. Although the BFI-10 data had a ‘miserable’ [54] KMO index (0.60), we pursued the EFA for this scale too to compare our results with [81].

**4.3.1 EFA with existing factor structures.** First, we performed EFA in line with existing theories, i.e. five factors for the BFI data and six factors for the Hexad data. Table 6 presents the rotated factor loadings of the BFI-44 items. Even though most items had loadings under 0.70 and communality under 0.40, some items loaded on multiple or no factors, and the total explained variance was only 0.36, the expected Big Five factor structure was clearly visible. Moreover, factors were barely correlated (at most 0.23).

Table 7 shows the loading results of the Hexad items. Most communalities were at least 0.40, but eight items loaded most heavily on an unexpected factor. The latter was particularly notable for items F1 to F4, whose largest absolute loadings were not in column F. In addition, five items had strong cross-loadings, i.e. the two largest absolute loadings differed less than 0.20. Due to these issues, the Hexad factor structure was only vaguely recognisable. In addition, the total explained variance was only 0.45, and three factor pairs had correlations over 0.30: Ph–Ac (0.50), Ph–S (0.54) and P–D (0.49).

The EFA for BFI-10 used oblique geomin rotation since promax rotation led to a Heywood case [9, 58]: N2 had a loading > 1. Table 8 shows cross-loading or low communality for at least one item in each personality trait. The factor pair A–E had the largest absolute correlation (0.37), and the total explained variance was 0.43.

**4.3.2 EFA with alternative factor structures.** Second, we departed from existing theories, and iteratively modified the items included in the scales to better reconstruct the observed data’s complexity. We only explored models for the BFI-44 and Hexad data because BFI-10 had too few items for pruning. In each iteration, starting from Tables 6 and 7, we removed one problematic item and re-computed loadings. In decreasing order of severity, removal criteria were: no significant loading on any factor, strong cross-loading, loading on an unexpected factor, and communality under 0.4. The Kaiser-Guttman criterion, Cattell’s scree test and parallel analysis gave inconclusive suggestions for the number of factors: between six and seven for BFI-44 and between three and seven for Hexad.

**Table 6: Rotated factor loadings and communalities based on the correlation matrix of the BFI-44 data.**

Item	1 (O)	2 (C)	3 (E)	4 (A)	5 (N)	Com.
O1	<b>0.460</b>					0.34
O2	<b>0.470</b>					0.29
O3	<b>0.489</b>					0.36
O4	<b>0.506</b>					0.33
O5	<b>0.457</b>					0.37
O6	<b>0.486</b>					0.24
O7						0.05
O8	<b>0.597</b>					0.39
O9	<b>0.376</b>					0.17
O10	<b>0.458</b>					0.21
C1		<b>0.677</b>				0.46
C2	-0.331	<b>0.403</b>				0.28
C3		<b>0.481</b>				0.34
C4		<b>0.650</b>				0.45
C5		<b>0.561</b>				0.40
C6		<b>0.659</b>				0.52
C7		<b>0.471</b>				0.35
C8		<b>0.566</b>				0.42
C9		<b>0.575</b>				0.40
E1			<b>0.598</b>	-0.310		0.46
E2			<b>0.677</b>			0.51
E3			<b>0.595</b>			0.47
E4			<b>0.566</b>			0.47
E5			<b>0.757</b>	-0.362		0.61
E6			<b>0.328</b>	-0.302	-0.205	0.31
E7			<b>0.637</b>			0.48
E8			<b>0.498</b>			0.46
A1				<b>0.598</b>		0.41
A2	0.237			<b>0.327</b>		0.22
A3				<b>0.567</b>		0.36
A4				<b>0.414</b>		0.20
A5			<b>0.259</b>	0.248		0.20
A6			<b>0.484</b>			0.32
A7				<b>0.521</b>		0.41
A8				<b>0.609</b>		0.40
A9			<b>0.372</b>	0.245		0.27
N1			<b>-0.477</b>			0.30
N2					<b>0.771</b>	0.58
N3					<b>0.614</b>	0.39
N4					<b>0.639</b>	0.48
N5					<b>0.541</b>	0.30
N6				<b>-0.386</b>		0.16
N7					<b>0.616</b>	0.40
N8					<b>0.691</b>	0.50

Note: |Loadings| ≤ 0.30 are suppressed unless the absolute difference with the maximal loading (bold) ≤ 0.20. Maximal loadings belonging to an unexpected factor and communalities < 0.40 are in red.

Tables 9 and 10 contain the pattern matrices resulting from our pruning procedure. For Hexad, all Free Spirit items were removed due to persistent cross-loadings, and the number of factors was further reduced to four because all Philanthropist and Socialiser items strongly cross-loaded. The adapted BFI-44 and Hexad models had no factor pairs with a |correlation| > 0.40, except P–D (0.48) and PhS–Ac (0.41).

**Table 7: Rotated factor loadings and communalities based on the correlation matrix of the Hexad data.**

Item	1 (F)	2 (Ac)	3 (S)	4 (Ph)	5 (P)	6 (D)	Com.
F1				<b>0.546</b>			<b>0.26</b>
F2			0.218		<b>0.345</b>		<b>0.29</b>
F3	0.370			<b>0.396</b>			0.40
F4				<b>0.357</b>		0.279	<b>0.30</b>
Ac1				<b>0.463</b>			<b>0.37</b>
Ac2		<b>0.575</b>					0.47
Ac3		<b>0.689</b>					0.43
Ac4		<b>0.497</b>					0.42
S1			<b>0.590</b>				0.48
S2			<b>0.898</b>				0.72
S3		0.351	<b>0.637</b>				0.50
S4			<b>0.760</b>				0.63
Ph1			0.330	<b>0.530</b>			0.51
Ph2				<b>0.528</b>			0.45
Ph3				<b>0.669</b>			0.40
Ph4	-0.233		<b>0.420</b>	0.397			0.47
P1	<b>0.544</b>						0.44
P2					<b>0.604</b>		0.47
P3					<b>0.516</b>		<b>0.34</b>
P4					<b>0.969</b>		0.76
D1						<b>0.506</b>	<b>0.32</b>
D2							<b>0.16</b>
D3						<b>0.889</b>	0.65
D4						<b>0.719</b>	0.46

Note: |Loadings| ≤ 0.30 are suppressed unless the absolute difference with the maximal loading (bold) ≤ 0.20. Maximal loadings belonging to an unexpected factor and communalities < 0.40 are in red.

**Table 8: Rotated factor loadings and communalities based on the correlation matrix of the BFI-10 data.**

Item	1 (O)	2 (C)	3 (E)	4 (A)	5 (N)	Com.
O4	<b>0.518</b>					0.41
O9	<b>0.546</b>					<b>0.32</b>
C1		<b>0.486</b>				<b>0.24</b>
C5		<b>0.718</b>				0.52
E2			0.757			0.62
E8				<b>0.648</b>		0.53
A1		<b>0.326</b>	-0.182			<b>0.14</b>
A5				<b>0.331</b>		<b>0.12</b>
N2					<b>0.998</b>	1.00
N8			-0.358		<b>0.506</b>	0.44

Note: |Loadings| ≤ 0.30 are suppressed unless the absolute difference with the maximal loading (bold) ≤ 0.20. Maximal loadings belonging to an unexpected factor and communalities < 0.40 are in red.

## 5 DISCUSSION

This study set out to empirically investigate whether the Big Five model of personality traits and the Hexad model of gamification user types are suitable instruments when tailoring gamified systems for adolescents. To that end, we collected, carefully screened and consequently analysed a unique data set of BFI-44 and Hexad questionnaire submissions from 402 adolescents.

**Table 9: Rotated factor loadings and communalities based on the correlation matrix of the adapted BFI-44 data.**

Item	1 (O)	2 (C)	3 (E)	4 (A)	5 (N)	Com.
O1	<b>0.494</b>					<b>0.38</b>
O4	<b>0.576</b>					0.41
O5	<b>0.360</b>	0.202			-0.223	<b>0.36</b>
O6	<b>0.670</b>					0.40
O8	<b>0.524</b>					<b>0.35</b>
O9	<b>0.628</b>					<b>0.37</b>
O10	<b>0.633</b>					<b>0.35</b>
C1		<b>0.674</b>				0.45
C3		<b>0.606</b>				<b>0.37</b>
C4		<b>0.511</b>				<b>0.35</b>
C5		<b>0.454</b>		0.274		<b>0.35</b>
C6		<b>0.747</b>				0.56
C7		<b>0.545</b>				<b>0.35</b>
C8		<b>0.632</b>				0.43
C9		<b>0.522</b>				<b>0.33</b>
E1			<b>0.642</b>			0.47
E2			<b>0.747</b>			0.53
E3			<b>0.594</b>			0.42
E4			<b>0.587</b>			0.46
E5			<b>0.813</b>			0.62
E6			<b>0.350</b>	-0.222		<b>0.29</b>
E7			<b>0.661</b>			0.46
E8			<b>0.490</b>			0.40
A1				<b>0.653</b>		0.47
A3				<b>0.612</b>		0.42
A4				<b>0.343</b>		<b>0.12</b>
A7				<b>0.395</b>		<b>0.28</b>
A8				<b>0.644</b>		0.45
N2					<b>0.774</b>	0.59
N3					<b>0.603</b>	<b>0.38</b>
N4					<b>0.614</b>	0.49
N5					<b>0.528</b>	<b>0.29</b>
N7					<b>0.611</b>	0.40
N8					<b>0.689</b>	0.50

Note: |Loadings| ≤ 0.30 are suppressed unless the absolute difference with the maximal loading (bold) ≤ 0.20. Communalities < 0.40 are in red.

### 5.1 BFI-10, BFI-44 and Hexad: Unfit for Adolescents

A rigorous CFA yielded fit indices below recommended thresholds: absolute indices were only found acceptable for BFI-10; comparative fit indices were problematic for all scales. Hence, we could not confirm the validity of the BFI-10, BFI-44 and Hexad scales for adolescents.

Therefore, we subsequently performed an EFA to empirically examine the factor structures of the BFI-10, BFI-44 and Hexad data. Tables 6 to 8 revealed multiple issues for all scales: items majorly cross-loaded on several factors, loaded on factors diverging from the ones prescribed by the models, or had low communality. We attempted to optimise the models by iteratively pruning the most problematic items. However, even the most acceptable resulting models in Tables 9 and 10 were not fully satisfying due to low communalities, low average loadings per factor, and low total explained variance. Even though the pruned models slightly improved the fit



**Table 10: Rotated factor loadings and communalities based on the correlation matrix of the adapted Hexad data.**

Item	1 (Ac)	2 (PhS)	3 (P)	4 (D)	Com.
Ac2	<b>0.534</b>				0.37
Ac3	<b>0.737</b>				0.48
Ac4	<b>0.488</b>				0.36
S1		<b>0.624</b>			0.48
S2		<b>0.846</b>			0.64
S3		<b>0.520</b>			0.45
S4		<b>0.786</b>			0.60
Ph1		<b>0.613</b>			0.38
Ph2		<b>0.543</b>			0.34
Ph4		<b>0.580</b>			0.34
P2			<b>0.623</b>		0.42
P3			<b>0.526</b>		0.33
P4			<b>0.948</b>		0.74
D1				<b>0.518</b>	0.32
D3				<b>0.892</b>	0.70
D4				<b>0.670</b>	0.46

Note: |Loadings|  $\leq 0.30$  are suppressed and all absolute differences with the maximal loading (bold) are  $> 0.20$ . Communalities  $< 0.40$  are in red.

indices, we decided not to report these as they were still far from acceptable and such a confirmatory validation actually requires testing with a new sample [30].

Our findings suggest that HCI researchers should be cautious when including the BFI and Hexad scales in their current form as they may be unsuitable for adolescents. However, we acknowledge that we cannot blindly rely on the thresholds for fit indices in Table 3 as if they were golden rules for objectively deciding whether a model has good fit [58, 65]. Bearing in mind the complexities of research with adolescents, one could even argue that it is neither feasible nor fair to expect to meet the recommended cut-off values with data from a population as complex as adolescents. What then are acceptable fit indices and cut-off values deserves further debate.

## 5.2 Theoretical Model or Instrument?

Given their unique developmental stage, it was expected that adolescents would not fully comply with instruments for adults when examining personality and gamification user type. However, further investigation is required to assess whether the lack of valid measurements is to be attributed to issues with the underlying theoretical models (i.e. non-existence or partial absence of the proposed types in an adolescent population) or to the instruments themselves (i.e. error and bias in answering the items).

**5.2.1 Big Five Personality Traits and Adolescents.** The internal consistency scores suggested that all Big Five traits were measured reliably; only Agreeableness was borderline acceptable ( $\alpha = 0.68$ ). However, a CFA could not confirm good fit. The EFA for BFI-10 and BFI-44 revealed strong cross-loadings in the pattern matrix, particularly for Agreeableness. As this trait has also been found to be the most problematic for adults [81], our findings are in line with other researchers using BFI-10.

Yet, several studies before us have investigated and confirmed the existence of the five-factor nomology for adolescents [48] and even

younger children [26]. Moreover, personality scales specifically designed for adolescents do exist. For example, the scale in [63] has shown good psychometric qualities and may be considered by HCI researchers instead of BFI-10 or BFI-44. However, this scale contains 91 items and may not be feasible in situations where a parsimonious scale is needed. Surprisingly, if a short scale is needed, we found that BFI-10 outperformed BFI-44 in terms of goodness of fit. Still, given the problems observed in our analysis and in [81], caution remains needed when using this shortened version.

**5.2.2 Hexad Gamification User Types and Adolescents.** As Hexad is a relatively recent instrument, no prior studies with adolescents are known to us. Our internal consistency analysis revealed that four out of six factors could not be measured reliably and that Free Spirit was particularly problematic. The EFA revealed that all its items loaded higher on a different factor. During the iterative scale pruning, we could only achieve a somewhat acceptable factor structure by entirely removing Free Spirit, and by merging Philanthropist with Socialiser, which suggests a need to simplify the theoretical model for gamification user types for adolescents.

Interestingly, while validating their scale in English with adults, Tondello et al. [97] also found that Free Spirit has low internal consistency ( $\alpha = 0.628$ ), and that Socialiser and Philanthropist strongly correlate. Moreover, we found that the Disruptor type had the lowest overall score, which has been reported for adults too [e.g. 98]. In fact, scores on all types were similar to those reported for adults. Overall, our study found patterns in the gamification user types resembling those in [97, 98], which is somewhat surprising as literature on the characteristics of adolescence highlights a stronger need for autonomy, more sensation-seeking and risk-taking, and higher susceptibility to peers [5]. This suggests that we would find different or more articulated user types, especially for Disruptor, Free Spirit and Socialiser. Instead, it seems that our study stumbled upon issues similar to those unveiled by earlier studies, yet that these issues were more pronounced with adolescents and thus rendered problematic.

## 5.3 Limitations and Future Work

Our study investigated the BFI-10, BFI-44 and Hexad questionnaires with adolescents, revealing both similarities and differences with existing results for adults. However, our study comes with limitations: the lack of valid measurements may partly be attributed to difficulties with the instruments and the study context.

First, we relied on back-translation to verify the Dutch variant of Hexad. Even though three language experts independently confirmed the adequacy of our translation, we did not conduct a full validity study with adults. To remain as close as possible to the original Hexad questionnaire, we chose to not alter wordings that showed difficult in the pilot studies, particularly ‘to provoke’ and ‘status quo’. Instead, the online Hexad questionnaire provided tooltips to clarify this uncommon terminology. Unfortunately, we cannot verify whether participants used or understood them, as we did not track interaction with tooltips. Our findings underscore that further efforts are needed to reformulate items in language that is closer to the understanding of adolescents. Aligning ourselves with *TeenCI* [62], HCI for adolescents, we deem it imperative that this happens in close collaboration with adolescents [83]. Moreover, it

is warranted that future researchers first validate our Dutch version for Hexad with adults, and compare the results with validations in English and Spanish [95]. Future studies with adolescents may also further investigate how understanding and factor structure are impacted when replacing the Free Spirit and Achiever items with alternative items that improved Hexad's reliability in English in [95, Table 29]

Second, we acknowledge that contextual bias may have impacted our measurements. Even though adolescents filled out the Big Five and Hexad questionnaires *before* interacting with our mathematics platform, participants might still have framed them in a context of mathematical education. This may restrict the conclusions of our validation to this particular context. At the same time, our study meets the rising demand for testing Hexad in different contexts, which allows to better understand how gamification user types vary over different application domains, tasks, expertise, etc. [93] Moreover, we raise the question whether studies involving self-reported personality scales can avoid contextual bias altogether because they are highly susceptible to contextual cues [59]. This seems particularly pertinent for adolescents, who are typically peer-influenced and may be more self-conscious when answering personality-related questionnaires. Therefore, future studies with adolescents on gamification user types and personality traits need to further investigate the influence of context.

Finally, we investigated BFI-10 based on a subset of the BFI-44 data. Data directly stemming from the short survey might be more reliable as adolescents may be more considerate about their answers. Hence, our BFI-10 results should be handled with caution.

## 6 CONCLUSION

No longer children, not yet adults, adolescents are often overlooked in HCI studies in general [62, 80] and specifically in studies on tailoring gamification. Our results, based on Dutch translations of the BFI-10, BFI-44 and Hexad scales, revealed that the current forms of these scales are inadequate when used with adolescents. We suggested improvements both in terms of item legibility (instrument level) and factor structure (model level). In particular, we proposed a simplified model structure for Hexad that may better fit adolescent data. Future research could verify whether this novel theoretical model outperforms the current Hexad model when studying adolescents' gamification user types. An additional contribution to the HCI field is a unique data set of BFI-44 and Hexad questionnaire submissions in Dutch from 402 adolescents, publicly available in the spirit of open science and to allow further exploration. This data set can be a benchmark in future model fit tests, and the translation of the Hexad questionnaire could guide Dutch studies concerning the Hexad framework. In conclusion, our study aimed to advance theory on the Big Five and Hexad models; turning these models into reliable instruments for adolescents will help to better tailor gamified applications. Ultimately, with our validation study, we hope to contribute to the debate in HCI research on how to tailor interactive systems for the complex yet fascinating adolescent population.

## ACKNOWLEDGMENTS

We are very grateful to the mathematics teachers from Atheneum Unesco Koekelberg, Don Bosco Groot-Bijgaarden, VKO Opwijk,

DvM Humaniora Aalst, GO! Atheneum Asse Vijverbeek, and Sint-Jozefcollege Woluwe, who invited us into their classrooms; all parents who allowed their teenage children to participate in our study; and all adolescents who filled out the questionnaires. We also thank the English language experts for back-translating the Hexad questionnaire and comparing it to the original version. This work is part of the PANACEA Gaming Platform (HBC.2016.0177) and Personal Health Empowerment (HBC.2018.2012) research projects financed by Flanders Innovation & Entrepreneurship.

## REFERENCES

- [1] Aisha Muhammad Abdullahi, Kiemute Oyibo, Rita Orji, and Abdullahi Abubakar Kawu. 2019. The Influence of Age, Gender, and Cognitive Ability on the Susceptibility to Persuasive Strategies. *Information* 10, 11 (2019), 352.
- [2] Özcan Erkan Akgün and Murat Topal. 2018. Adaptation of the Gamification User Types Hexad Scale into Turkish. *International Journal* 5, 3 (2018), 389–402.
- [3] Maximilian Altmeyer, Pascal Lessel, Marc Schubhan, and Antonio Krüger. 2019. Towards Predicting Hexad User Types from Smartphone Data. In *Extended Abstracts of the Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts* (Barcelona, Spain) (CHI PLAY '19 Extended Abstracts). Association for Computing Machinery, New York, NY, USA, 315–322. <https://doi.org/10.1145/3341215.3356266>
- [4] Amy S. Beavers, John W. Lounsbury, Jennifer K. Richards, and Schuyler W. Huck. 2013. Practical considerations for using exploratory factor analysis in educational research. *Practical Assessment, Research, and Evaluation* 18, 1 (2013), 6.
- [5] Beth T. Bell. 2016. Understanding Adolescents. In *Perspectives on HCI Research with Teenagers*, Linda Little, Daniel Fitton, Beth T. Bell, and Nicola Toth (Eds.). Springer International Publishing, Cham, 11–27. [https://doi.org/10.1007/978-3-319-33450-9\\_2](https://doi.org/10.1007/978-3-319-33450-9_2)
- [6] Sheri A. Berenbaum, Adriene M. Beltz, and Robin Corley. 2015. The Importance of Puberty for Adolescent Development: Conceptualization and Measurement. In *Advances in Child Development and Behavior*. Vol. 48. Elsevier Academic Press inc., 525 B Street, Suite 1900, San Diego, CA 92101-4495 USA, 53–92. <https://doi.org/10.1016/bs.acdb.2014.11.002>
- [7] Max V. Birk, Maximilian A. Friehs, and Regan L. Mandryk. 2017. Age-Based Preferences and Player Experience: A Crowdsourced Cross-Sectional Study. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play* (Amsterdam, The Netherlands) (CHI PLAY '17). Association for Computing Machinery, New York, NY, USA, 157–170. <https://doi.org/10.1145/3116595.3116608>
- [8] Maura Bouça. 2012. Mobile Communication, Gamification and Ludification. In *Proceeding of the 16th International Academic MindTrek Conference* (Tampere, Finland) (MindTrek '12). Association for Computing Machinery, New York, NY, USA, 295–301. <https://doi.org/10.1145/2393132.2393197>
- [9] Timothy A. Brown. 2015. *Confirmatory factor analysis for applied research*. The Guilford Press, New York, NY.
- [10] Marc Busch, Elke Mattheiss, Rita Orji, Andrzej Marczewski, Wolfgang Hochleitner, Michael Lankes, Lennart E. Nacke, and Manfred Tscheligi. 2015. Personalization in Serious and Persuasive Games and Gamified Interactions. In *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play* (London, United Kingdom) (CHI PLAY '15). Association for Computing Machinery, New York, NY, USA, 811–816. <https://doi.org/10.1145/2793107.2810260>
- [11] James N. Butcher. 2010. Minnesota Multiphasic Personality Inventory. In *The Corsini Encyclopedia of Psychology*. John Wiley & Sons, Inc., Hoboken, New Jersey, 1–3. <https://doi.org/10.1002/9780470479216.corpsy0573>
- [12] Joseph A. Cafazzo, Mark Casselman, Nathaniel Hamming, Debra K. Katzman, and Mark R. Palmert. 2012. Design of an mHealth app for the self-management of adolescent type 1 diabetes: a pilot study. *Journal of medical Internet research* 14, 3 (2012), e70.
- [13] Stefano Carrino, Maurizio Caon, Leonardo Angelini, Elena Mugellini, Omar Abou Khaled, Silvia Orte, Eloisa Vargiu, Neil Coulson, Jose C. E. Serrano, Sarah Tabozzi, Claudio Lafortuna, and Giovanna Rizzo. 2014. PEGASO: A Personalised and Motivational ICT System to Empower Adolescents Towards Healthy Lifestyles. *Studies in Health Technology and Informatics* 207 (2014), 350–359. <https://doi.org/10.3233/978-1-61499-474-9-350>
- [14] Raymond B. Cattell. 1966. The scree test for the number of factors. *Multivariate behavioral research* 1, 2 (1966), 245–276.
- [15] Dennis Child. 1975. *The essentials of factor analysis*. Holt, Rinehart and Winston, London.
- [16] Chih-Ping Chou and Peter M. Bentler. 1995. Estimates and tests in structural equation modeling. In *Structural equation modeling: Concepts, issues, and applications*, Rick H. Hoyle (Ed.). Sage Publications, Inc, Thousand Oaks, CA, 37–55.

- [17] Paul Costa and Robert McCrae. 2008. The revised NEO personality inventory (NEO-PI-R). *The SAGE Handbook of Personality Theory and Assessment 2* (01 2008), 179–198. <https://doi.org/10.4135/9781849200479.n9>
- [18] Anna B. Costello and Jason Osborne. 2005. Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical assessment, research, and evaluation* 10, 1 (2005), 7.
- [19] Karen W. Cullen, Kathy Watson, Tom Baranowski, Janice H. Baranowski, and Issa Zakeri. 2005. Squire's Quest: intervention changes occurred at lunch and snack meals. *Appetite* 45, 2 (2005), 148–151.
- [20] Patrick J. Curran, Stephen G. West, and John F. Finch. 1996. The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological methods* 1, 1 (1996), 16–29.
- [21] Robin De Croon, Alemitu Mequanint Bezabih, Jonas Geuens, Davina Wildemeersch, Dries Oeyen, Katrien Verbert, and Vero Vanden Abeele. 2019. Motivational design techniques to increase adherence to a telemonitoring therapy a study with adolescent pectus patients. In *2019 IEEE International Conference on Healthcare Informatics (ICHI)*. IEEE, Xi'an, China, 1–12.
- [22] Edward L. Deci and Richard M. Ryan. 2012. *Self-determination theory*. Sage Publications Ltd, Thousand Oaks, CA, 416–436.
- [23] Jaap A. Denissen, Rinie Geenen, Marcel Aken, Samuel Gosling, and Jeff Potter. 2008. Development and Validation of a Dutch Translation of the Big Five Inventory (BFI). *Journal of personality assessment* 90 (4 2008), 152–157. <https://doi.org/10.1080/00223890701845229>
- [24] Sebastian Deterding, Dan Dixon, Rilla Khaled, and Lennart Nacke. 2011. From Game Design Elements to Gamefulness: Defining “Gamification”. In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments* (Tampere, Finland) (*MindTrek '11*). ACM, New York, NY, USA, 9–15. <https://doi.org/10.1145/2181037.2181040>
- [25] Darina Dicheva, Christo Dichev, Gennady Agre, and Galia Angelova. 2015. Gamification in Education: A Systematic Mapping Study. *Educational Technology & Society* 18, 3 (2015), 75–88.
- [26] John M Digman and Jillian Inouye. 1986. Further specification of the five robust factors of personality. *Journal of personality and social psychology* 50, 1 (1986), 116–123.
- [27] Christine DiStefano. 2002. The impact of categorization with confirmatory factor analysis. *Structural equation modeling* 9, 3 (2002), 327–346.
- [28] Christine DiStefano and Brian Hess. 2005. Using confirmatory factor analysis for construct validation: An empirical review. *Journal of Psychoeducational Assessment* 23, 3 (2005), 225–241.
- [29] Adrián Domínguez, Joseba Saenz de Navarrete, Luis de Marcos, Luis Fernández-Sanz, Carmen Pagés, and José-Javier Martínez-Herráiz. 2013. Gamifying learning experiences: Practical implications and outcomes. *Computers & Education* 63 (2013), 380 – 392. <https://doi.org/10.1016/j.compedu.2012.12.020>
- [30] Brian Everitt and Torsten Hothorn. 2011. *An introduction to applied multivariate analysis with R*. Springer Science & Business Media, New York, NY.
- [31] Leandre R. Fabrigar, Duane T. Wegener, Robert C. MacCallum, and Erin J. Strahan. 1999. Evaluating the use of exploratory factor analysis in psychological research. *Psychological methods* 4, 3 (1999), 272–299.
- [32] Daniel Fitton, Linda Little, and Beth T. Bell. 2016. Introduction: HCI Reaches Adolescence. In *Perspectives on HCI Research with Teenagers*, Linda Little, Daniel Fitton, Beth T. Bell, and Nicola Toth (Eds.). Springer International Publishing, Cham, 1–9.
- [33] Daniel Fitton, Janet C. Read, and Matthew Horton. 2013. The Challenge of Working with Teens as Participants in Interaction Design. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems* (Paris, France) (*CHI EA '13*). Association for Computing Machinery, New York, NY, USA, 205–210. <https://doi.org/10.1145/2468356.2468394>
- [34] Elia Gabarron, Thomas Schopf, J. Artur Serrano, Luis Fernández Luque, and Enrique Dorronzoro Zubiete. 2013. Gamification strategy on prevention of STDs for youth. In *Medinfo2013: The 14th World Congress on Medical and Health Informatics*. IMIA, Copenhagen, Denmark, 1066.
- [35] Lewis R. Goldberg. 1993. The structure of phenotypic personality traits. *American psychologist* 48, 1 (1993), 26–34.
- [36] Samuel D. Gosling, Peter J. Rentfrow, and William B. Swann Jr. 2003. A very brief measure of the Big-Five personality domains. *Journal of Research in personality* 37, 6 (2003), 504–528.
- [37] Anton Gustafsson, Magnus Bång, and Mattias Svahn. 2009. Power Explorer: A Casual Game Style for Encouraging Long Term Behavior Change among Teenagers. In *Proceedings of the International Conference on Advances in Computer Entertainment Technology* (Athens, Greece) (*ACE '09*). Association for Computing Machinery, New York, NY, USA, 182–189. <https://doi.org/10.1145/1690388.1690419>
- [38] Stuart Hallifax, Audrey Serna, Jean-Charles Marty, Guillaume Lavoué, and Elise Lavoué. 2019. Factors to Consider for Tailored Gamification. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play* (Barcelona, Spain) (*CHI PLAY '19*). ACM, New York, NY, USA, 559–572. <https://doi.org/10.1145/3311350.3347167>
- [39] Juho Hamari, Jonna Koivisto, and Harri Sarsa. 2014. Does Gamification Work? – A Literature Review of Empirical Studies on Gamification. In *2014 47th Hawaii International Conference on System Sciences*. IEEE, Waikoloa, HI, 3025–3034.
- [40] James C. Hayton, David G. Allen, and Vida Scarpello. 2004. Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis. *Organizational research methods* 7, 2 (2004), 191–205.
- [41] Richard M. Heiberger and Naomi B. Robbins. 2014. Design of diverging stacked bar charts for Likert scales and other applications. *Journal of Statistical Software* 57, 5 (2014), 1–32.
- [42] Daire Hooper, Joseph Coughlan, and Michael R. Mullen. 2008. Structural equation modelling: Guidelines for determining model fit. *Electronic journal of business research methods* 6, 1 (2008), 53–60.
- [43] John L. Horn. 1965. A rationale and test for the number of factors in factor analysis. *Psychometrika* 30, 2 (1965), 179–185.
- [44] Li-tze Hu and Peter M. Bentler. 1999. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal* 6, 1 (1999), 1–55.
- [45] Kai Huotari and Juho Hamari. 2017. A definition for gamification: anchoring gamification in the service marketing literature. *Electronic Markets* 27, 1 (2017), 21–31.
- [46] Yuan Jia, Bin Xu, Yamini Karanam, and Stephen Voids. 2016. Personality-targeted Gamification: A Survey Study on Personality Traits and Motivational Affordances. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) (*CHI '16*). ACM, New York, NY, USA, 2001–2013. <https://doi.org/10.1145/2858036.2858515>
- [47] Oliver P. John. 1990. The “Big Five” factor taxonomy: Dimensions of personality in the natural language and in questionnaires. In *Handbook of personality: Theory and research*, L.A. Pervin (Ed.). The Guilford Press, New York, NY, 66–100.
- [48] Oliver P. John, Avshalom Caspi, Richard W. Robins, Terrie E. Moffitt, and Magda Stouthamer-Loeber. 1994. The “little five”: Exploring the nomological network of the five-factor model of personality in adolescent boys. *Child development* 65, 1 (1994), 160–178.
- [49] Oliver P John, Laura P Naumann, and Christopher J Soto. 2008. Paradigm shift to the integrative big five trait taxonomy. *Handbook of personality: Theory and research* 3, 2 (2008), 114–158.
- [50] Oliver P. John and Sanjay Srivastava. 1999. The Big Five trait taxonomy: History, measurement, and theoretical perspectives. *Handbook of personality: Theory and research* 2, 1999 (1999), 102–138.
- [51] Daniel Johnson, Sebastian Deterding, Kerri-Ann Kuhn, Aleksandra Staneva, Stoyan Stoyanov, and Leanne Hides. 2016. Gamification for health and wellbeing: A systematic review of the literature. *Internet interventions* 6 (2016), 89–106.
- [52] Daniel Johnson, Ella Horton, Rory Mulcahy, and Marcus Foth. 2017. Gamification and serious games within the domain of domestic energy consumption: A systematic review. *Renewable and Sustainable Energy Reviews* 73 (2017), 249–264.
- [53] Henry F. Kaiser. 1960. The application of electronic computers to factor analysis. *Educational and psychological measurement* 20, 1 (1960), 141–151.
- [54] Henry F. Kaiser. 1974. An index of factorial simplicity. *Psychometrika* 39, 1 (1974), 31–36.
- [55] Maurits Kaptein, Boris De Ruyter, Panos Markopoulos, and Emile Aarts. 2012. Adaptive Persuasive Systems: A Study of Tailored Persuasive Text Messages to Reduce Snacking. *ACM Trans. Interact. Intell. Syst.* 2, 2, Article 10 (June 2012), 25 pages. <https://doi.org/10.1145/2209310.2209313>
- [56] Yamini Karanam, Leslie Filko, Lindsay Kaser, Hanan Alotaibi, Elham Makhsoom, and Stephen Voids. 2014. Motivational Affordances and Personality Types in Personal Informatics. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication* (Seattle, Washington) (*UbiComp '14 Adjunct*). ACM, New York, NY, USA, 79–82. <https://doi.org/10.1145/2638728.2638800>
- [57] Christine Keung, Alexa Lee, Shirley Lu, and Megan O’Keefe. 2013. BunnyBolt: A Mobile Fitness App for Youth. In *Proceedings of the 12th International Conference on Interaction Design and Children (IDC '13)*. Association for Computing Machinery, New York, NY, USA, 585–588. <https://doi.org/10.1145/2485760.2485871>
- [58] Rex B. Kline. 2015. *Principles and practice of structural equation modeling* (fourth ed.). The Guilford Press, New York, NY.
- [59] Barbara Krahé, Jana Becker, and Jana Zöllner. 2008. Contextual cues as a source of response bias in personality questionnaires: The case of the NEO-FFI. *European Journal of Personality: Published for the European Association of Personality Psychology* 22, 8 (2008), 655–673.
- [60] Brian D. Leany. 2013. Brain Development and Health Implications in Adolescents. In *Handbook of Adolescent Health Psychology*, William T. O’Donohue, Lorraine T. Benuto, and Lauren Woodward Tolle (Eds.). Springer New York, New York, NY, 235–244.
- [61] Annulla Linders. 2017. Deconstructing Adolescence. In *International Handbook on Adolescent Health and Development: The Public Health Response*, Andrew L. Cherry, Valentina Baltag, and Mary E. Dillon (Eds.). Springer International Publishing, Cham, 15–28.

- [62] Linda Little, Daniel Fitton, Beth T. Bell, and Nicola Toth. 2016. *Perspectives on HCI research with teenagers*. Springer International Publishing, Switzerland.
- [63] John W Lounsbury, Holly Tatum, Lucy W Gibson, Soo-Hee Park, Eric D Sundstrom, Frances L Hamrick, and Denise Wilburn. 2003. The development of a Big Five adolescent personality inventory. *Journal of Psychoeducational Assessment* 21, 2 (2003), 111–133.
- [64] Andrzej Marczewski. 2015. User types. *Even Ninja Monkeys Like to Play: Gamification, Game Thinking and Motivational Design 1* (2015), 65–80.
- [65] Herbert W. Marsh, Kit-Tai Hau, and Zhonglin Wen. 2004. In Search of Golden Rules: Comment on Hypothesis-Testing Approaches to Setting Cutoff Values for Fit Indexes and Dangers in Overgeneralizing Hu and Bentler's (1999) Findings. *Structural Equation Modeling: A Multidisciplinary Journal* 11, 3 (2004), 320–341. [https://doi.org/10.1207/s15328007sem1103\\_2](https://doi.org/10.1207/s15328007sem1103_2)
- [66] Anna Lisa Martin-Niedecken and Ulrich Götz. 2016. Design and Evaluation of a Dynamically Adaptive Fitness Game Environment for Children and Young Adolescents. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts* (Austin, Texas, USA) (*CHI PLAY Companion '16*). Association for Computing Machinery, New York, NY, 205–212. <https://doi.org/10.1145/2968120.2987720>
- [67] Masaki Matsunaga. 2010. How to Factor-Analyze Your Data Right: Do's Don'ts, and How-To's. *International Journal of Psychological Research* 3, 1 (2010), 97–110.
- [68] Rudy McDaniel, Robb Lindgren, and Jon Friskics. 2012. Using badges for shaping interactions in online learning environments. In *2012 IEEE international professional communication conference*. IEEE, Orlando, FL, USA, 1–4.
- [69] Alberto Mora, Gustavo F. Tondello, Laura Calvet, Carina González, Joan Arnedo-Moreno, and Lennart E. Nacke. 2019. The Quest for a Better Tailoring of Gameful Design: An Analysis of Player Type Preferences. In *Proceedings of the XX International Conference on Human Computer Interaction* (Donostia, Gipuzkoa, Spain) (*Interacción '19*). Association for Computing Machinery, New York, NY, USA, 1–8. <https://doi.org/10.1145/3335595.3335625>
- [70] Bengt Muthén and David Kaplan. 1985. A comparison of some methodologies for the factor analysis of non-normal Likert variables. *Brit. J. Math. Statist. Psych.* 38, 2 (1985), 171–189.
- [71] Lennart E. Nacke, Chris Bateman, and Regan L. Mandryk. 2011. BrainHex: Preliminary Results from a Neurobiological Gamer Typology Survey. In *Entertainment Computing – ICEC 2011*, Junia Coutinho Anacleto, Sidney Fels, Nicholas Graham, Bill Kapralos, Magy Saif El-Nasr, and Kevin Stanley (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 288–293.
- [72] Lennart E. Nacke, Chris Bateman, and Regan L. Mandryk. 2014. BrainHex: A neurobiological gamer typology survey. *Entertainment computing* 5, 1 (2014), 55–62.
- [73] Lennart E. Nacke and Sebastian Deterding. 2017. The maturing of gamification research. *Computers in Human Behavior* 71, June (2017), 450–454. <https://doi.org/10.1016/j.chb.2016.11.062>
- [74] Oded Nov and Ofer Arazy. 2013. Personality-targeted Design: Theory, Experimental Procedure, and Preliminary Results. In *Proceedings of the 2013 Conference on Computer Supported Cooperative Work* (San Antonio, Texas, USA) (*CSCW '13*). ACM, New York, NY, USA, 977–984. <https://doi.org/10.1145/2441776.2441887>
- [75] William T. O'Donohue, Lorraine T. Benuto, and Lauren Woodward Tolle. 2013. *Handbook of adolescent health psychology*. Springer Science & Business Media, New York, NY.
- [76] Rita Orji, Lennart E. Nacke, and Chrysanne Di Marco. 2017. Towards Personality-driven Persuasive Health Games and Gamified Systems. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) (*CHI '17*). ACM, New York, NY, USA, 1015–1027. <https://doi.org/10.1145/3025453.3025577>
- [77] Rita Orji, Gustavo F. Tondello, and Lennart E. Nacke. 2018. Personalizing Persuasive Strategies in Gameful Systems to Gamification User Types. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) (*CHI '18*). ACM, New York, NY, USA, Article 435, 14 pages. <https://doi.org/10.1145/3173574.3174009>
- [78] Jason W. Osborne. 2015. What is rotating in exploratory factor analysis? *Practical Assessment, Research, and Evaluation* 20, 1 (2015), 2.
- [79] A.C. Petersen. 1988. Adolescent Development. *Annual Review of Psychology* 39, 1 (1988), 583–607. <https://doi.org/10.1146/annurev.ps.39.020188.003055>
- [80] Erika S. Poole and Tamara Peyton. 2013. Interaction Design Research with Adolescents: Methodological Challenges and Best Practices. In *Proceedings of the 12th International Conference on Interaction Design and Children* (New York, New York, USA) (*IDC '13*). Association for Computing Machinery, New York, NY, USA, 211–217. <https://doi.org/10.1145/2485760.2485766>
- [81] Beatrice Rammstedt and Oliver P. John. 2007. Measuring Personality in One Minute or Less: A 10-Item Short Version of the Big Five Inventory in English and German. *Journal of Research in Personality* 41 (02 2007), 203–212. <https://doi.org/10.1016/j.jrp.2006.02.001>
- [82] Amon Rapp, Frank Hopfgartner, Juho Hamari, Conor Linehan, and Federica Cena. 2019. Strengthening gamification studies: Current trends and future opportunities of gamification research.
- [83] Janet C. Read and Matthew Horton. 2016. Future Directions for Quality TeenCI Research. In *Perspectives on HCI Research with Teenagers*, Linda Little, Daniel Fitton, Beth T. Bell, and Nicola Toth (Eds.). Springer International Publishing, Cham, 269–278. [https://doi.org/10.1007/978-3-319-33450-9\\_11](https://doi.org/10.1007/978-3-319-33450-9_11)
- [84] Anna Rohde, Anja Duensing, Christine Dawczynski, Jasmin Godemann, Stefan Lorkowski, and Christine Brombach. 2019. An App to Improve Eating Habits of Adolescents and Young Adults (Challenge to Go): Systematic Development of a Theory-Based and Target Group-Adapted Mobile App Intervention. *JMIR mHealth and uHealth* 7, 8 (Aug. 2019), e11575. <https://doi.org/10.2196/11575>
- [85] Lamyae Sardi, Ali Idri, and José Luis Fernández-Alemán. 2017. A systematic review of gamification in e-Health. *Journal of biomedical informatics* 71 (2017), 31–48.
- [86] Neal Schmitt. 1996. Uses and abuses of coefficient alpha. *Psychological assessment* 8, 4 (1996), 350–353.
- [87] Thomas A. Schmitt. 2011. Current methodological considerations in exploratory and confirmatory factor analysis. *Journal of Psychoeducational Assessment* 29, 4 (2011), 304–321.
- [88] Katie Seaborn and Deborah I Fels. 2015. Gamification in theory and action: A survey. *International Journal of human-computer studies* 74 (2015), 14–31.
- [89] Christopher J Soto, Oliver P. John, Samuel D. Gosling, and Jeff Potter. 2011. Age differences in personality traits from 10 to 65: Big Five domains and facets in a large cross-sectional sample. *Journal of personality and social psychology* 100, 2 (2011), 330–348.
- [90] Pedro Sousa, Emilia Duarte, Regina Ferreira, Ana Esperanca, Roberta Frontini, Rita Santos-Rocha, Luis Luis, Sara S. Dias, and Nuno Marques. 2019. An mHealth intervention programme to promote healthy behaviours and prevent adolescent obesity (TeenPower): A study protocol. *JOURNAL OF ADVANCED NURSING* 75, 3 (March 2019), 683–691. <https://doi.org/10.1111/jan.13905>
- [91] Jennifer N. Stinson, Lindsay A. Jibb, Cynthia Nguyen, Paul C. Nathan, Anne Marie Maloney, L. Lee Dupuis, J. Ted Gerstle, Benjamin Alman, Sevan Hopyan, Caron Strahlendorf, et al. 2013. Development and testing of a multi-dimensional iPhone pain assessment application for adolescents with cancer. *Journal of medical Internet research* 15, 3 (2013), e51.
- [92] Sujit Subhash and Elizabeth A Cudney. 2018. Gamified learning in higher education: A systematic review of the literature. *Computers in Human Behavior* 87 (2018), 192–206.
- [93] Gustavo Fortes Tondello. 2019. *Dynamic Personalization of Gameful Interactive Systems*. Ph.D. Dissertation. University of Waterloo, Waterloo, Ontario, Canada.
- [94] Gustavo F. Tondello, Karina Arrambide, Giovanni Ribeiro, Andrew Jian-lan Cen, and Lennart E. Nacke. 2019. "I Don't Fit into a Single Type": A Trait Model and Scale of Game Playing Preferences. In *Human-Computer Interaction – INTERACT 2019*, David Lamas, Fernando Loizides, Lennart Nacke, Helen Petrie, Marco Winckler, and Panayiotis Zaphiris (Eds.). Springer International Publishing, Cham, 375–395.
- [95] Gustavo F. Tondello, Alberto Mora, Andrzej Marczewski, and Lennart E. Nacke. 2019. Empirical validation of the Gamification User Types Hexad scale in English and Spanish. *International Journal of Human-Computer Studies* 127 (2019), 95–111. <https://doi.org/10.1016/j.ijhcs.2018.10.002>
- [96] Gustavo F. Tondello, Alberto Mora, and Lennart E. Nacke. 2017. Elements of Gameful Design Emerging from User Preferences. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play* (Amsterdam, The Netherlands) (*CHI PLAY '17*). ACM, New York, NY, USA, 129–142. <https://doi.org/10.1145/3116595.3116627>
- [97] Gustavo Fortes Tondello and Lennart E. Nacke. 2018. Towards Customizing Gameful Systems by Gameful Design Elements. In *Proceedings of the Personalization in Persuasive Technology Workshop*. CEUR-WS, Waterloo, Canada, 101–109.
- [98] Gustavo F. Tondello, Rina R. Wehbe, Lisa Diamond, Marc Busch, Andrzej Marczewski, and Lennart E. Nacke. 2016. The Gamification User Types Hexad Scale. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play* (Austin, Texas, USA) (*CHI PLAY '16*). ACM, New York, NY, USA, 229–243. <https://doi.org/10.1145/2967934.2968082>
- [99] Stephen G. West, John F. Finch, and Patrick J. Curran. 1995. Structural equation models with nonnormal variables: Problems and remedies. In *Structural equation modeling: Concepts, issues, and applications*, Rick H. Hoyle (Ed.). Sage Publications, Inc, Thousand Oaks, CA, 56–75.
- [100] Daniel Zelterman. 2015. *Applied multivariate statistics with R*. Springer International Publishing, Switzerland.

**Table 11: Dutch translation of the Hexad questionnaire [98], found adequate through a back-translation procedure. The questionnaire is a 7-point Likert scale (Completely disagree, Disagree, Rather disagree, Neutral, Rather agree, Agree, Completely agree).**

No.	English original	Dutch translation
Ph1	It makes me happy if I am able to help others.	Het maakt me gelukkig als ik anderen kan helpen.
Ph2	I like helping others to orient themselves in new situations.	Ik hou ervan anderen te helpen met zich oriënteren in nieuwe situaties.
Ph3	I like sharing my knowledge.	Ik hou ervan om mijn kennis te delen.
Ph4	The wellbeing of others is important to me.	Ik vind het welzijn van anderen belangrijk.
S1	Interacting with others is important to me.	Ik vind interactie met anderen belangrijk.
S2	I like being part of a team.	Ik vind het leuk om deel te zijn van een team.
S3	It is important to me to feel like I am part of a community.	Ik vind het belangrijk om me een deel van een gemeenschap te voelen.
S4	I enjoy group activities.	Ik geniet van groepsactiviteiten.
F1	It is important to me to follow my own path.	Ik vind het belangrijk om mijn eigen weg te volgen.
F2	I often let my curiosity guide me.	Ik laat me vaak leiden door mijn nieuwsgierigheid.
F3	I like to try new things.	Ik vind het leuk om nieuwe dingen uit te proberen.
F4	Being independent is important to me.	Ik vind onafhankelijk zijn belangrijk.
Ac1	I like defeating obstacles.	Ik vind het leuk om obstakels te overwinnen.
Ac2	It is important to me to always carry out my tasks completely.	Ik vind het belangrijk om taken altijd volledig uit te voeren.
Ac3	It is difficult for me to let go of a problem before I have found a solution.	Ik vind het moeilijk om een probleem los te laten voordat ik een oplossing heb gevonden.
Ac4	I like mastering difficult tasks.	Ik vind het leuk om moeilijke taken de baas te kunnen.
P1	I like competitions where a prize can be won.	Ik hou van competities waarbij een prijs gewonnen kan worden.
P2	Rewards are a great way to motivate me.	Beloningen zijn een goede manier om me te motiveren.
P3	Return of investment is important to me.	Ik vind het belangrijk om iets terug te krijgen voor mijn inspanningen.
P4	If the reward is sufficient I will put in the effort.	Als de beloning voldoende is, dan zal ik de inspanningen leveren.
D1	I like to provoke.	Ik vind het leuk om te provoceren.
D2	I like to question the status quo.	Ik vind het leuk om de status quo in vraag te stellen.
D3	I see myself as a rebel.	Ik zie mezelf als een rebel.
D4	I dislike following rules.	Ik hou niet van het volgen van regels.

**Table 12: The BFI-44 questionnaire, which is a 5-point Likert scale (Completely disagree, Disagree, Agree nor disagree, Agree, Completely agree). Items whose number ends with an “r” are reverse-scored. The BFI-10 item numbers are in bold.**

No.	English original	No.	English original
O1	Is original, comes up with new ideas	E4	Generates a lot of enthusiasm
O2	Is curious about many different things	E5r	Tends to be quiet
O3	Is ingenious, a deep thinker	E6	Has an assertive personality
<b>O4</b>	Has an active imagination	E7r	Is sometimes shy, inhibited
O5	Is inventive	<b>E8</b>	Is outgoing, sociable
O6	Values artistic, aesthetic experiences	<b>A1r</b>	Tends to find fault with others
O7r	Prefers work that is routine	A2	Is helpful and unselfish with others
O8	Likes to reflect, play with ideas	A3r	Starts quarrels with others
<b>O9r</b>	Has few artistic interests	A4	Has a forgiving nature
O10	Is sophisticated in art, music, or literature	<b>A5</b>	Is generally trusting
<b>C1</b>	Does a thorough job	A6r	Can be cold and aloof
C2r	Can be somewhat careless	A7	Is considerate and kind to almost everyone
C3	Is a reliable worker	A8r	Is sometimes rude to others
C4r	Tends to be disorganized	A9	Likes to cooperate with others
<b>C5r</b>	Tends to be lazy	N1	Is depressed, blue
C6	Perseveres until the task is finished	<b>N2r</b>	Is relaxed, handles stress well
C7	Does things efficiently	N3	Can be tense
C8	Makes plans and follows through with them	N4	Worries a lot
C9r	Is easily distracted	N5r	Is emotionally stable, not easily upset
E1	Is talkative	N6	Can be moody
<b>E2r</b>	Is reserved	N7r	Remains calm in tense situations
E3	Is full of energy	<b>N8</b>	Gets nervous easily