The role of gameful perception as a mediator for intrinsically motivating gamification

Henrik Wesseloh[a], Niklas Buddensiek[a], Tim Pantel[a], Felix M. Stein[a], Phillip Szelat[a] and Matthias Schumann[a]

[a] *University of Göttingen, Platz der Göttinger Sieben 5, Göttingen, 37073, Germany*

Abstract

Scientific studies have shown that the use of gamification is effective in promoting motivation and behavior. However, reasons for the effectiveness of the implemented game elements are analyzed insufficiently. The following article therefor addresses the issue of gameful perception induced through gamification by using an experimental, single factor between-subjects design. First, we draw on *Self-Determination Theory* to derive a grounded research model. Then, we investigate two groups by using a learning application with a gamified and non-gamified version to test our hypotheses. Our results show that the users' gameful perception of an activity mediates the motivational outcomes. We discuss our findings and relate these to the app's design to outline pitfalls and the importance of evoking a conscious gameful experience.

Keywords

Gamification, Intrinsic Motivation, SDT, Gameful Perception, SEM

1. Introduction

As an important instrument for fostering or maintaining human motivation and behavior, gamification, *i.e.* "the use of game design elements in a non-game context" [6] to evoke "gameful experiences" [11], has gained more attention in recent years. Gamification is used across different domains, *e.g.* education, health, or crowdsourcing, to support psychological and behavioral outcomes, like *e.g.* engagement or participation [13]. Previous research has shown that gamification can also help to satisfy intrinsic psychological needs like autonomy, competence or relatedness and thus support intrinsic motivation [4, 15, 17, 23].

However, the knowledge about *how* exactly gamification works is still incomplete [19]. Systematic reviews on gamification research mostly report mixed results for the approach's effectiveness [13]. Even though the number of studies analyzing gamifications effectiveness is growing, academia still misses insights: It is yet unknown to what extent certain factors besides gameful affordances influence the motivational impact. These aspects are however needed to explain the different outcomes among similar game design elements [13, 18]. For this study we therefor analyze whether the motivational effect of gamification depends on a gameful perception. We assume that in order to benefit from gamification, it is mandatory that users also perceive the activity as gameful. Thus, we address the following research question:

RQ: *How does a gameful perception influence the motivational effect of gamification?*

To address this question, we briefly want to give an overview of related research and derive a research model based on *Self-Determination Theory* (SDT) [4] as theoretical foundation. We then present our single factor between-subjects research design by describing the experiment procedure, the measures for data collection and provide a descriptive analysis of participants.

5th International GamiFIN Conference 2021 (GamiFIN 2021), April 7-10, 2021, Finland.
EMAIL: henrik.wesseloh@uni-goettingen.de

To validate the mediation model, we follow a structural equation modeling approach [2, 12]. Moreover, we do a mean comparison, to find differences in motivational effect of a gamified and a non-gamified application. Finally, we discuss our findings and conclude this paper to provide implications for future research.

Our findings contribute to current research analyzing the effect of gamification on intrinsic motivation and need satisfaction, because they provide an understanding of why gamification studies might fail measuring positive outcomes [13]: When analyzing the effect of (individual) game elements on psychological outcomes, we often miss on the actual emersion of a conscious gameful perception within the participants.

2. Theoretical foundation

2.1. Related research

Recent empirical studies on the effects of gamification on intrinsic motivation are mostly based on SDT and its subtheories and therefore also examine the basic needs for autonomy, competence or relatedness [cf. 15, 17, 20, 23].

Autonomy refers to a psychological freedom to perform an activity at own will, while *competence* describes the effectiveness of an individual doing this activity. *Relatedness* covers the social aspect of an activity. In combination, these factors build the concept of *intrinsic motivation,* which refers to an inherent interest or enjoyment towards an activity itself and not for its instrumental value [3, 4].

While an older study from Hanus & Fox [8] showed that gamification had a negative effect on intrinsic motivation recent studies presented positive effects. We outline a few contributions which stressed the importance of gamification's perception for intrinsic need satisfaction in a chronological order.

Lieberoth (2015) hypothesized that framing an activity as a game can lead to behavioral changes and subjective experiences of intrinsic motivation [14]. The study showed that the perception of a playful frame for an mundane task has a significant impact on enjoyment.

Deterding (2015), suggested gamification to support experiences of competence or inhibit autonomy, depending on the user's perception. He draws on SDT to argue that, if gamification is perceived by users as controlling, this would have a negative influence on the experience of autonomy. If the gameful feedback in turn is perceived as informative, it would promote the experience of competence [4, 5].

Mekler et al. (2017) analyzed motivational effects of individual game elements and found that, despite their supposed external incentive, *points*, *level* and the *leaderboard* did not lead to a negative (overjustification) effect on intrinsic motivation [15]. They attributed this result to a rather low gameful experience during the task.

Sailer et al. (2017) did a comparable study and found positive effects of gamification on competence and relatedness. Furthermore, they noticed that some participants did not recognize the game elements. The authors attributed this result to potential flaws in their gamification design and assumed that it had resulted in a lower motivational effect [17].

Furthermore, the contributions from Suh et al. (2018) and Xi & Hamari (2019) both provide positive empirical insights on how gamification satisfies the three intrinsic needs [20, 23]. These authors point out that the need for further research on potential moderators or mediators remains.

2.2. Research model

This section depicts the research model, which we briefly outline in the following.

For our model, we consider gamification as a manifest, exogenous variable that represents the source of a gameful perception. To simplify, gamification is modelled as binary variable, so an application or service will only vary between the gamified (1) and the non-gamified (0) state.

Perception of gamefulness describes to what extent individuals consciously notice a game-like situation when interacting with a system or service. This perception is supported *e.g.* by feelings of challenge, competition, immersion and social experiences, as well as playfulness [9]. As gameful designers, we can address these different experiences when implementing game mechanics in a utilitarian system or service [10]. Therefore, we state our first hypothesis:

H_1: *Gamification has a positive impact on the perception of gamefulness.*

Furthermore, Deci & Ryan's SDT (precisely Cognitive Evaluation Theory, CET) postulates that psychological need satisfaction consists of the fulfillment of autonomy, competence and relatedness [4]. As an individual's environment influences the satisfaction of intrinsic needs [4], gamification can be used to create beneficial conditions that support feelings of autonomy, competence and relatedness [5, 15, 17, 20, 23]. For example, *points*, *badges,* or *rankings* are common feedback elements and should thus support the need of competence [4]. These elements, however, do not have to inevitably originate from games and thus might not be perceived per se as gameful affordance. We take this instant as a reason why gameful approaches led to different results despite using comparable game elements within similar cases [13, 18]. So, to benefit from the "motivational pull" [16] of games, we assume that users need to perceive the use of a gamified application as gameful. Consequently, we state our second research hypothesis:

H2a-c: *Perception of gamefulness mediates gamification's effect on needs of autonomy (a), competence (b) and relatedness (c).*

Moreover, we want to address the support of intrinsic motivation as a desirable motivational effect. In theory, the three psychological needs of autonomy, competence and relatedness are required determinants for intrinsic motivation [4]. To assess the user's intrinsic motivation, SDT suggests to measure enjoyment, as they are closely associated [3]. Recent studies on gamification have used this operationalization to prove its motivational effects and concluded that supporting need satisfaction may improve enjoyment and, thus, intrinsic motivation [20]. For that reason, we hypothesize:

H3a-c: *Satisfying the need of autonomy (a), competence (b), and relatedness (c), has a positive effect on intrinsic motivation.*

Figure 1 visualizes our derived research model with the different hypotheses to test.

3. Study design

We used a single factor between-subjects study design to answer the research question and test our model hypotheses. Hence, the participants were divided into two groups: *experimental* and *control*. Both groups performed the same experiment, in which they answered general knowledge questions with our prototypical application. We treated the experimental group with a gamified version, while the control group used a non-gamified version. After the experiment, both groups were asked to fill out an online survey to determine their experiences. These survey results were then analyzed using statistical methods.

3.1. Experiment

The laboratory experiment is based on the use of a mobile web application for knowledge retrieval that uses collaborative and competitive game design elements to motivate learning. We developed this learning app for lectures [22], so mainly groups of 10 or more students use it. During a session, participants answer randomly assigned, multiple-choice questions prepared by lecturer on their mobile device, in their own pace, with a maximum duration of 30 seconds per question. These question sessions are set to last between three to five minutes. After a session, the participants receive a summary of their performance and solutions for their individual questions, while lecturers get access to aggregated performance data.

The gamified version additionally includes different *game design elements* to provide a gameful experience for its users and thus support the feeling of autonomy, competence, and relatedness. We integrated feedback elements like, *points* (for correct answers incl. streaks), *badges*, and a *leaderboard* to support

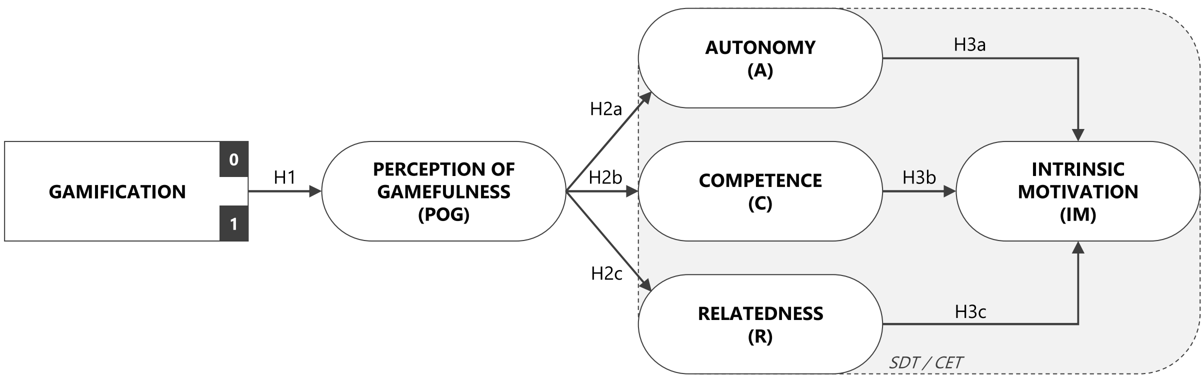

Figure 1: Research model

experiences of achievement and competition. To support immersion, users find a *narrative* setting and can choose their respective gameful *avatar*. Furthermore, the session resembles a challenging *quest*, in which the students *team* up to encounter a virtual *boss enemy*, which resembles the question pool. During this boss battle, each student contributes individually to overcome the boss by answering questions correctly but can be eliminated after a defined number of mistakes (*virtual lives*). The students win the boss event collaboratively by solving a defined number of questions during the set time. Overall, we assume, these achievement-related, immersive and social features are able to address the users' intrinsic need satisfaction in the educational context [23].

In contrast, the non-gamified version only shows a timer to indicate the remaining time during a question session. All gamification features like the narrative or the feedback elements are omitted. Figure 2 shows the visual differences between the gamified and non-gamified application. The screenshots display the quiz question design on a mobile device (*e.g.*, smartphone) [#1,2], the lecturer screen during a question session, which is presented to the participants via projector [#3,4], the start screen for participants [#5,6], the gameful feedback for participants [#7] and the gameful post-game screen with the leaderboard [#8].

Our plan was to use the application in two distinct lectures with live attendees, to set up the two groups for the experiment. Due to the ongoing corona pandemic, however, no live lectures were held, so we decided to adapt the experiment to take place online virtually. To simulate other participants during the question session, we integrated bots in our application. This way, we were able to allow individuals or smaller groups to participate in the online experiment and still maintain the social aspect of our gamification concept. Additionally, we setup a question pool of 45 general knowledge questions in the field of sports, politics, history, and science to reach out for a broader audience.

We shared our invitation to participate in the experiment via social media. The individuals and small groups that joined our experiment via a video conference platform that allows screen sharing (of the lecturer screen) were randomly assigned to the experimental or control group. Before each question session, we informed the participants about the procedure of the

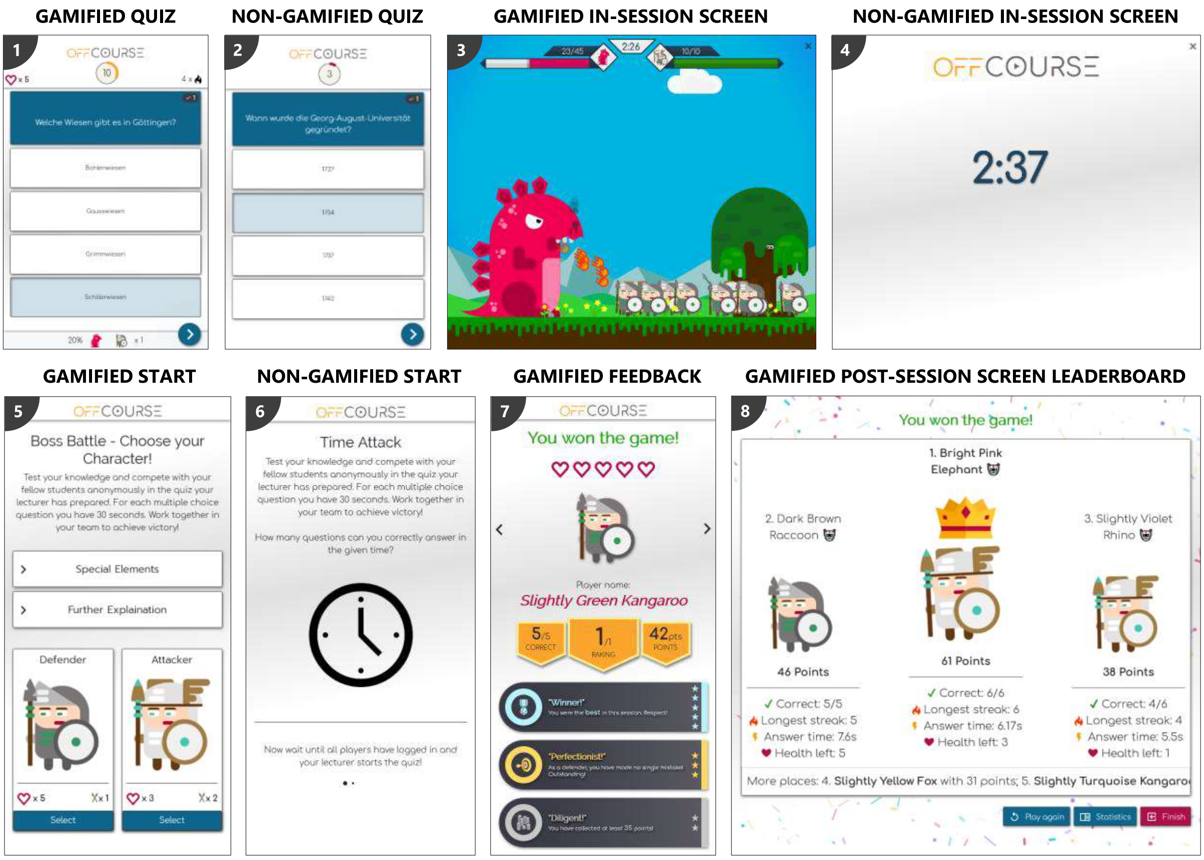

Figure 2: Comparison of the gamified and non-gamified application versions

experiment and briefly presented the respective application. Then, the question session started with a duration of 4 minutes (experimental group) or 3 minutes (control group)[2].

After participants had seen their individual (gameful) results, the app provided a button linked to the online survey. We used this link to add a parameter of each participant's ID to be able to connect the self-reported survey data with actual performance and session data provided be the app.

3.2. Measurement

The data collection following the experiment was carried out using an online survey. All items were measured with a seven-point Likert scale (completely disagree (1) to completely agree (7)). We used five items per latent construct to ensure an appropriate length of the questionnaire but also to leave additional room for the removal of weak or unsuitable items. As all items are self-reported experience measures, we carefully translated the items to German for participants to comprehend.

To measure the perception of gamefulness, we selected items from the GAMEFULQUEST scale for '*playfulness*' [9]. We did not consider other game-specific items to prevent confusion in the control group. To operationalize the SDT perspective of intrinsic motivation, we chose items from the Intrinsic Motivation Inventory (IMI) and respective contributions [1, 3, 21]. For the construct of autonomy the Ubisoft Perceived Experience Questionnaire (UPEQ) was used, which adapts the autonomy measures to video games [1]. The construct of relatedness was measured using an adapted questionnaire of the IMI's relatedness scale for group work [21]. The third SDT construct, competence, was measured using the '*perceived competence*' scale from the IMI [3]. Additionally, we used the IMI's scale for '*enjoyment / interest*' to operationalize intrinsic motivation [3].

Table 1
Measures for Latent Constructs

Construct	Item	Source
Perception of Gamefulness (POG)	The activity… POG1: …gives me an overall playful experience POG2: …taps into my imagination. *POG3: …feels like a mystery to reveal.* POG4: …makes me feel like I discover new things. POG5: …appeals to my curiosity.	**GAMEFUL-QUEST** [9]
Autonomy (A)	A1: I was free to decide how I wanted to do the activity. A2: I could approach the activity in my own way. *A3: I could make important decisions during the activity.* *A4: The choices I made influenced what happened.* *A5: My actions had an impact on the activity.*	**UPEQ** [1]
Competence (C)	C1: I think I am pretty good at this activity. C2: I think I did pretty well at this activity, compared to other students. C3: After doing this activity for a while, I felt pretty competent. C4: I am satisfied with my performance at this activity. *C5: This was an activity that I could not do very well. (R)*	**IMI** [3]
Relatedness (R)	*R1: I felt really distant to other participants. (R)* *R2: I feel close to other participants.* R3: Doing this activity, I feel I can learn with other participants. R4: With this activity, I feel I can relate with other participants. R5: Participating in this activity is important to make me feel that I belong to a community.	adapted **IMI** [21]
Intrinsic Motivation (IM)	IM1: I enjoyed doing this activity very much. IM2: This activity was fun to do. *IM3: I thought this was a boring activity. (R)* *IM4 This activity did not hold my attention at all. (R)* IM5: I would describe this activity as very interesting.	**IMI** [3]

Note. *(R): Reversed item; IMI: Intrinsic Motivation Inventory; UPEQ: Ubisoft Perceived Experience Questionnaire; Gray italic items were removed during factor analysis.*

[2] We chose this distinction, as the gamified session can end early due to its winning condition. In average, this resulted in the same playing time.

In addition to these measures, we asked six control questions to differentiate between the respondents. These controls include gender, age and highest educational level. Furthermore, the general affinity towards play was queried, with users specifying how often they play games during the week and whether participants had used our application before.

Due to cross-loadings (*cf.* 4.1 Measurement Model) and/or poor factor loadings (< 0.70) we had to remove certain items during data analysis to ensure high construct reliability. Cronbach's alpha confirms internal consistency for each construct ($\alpha \geq 0.828$).

Table 1 shows the items used as reflective measures for the latent constructs (*italic items* had to be removed during factor analysis).

3.3. Participants

Overall, the distribution of participants between the experimental and comparison group was balanced with 105 and 93 respondents. The experiment was taken by 43 individuals, 23 groups of two, 15 groups of three and 12 groups with four or more participants. The largest group in our sample held twelve participants.

The age of the participants was highly concentrated in the range of 20 to 30 years, the average age was 30.6 years (SD = 13.57), the median was 25 years. Whereas 134 participants were between the age of 20 and 30 years, only 48 participants were older than 30. Males were slightly overrepresented (60%) compared with females (39%), while two participants replied their gender as diverse (1%).

The participants mainly have an academic background: 43% had a bachelor's degree, 22% a master's or diploma. 21% of the respondents indicated that their highest educational level was the high school graduation. The most common field of study was *'economics'* (51%), while other fields of study were mainly in *'humanities'* and *'social sciences'* as well as *'teacher training'* and *'computer science'*. During the question sessions, the participants answered 12.8 (SD = 4.81) questions after 14.6 (SD = 5.39) seconds on average. The relative success rate, *i.e.*, the ratio of correct responses compared to the total number of answered questions, was 54%.

About 22% of the participants had a high affinity towards play and stated that they played games such as board games or video games *'every day'*, whereas 19% of participants said they played *'several times a week'* and 26% *'several times a month'*. The largest proportion of respondents (33%) said they played games *'less often'* or *'never'*. Some participants (34%) also had used our application before, but for the majority (64%) the app was new.

4. Results from data analysis

To test the postulated hypotheses from our research model, we used a structural equation modelling (SEM) approach [2, 12]. The data analysis was done in *Stata 16* and *SPSS 26*. To assess the quality of the measurement model, we performed an exploratory factor analysis in *SPSS*. Furthermore, we compared the survey results between the two groups with a mean comparison. Subsequently, the structural model was estimated using the maximum likelihood estimation (MLE) as the default method [12] in *Stata*. Accordingly, the SEM requirements of a large sample size, multivariate normality and correct model specification [12] were met, as the next section shows.

4.1. Measurement model

The SEM approach requires, the constructs to be tested for their *content*, *convergent*, and *discriminant* validity in order to assess the fit of the model with the collected data [2].

First, we assume content validity, because of our item selection from validated, existing scales like IMI [3] and GAMEFULQUEST [9].

Second, we tested empirically our sample for convergent validity by determining the individual item reliability, composite reliability (CR) and average variance extracted (AVE) by executing a factor analysis. The sample data was suitable for factor analysis, as the Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO) is greater than 0.60 (KMO = 0.834). In the case of individual item reliability, we follow the criteria of high factor loadings (> 0.70) [12], which is why a few items (POG3; A3; A4; A5; C5; IM3; IM4) had to be removed from the initial model. Afterwards we determined values for CR that ranged from 0.764 to 0.932, which is above the acceptable limit (> 0.70) [12]. Additionally, we calculated the AVE for each construct, which also exceeded the required threshold (> 0.50) in all cases, except for the perception of gamefulness (AVE = 0.462), which is slightly below the limit. Still, this constructs convergent validity can be regarded adequate [7], due to its higher composite reliability (CR = 0.764) .

Third, we determined discriminant validity by checking for possible cross-loadings via the pattern matrix. For the adjusted model, all items showed the highest loadings for their respective construct. Additionally, we draw upon the criterion from Fornell & Larcker [7] to show that the AVE of each construct is higher than any squared correlation with another construct, which serves as a prove of discriminant validity at construct level.

Table 2 and Table 3 present the results from our factor analysis, showing the relevant scores for CR, AVE, Cronbach's Alpha (α), as well as cross-loadings and inter-construct correlations for the Fornell-Larcker-criterion [7].

Table 2

Composite reliability (CR), average variance extracted (AVE) and inter-construct correlations

Construct	CR	AVE	POG	A	C	R	IM
Perception of Gamefulness (POG)	.764	.462	**.680**				
Autonomy (A)	.932	.873	.411	**.935**			
Competence (C)	.922	.748	.150	.009	**.865**		
Relatedness (R)	.839	.638	.620	.296	.155	**.799**	
Intrinsic Motivation (IM)	.832	.635	.604	.272	.363	.429	**.797**

Note. *Extraction method: Principal Axis Factoring. Rotation method: Promax with Kaiser normalization.*
AVE = Average Variance Extracted; CR = Composite Reliability; Diagonal ***bold numbers*** *show the square root of AVE [7]*

Table 3

Pattern matrix with cross-loadings

Item	POG	A	C	R	IM
α	.831	.923	.918	.828	.843
POG1	**.454**	.101	-.050	.078	.224
POG2	**.543**	.111	-.015	.272	-.056
POG4	**.849**	-.024	-.017	-.066	-.091
POG5	**.791**	-.058	.063	-.026	.139
A1	-.097	**.985**	-.005	.008	.045
A2	.114	**.882**	.007	-.055	-.058
C1	-.091	.000	**.929**	-.018	.109
C2	-.080	.086	**.947**	.016	-.055
C3	.053	.009	**.765**	.102	.007
C4	.130	-.098	**.804**	-.085	-.067
R3	.153	-.026	.017	**.696**	-.078
R4	-.078	-.055	-.004	**.891**	.008
R5	-.064	.018	-.004	**.797**	.046
IM1	-.151	.006	-.031	-.025	**1.024**
IM2	.217	-.026	.028	-.097	**.728**
IM5	.078	-.003	.012	.181	**.572**

Note. *Extraction method: Principal Axis Factoring. Rotation method: Promax with Kaiser normalization. α: Cronbach's Alpha; POG: Perception of Gamefulness; A: Autonomy; C: Competence; R: Relatedness; IM: Intrinsic Motivation;*

Table 4

Results of mean comparison

Item	μ_0	μ_1	Dif. (CI 95 %) Lower	Upper	Sig.
POG1	5.18	5.71	-.952	-.111	**.013***
POG2	3.59	4.31	-1.189	-.256	**.003****
POG4	3.78	4.18	-.857	.065	.092
POG5	4.42	4.92	-.957	-.051	**.029***
A1	4.88	4.82	-.461	.586	.814
A2	4.63	4.65	-.522	.496	.959
C1	3.86	4.08	-.720	.289	.400
C2	3.81	3.95	-.634	.342	.556
C3	3.57	3.77	-.650	.247	.376
C4	3.47	3.85	-.859	.110	.129
R3	3.55	4.28	-1.194	-.262	**.002****
R4	3.28	4.06	-1.221	-.334	**.001****
R5	2.97	3.86	-1.350	-.429	**.000****
IM1	5.10	5.55	-.848	-.064	**.023***
IM2	5.33	5.72	-.749	-.032	**.033***
IM5	5.09	5.32	-.623	.147	.225

Note. *Scale: Likert (1-7); μ_x: Mean of Group Sample; CI: Confidence Interval; Sig.: ** $p \leq 0.01$; * $p \leq 0.05$; POG: Perception of Gamefulness; A: Autonomy; C: Competence; R: Relatedness; IM: Intrinsic Motivation;*

As last data analysis in the measurement model, we did an independent-samples t-test with a 95% confidence interval to compare the means of the survey results between the gamified, experimental group (n = 105) and the non-gamified, control group (n = 93). In advance, we did a Levene's test to check the requirement of equal variances.

On the one hand, the results of the mean comparison identified differences between the gamified and non-gamified group considering a gameful perception and intrinsic motivation (both $p \leq 0.05$). Especially for the construct of relatedness, the item means differ significantly ($p \leq 0.01$). On the other hand, we were, not able to measure any differences for the constructs of autonomy or competence.

Table 4 shows the results of the mean comparison with sample means for the control group (μ_0) and experimental group (μ_1) and their respective confidence intervals (CI).

4.2. Structural Model

We then estimated the revised model with MLE and determined the model fit indices. The results for our structual model indicate effects of gamification on gameful perception (POG) (β = 0.254; p = 0.001), which was expectable. Thus, hypothesis H1 is supported. Furthermore, we found that gameful perception affected the experience of some basic psychological needs. The influence of POG on relatedness showed the strongest effect (β = 0.654; p = 0.000), whereas the impact on experience of autonomy seemed moderate (β = 0.451; p = 0.000). In contrast, the relationship between a gameful perception and experiencing competence could not be confirmed (β = 0.149; p = 0.062). When analyzing for mediating effects, we did not find significant direct effects between gamification and either autonomy (β = −0.102; p = 0.150) or competence (β = 0.034; p = 0.647). In case of relatedness, however, gamification showed a significant direct effect (β = 0.163; p = 0.013). In addition, we determined significant indirect effects of gamification on autonomy (β = 0.114; p = 0.004), as well as on relatedness (β = 0.166; p = 0.002), but not on competence (β = 0.038; p = 0.107). According to this result, POG *fully* mediates the relationship between gamification and autonomy but only *partially* mediates the effect of gamification on relatedness. Thus, our hypotheses (H2a-c) are just partly supported.

As postulated by SDT [4], we are also able to prove the relationship of intrinsic motivation and needs of autonomy (β = 0.163; p = 0.021), competence (β = 0.316; p = 0.000) and social relatedness (β = 0.395; p = 0.000). Thus, the hypotheses derived from SDT (H3a-c) are also supported. Moreover, neither of the measured control variables has a significant positive association ($p \leq 0.05$) with POG or any other latent construct.

In terms of model fit, our structural model shows good to adequate values for different quality criteria [12]: RMSEA = 0.073 (< 0.08), CFI = 0.912 (> 0.9), TLI = 0.896 (~ 0.9), and SRMR = 0.077 (< 0.08). Figure 3 shows the estimated model with standardized coefficients.

5. Discussion

Overall, we found an evidence for a mediating effect of a user's gameful perception between gamification and intrinsic need satisfaction. We interprete that the reason for the *full* mediation in case of autonomy lies in the characteristic of play itself, which is regarded as a voluntary activity (and thus autonomous) [16]. Therefore, only if users perceive their activity as gameful, it's likely that they actually feel self-determined [4]. The reason for *partial* mediation in case of relatedness can be explained by the design of

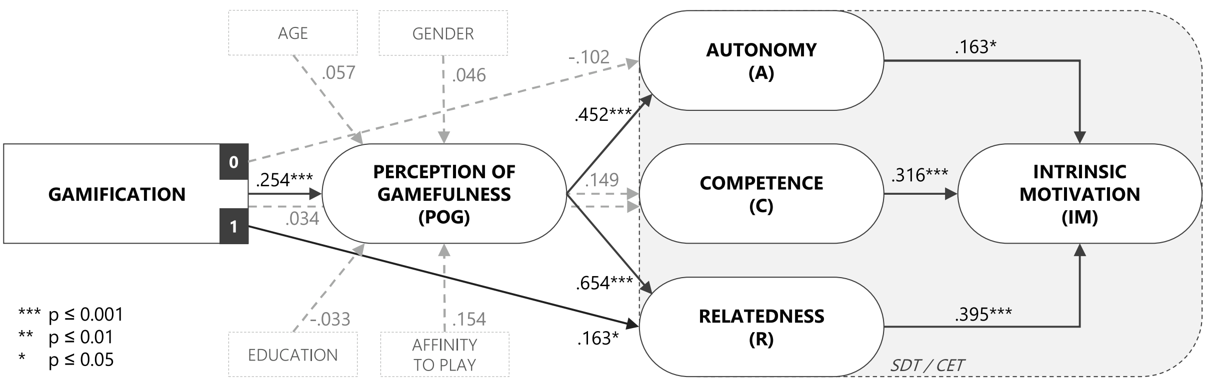

Figure 3: Results for structural model including control variables

our gamified app. The leaderboard or avatars, *e.g.*, reveal that other users are also involved in a session. Therefore, regardless of a gameful perception, users can still feel ”connected” [4]. Additionally, the relatedness is reinforced by a collaborative competition to beat the boss and win the quiz, which likely fuels their gameful perception [9]. This seems reasonable for us, but, some unexpected findings need discussion.

First, the mean comparison did not show differences for autonomy and competence between the groups. In the case of competence, the bots in the gamified version could be a possible cause for the missing change, as they outperformed our participants. The bots' average response time was 5.7 seconds lower, while having a success rate of 75%. In comparison to the 52% success rate of the participants, the integrated bots seemed a bit 'overpowered'. As the bots appeared on the top of the leaderboard, they may have inhibited the participants' experience of competence that we expected to emerge from the gameful feedback [22]. This flaw in our app design might have led to the low (non-significant) mediating effect. One reason for the lack of change in autonomy could also lie in the gameful design of the application: The gamified app might not have provided enough choice to support a strong feeling of autonomy [17]. For this aspect, the two groups were only treated differently due to the gamified character selection.

Second, the experience of relatedness differed strongly between the groups. In this case, the effect of gamification might be skewed, because the group sizes were slightly larger in the experimental group with 4.22 participants than in the comparison group with 2.40 participants in average. In general, this small difference in the social environment during the experiment might have moderated the effect. It therefore remains uncertain to what extent gamification was the underlying factor for the change in social relatedness.

Third, the control variable of affinity to play was slightly below the borderline to show a significant effect on gameful perception ($\beta = 0.154$; $p = 0.054$). We think that the affinity to play, however, can have an impact on gameful perception, as gamers have a wider experience of what an actual game situation feels like compared to non-gamers. Therefore, gamers might perceive a gameful situation easier as they can relate them to other game experiences. A possible reason, why this effect was statistically not significant could lie in the group composition, as the affinity to play was about 15% lower in the gamified group ($M = 2.486$) compared to the non-gamified group ($M = 2.925$).

6. Conclusion

The objective of this study was to analyze how a gameful perception influences gamification's motivational effect. Our results indicate the importance of a gameful perception in context of a gamified learning app to foster intrinsic motivation. The results, however, underly a few limitations which we briefly want to address to provide some ideas for future research.

During the experiment, both applications for knowledge retrieval represented a quiz, which can be considered per se as a game. In context of education, however, we want to point out the apps' ulitarian (and not hedonic) purpose to determine learning outcomes. That is why we refer to the control group as non-gamified in this study. Still, in other scenarios it might be reasonable to compare low and high degrees of gamification to determine gameful perception.

Furthermore, we want to stress out, that we selected and translated items from different, validated scales. To measure the user's gameful perception, we only considered *playfulness* as an indicator. Still, there are also other possible dimensions that induce a gameful experience which researchers can use to employ a fuller measure of gamefulness as a mediator [9].

Finally, we want to encourage researchers to consider gameful perception when analyzing motivational effects of gamification in future (replication) studies. Studies that examine the circumstances where individual game elements evoke a gameful perception and thus support intrinsic needs are still scarce. We also suggest to compare gamification's impact on gamers and non-gamers need satisfaction, as both target groups seem to have a different perspective on gamification.

References

[1] A. Azadvar, A. Canossa, UPEQ: Ubisoft perceived experience questionnaire: A self-determination evaluation tool for video games, *ACM Internat. Conference Proceeding Series* (2018) 1–8.

[2] K. A. Bollen. Structural Equations with Latent Variables, Wiley, New York, 1989.

[3] E. L. Deci, H. Eghrari, B. C. Patrick, D. R. Leone, Facilitating Internalization: The Self-Determination Theory Perspective, *Journal of Personality* 62 (1994) 119–142.

[4] E. L. Deci, R. M. Ryan, The "What" and "Why" of Goal Pursuits: Human Needs and the Self-Determination of Behavior, *Psychological Inquiry* 11 (2000) 227–268.

[5] S. Deterding, The Lens of Intrinsic Skill Atoms: A Method for Gameful Design, *Human-Computer Interaction* 30 (2015) 294–335.

[6] S. Deterding, D. Dixon, R. Khaled, L. E. Nacke, From game design elements to gamefulness: Defining "Gamification", *Proceedings of the 15th International Academic MindTrek Conference on Envisioning Future Media Environments* (2011) 1–7.

[7] C. Fornell, D. F. Larcker, Evaluating Structural Equation Models with Unobservable Variables and Measurement Error, *Journal of Marketing Research* 18 (1981) 39–50.

[8] M. D. Hanus, J. Fox, Assessing the effects of gamification in the classroom: A longitudinal study on intrinsic motivation, social comparison, satisfaction, effort, and academic performance, *Computers and Education* 80 (2015) 152–161.

[9] J. Högberg, J. Hamari, E. Wästlund, Gameful Experience Questionnaire (GAMEFULQUEST): an instrument for measuring the perceived gamefulness of system use, *User Modeling and User-Adapted Interaction* (2019) 619–660.

[10] R. Hunicke, M. LeBlanc, R. Zubek, MDA: A Formal Approach to Game Design and Game Research, *Workshop on Challenges in Game AI* (2004) 1–4.

[11] K. Huotari, J. Hamari, A definition for gamification: Anchoring gamification in the service marketing literature, *Electronic Markets* 27 (2017) 21–31.

[12] R. B. Kline. Principles and Practice of Structural Equation Modeling, The Guilford Press, New York, 2011.

[13] J. Koivisto, J. Hamari, The rise of motivational information systems: A review of gamification research, *International Journal of Information Management* 45 (2019) 191–210.

[14] A. Lieberoth, Shallow Gamification Testing Psychological Effects of Framing an Activity as a Game, *Games and Culture* 10 (2015) 229–248.

[15] E. D. Mekler, F. Brühlmann, A. N. Tuch, K. Opwis, Towards understanding the effects of individual gamification elements on intrinsic motivation and performance, *Computers in Human Behavior* 71 (2017) 525–534.

[16] R. M. Ryan, C. S. Rigby, A. Przybylski, The motivational pull of video games: A self-determination theory approach, *Motivation and Emotion* 30 (2006) 347–363.

[17] M. Sailer, J. U. Hense, S. K. Mayr, H. Mandl, How gamification motivates: An experimental study of the effects of specific game design elements on psychological need satisfaction, *Computers in Human Behavior* 69 (2017) 371–380.

[18] M. Sailer, L. Homner, The Gamification of Learning: a Meta-analysis, *Educational Psychology Review* 32 (2019) 77–112.

[19] S. Schöbel, A. Janson, K. Jahn, B. Kordyaka, O. Turetken, N. Djafarova, M. Saqr, D. Wu, M. Söllner, M. Adam, P. Heiberg Gad, H. Wesseloh, J. Leimeister, A Research Agenda for the Why, What, and How of Gamification Designs: Outcomes of an ECIS 2019 Panel, *Communications of the Association for Information Systems* 46 (2020) 706–721.

[20] A. Suh, C. Wagner, L. Liu, Enhancing User Engagement through Gamification, *Journal of Computer Information Systems* 58 (2018) 204–213.

[21] P. Tiago, M. J. Gouveia, C. Capinha, M. Santos-Reis, H. M. Pereira, The influence of motivational factors on the frequency of participation in citizen science activities, *Nature Conservation* 18 (2017) 61–78.

[22] H. Wesseloh, F. M. Stein, P. Szelat, M. Schumann, Boss Fights in Lectures! - A Longitudinal Study on a Gamified Application for Testing Factual Knowledge, *Proceedings of the 4th International GamiFIN Conference* (2020) 31–40.

[23] N. Xi, J. Hamari, Does gamification satisfy needs? A study on the relationship between gamification features and intrinsic need satisfaction, *International Journal of Information Management* 46 (2019) 210–221.