

A Systematic Review of the Effect of Gamification on Adherence across Disciplines

Robin De Croon, Jonas Geuens, Katrien Verbert, and Vero Vanden Abeele

KU Leuven, Department of Computer Science, Celestijnenlaan 200a, Leuven, Belgium

Abstract. Systematic reviews of gamification research often focus on effects on motivation and engagement. Fewer studies systematically investigate the effect of gamification on ‘adherence’, the extent to which individuals use a gamified service and experience its content, as envisioned by the creators, to derive a certain benefit. To this end, this paper presents a systematic review of the effect of gamification on adherence across disciplines, including only studies with experimental designs and empirical measurements of adherence. The results lend support to the hypothesis that gamification has a positive effect on adherence: 19 out of 27 studies report a significantly positive effect or trend. However, we also demonstrate that further debate is necessary on how to conceptualize and measure adherence across disciplines, and suggest that studies on gamification on adherence not only measure the usage behavior of individuals, but equally provide an *a priori* operationalization of intended use along with a justification.

Keywords: Gamification · adherence · retention rate · attrition · drop out · intended usage.

1 Introduction

Gamification, “*the intentional use of game elements for a gameful experience of non-game tasks and context,*” [1] is an interdisciplinary research area; a combination of game design, user experience design, behavioral economics and motivational psychology [2]. Gamification is most often a means to an end, i.e., many gamified services aim to increase motivation and engagement with the ultimate aspiration of promoting a certain behavior. To date, review studies that systematically investigate the effect of gamification mostly focus on a combination of psychological outcomes (e.g., intrinsic motivation, engagement), and/or behavioral outcomes at large (i.e., usage, retention) [3].

As the gamification field matures [4, 5], there is a call for a stronger effort in defining precise research questions on the basis of existing theorizations, and to further refine the understanding of the kind and size of the effects gamification has on individuals [5]. Hence, to further the field, it is needed to more clearly articulate the type of psychological and behavioral outcomes expected.

In this study, we focus specifically on one behavioral outcome of gamification, *adherence*, i.e., the extent to which individuals use a gamified service or system

and experience its content, as defined or implied by its creators, in order derive a certain benefit [6–8]. Thus far, only few gamification studies focus on *adherence* itself, and the studies that do typically limit themselves to one specific product and application domain, such as health apps [9, 10], online learning software [11, 12], or programming [13]. To date, there has been no study that systematically investigates the effect of gamification on adherent behavior across diverse implementations and disciplines.

The contributions of the paper are twofold. The first contribution is empirical in nature, as we report on the impact of gamification on adherence across disciplines, from 27 academic papers. The results of our systematic review lend support to the hypothesis that gamification has a positive effect on adherence. The second contribution is theoretical in nature, as we further the understanding of adherence in gamification studies and promote both a more refined conceptualization as well as a set of standard elements to be present in any gamification study on adherence.

2 Background

We first discuss the concept of adherence and define it as one specific behavioral outcome, that is pertinent for much gamification research yet different from engagement. We end this section with the research objectives of the paper.

2.1 The concept of adherence

Although the term adherence is perhaps most established in the health domain [14], the concept itself has long been used in other domains as well. The Oxford dictionary defines adherence in layperson terms as “*attachment or commitment to a person, cause, or belief*”. From this definition, we learn that adherence is not about short lived, single-point-in-time behavior, attachment and commitment imply *sustained behavior*. The Cambridge dictionary defines adherence as: “*the fact of someone behaving exactly according to rules, beliefs, etc.*” This definition highlights the existence of rules and beliefs, in other words, there is an envisioned *intended behavior*. Hence, from these two definitions, the reader can understand that adherence implies both a temporal aspect (behavior as it unfolds over a longer period) and an intended usage aspect (behavior as according to rules, beliefs).

More targeted definitions that elaborate on these temporal and behavioral aspects can be found in the health domain. For example, the World Health Organization (WHO) defines adherence as “*the extent to which a person’s behaviour - taking medication, following a diet, and/or executing lifestyle changes, corresponds with agreed recommendations from a health care provider*.” [14] With respect to eHealth, as a form of interactive technology close to the realm of gamified technology, Christensen et al. [15] put forward the following definition of adherence: “*the extent to which individuals experience the content of the Internet intervention*.” Again, these definitions encompass a temporal

aspect (experiencing a certain ‘dosage’ of content) and intended usage (following agreed recommendations). To this end, Kelders et al. [6] promote a definition of adherence for interactive systems and services in eHealth that encapsulates both aspects, as “*the extent to which individuals experience the content to derive maximum benefit from the intervention, as defined or implied by its creators.*” Interestingly, the aforementioned definition of adherence lends itself well to the domain of gamification, as “*the extent to which individuals use a gamified service or system and experience its content, as defined or implied by its creators, in order to derive a certain benefit,*” and can also apply to situations beyond eHealth, such as e-learning and customer loyalty.

Basing themselves on Kelders’ definition of adherence, Sieverink et al. [7] suggest three elements to be present in any adherence studies:

1. The ability to measure the usage behavior of individuals.
2. An operationalization of intended use.
3. An empirical, theoretical, or rational justification of the intended use.

In combination, these three elements not only ensure empirical measurement of the behavioral outcome, but also that this outcome is compared to a pre-specified value or threshold (~intended use) and that authors need to be able to justify this value or threshold.

2.2 Adherence versus Engagement

At the heart of gamification lie the interrelated concepts of engagement and behavior change [1]. Therefore, gamification scholars may consider adherent behavior simply as an outcome of increased engagement. While related, adherence and engagement are different concepts. For example, people can be engaged to lose weight, but still not adhere to their weight plan. As aforementioned, adherence starts from the notion of usage behavior that is sustained and as intended, whereas engagement on the other hand, foreground the affective, psychological experience. For example, Brown and Cairns [16], Brockmyer et al. [17], and Denisova et al. [18] conceptualize engagement as a multi-dimensional construct encompassing a user’s *absorption*, *flow*, *presence*, and *immersion*. Hence, it is possible for a gamified intervention to have engaged users who do not fully experience the content of the services as intended by its creators or prescribers (and are thus not adherent). Vice versa, users may show adherence to the recommended gamified plan, yet show a lack of engagement. To date, as the lines are blurry, the elusive concept of ‘engagement’ is frequently measured through behavioral variables such as ‘returning visits’ or ‘regular use’ [19]. Also the opposite exists, where a less strict understanding of the concept of adherence equates it to measures of engagement. For example, [20] and [21] measure non-adherence as a “*lack of participant engagement*,” i.e., they are using engagement as a proxy for adherence. Yet as the field of gamification matures, it is beneficial to more clearly delineate and separate these theoretical concepts, to refine our understanding and measurement of the impact of specific gamification strategies.

2.3 Studies on gamification and adherence

There are multiple review studies on the effect of gamification on psychological (e.g., engagement, motivation), behavioral outcomes (e.g., retention, increased usage) outcomes, e.g., [1, 22–25]. However, only the systematic review study of Brown et al. [26] on Web-Based Mental-Health Interventions explicitly studies adherence. The authors reported that web-based health interventions incorporating gamification features had a higher mean adherence rate. Yet, they also found both adherence and usage data were inconsistent or underreported.

In sum, the aforementioned study [26] suggests that gamification has a positive impact on adherence. However, it is limited to the mental health domain. To this end, this study sets out to broaden this scope and systematically investigate the effect of gamification on adherent behavior, across applications and/or disciplines. We set out to explore to what extent current studies conceptualize adherence according to the definition put forward by Kelders et al. [6] and adhere to standards recommended by Sieverink et al. [7], i.e., measuring usage behavior, operationalization of intended use, and a justification of intended use. Additionally, we aim to broaden our understanding of which gamification techniques are most popular and have the strongest impact, and which disciplines perform most studies on gamification and adherent behavior.

3 Material and Methods

3.1 Search String

The protocol that was used to find and review the studies was developed according to the PRISMA guidelines [27]. In this systematic review we solely focus on gamification techniques hence the truncated keyword ‘*gamif**’. We used Brown et al. [26]’s synonyms for adherence. However, we modified ‘retention rate’ to ‘retention’ to also include studies that, for example, report on customer and employee retention. Finally, ‘compliance’ and ‘concordance’ were added to include papers that use a less authoritative approach to describe adherence [28]. We therefore built on their work and extended the search string:

gamif AND (adherence OR attrition OR dropout OR drop-out OR non-completers OR non-completers OR “lost to follow up” OR withdrawal OR nonresponse OR non-response OR “completion rate” OR “did not complete” OR retention OR loss OR compliance OR concordance).*

3.2 Data collection

To find gamification studies across disciplines, a comprehensive search of seven electronic databases was conducted and produced a set of 1122 papers: Scopus (life sciences, social sciences, physical sciences and health sciences, n=300), PubMed (life sciences and biomedical, n=86), ACM Digital Library (all computing and information technology domains, n=222), IEEExplore (computer science, electrical engineering and electronics, n=56), Web-of-Science (multidisciplinary, n=193), ScienceDirect (physical sciences and engineering, life sciences, health sciences, and social sciences and humanities, n=12), and ProQuest (multidisciplinary, n=253).

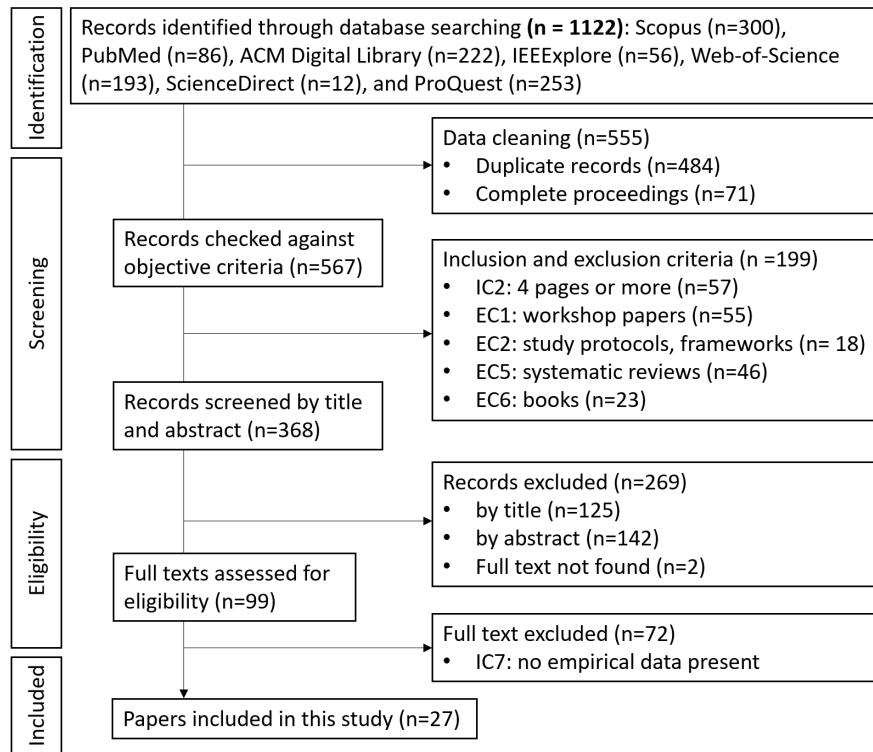


Fig. 1: Flow diagram according to the PRISMA guidelines.

3.3 Inclusion criteria

Our review focused on high-quality research reporting original work on the effect of gamification on adherence. From this perspective, we developed the following inclusion criteria:

1. Peer-reviewed conference or journal papers.
2. Full papers (minimum length of four pages).
3. Explained research methods.
4. Researched effect of gamification on adherence as main research subject.
5. Reported how gamification was applied.
6. Reported the effect of gamification on adherence.
7. Reported behavioral or attitudinal measurements.

Criteria 1-2 were chosen to maximize the inclusion of high-quality and original research. Criteria 3-4 were included to enable an assessment of the quality of the work. Criteria 5 ensured that the included papers report on gamification, and not on serious games or persuasive technology. Finally, criteria 6-7 were chosen to ensure the included papers research the effect of gamification on adherence in a user study, and not only provide a conceptual discussion.

3.4 Exclusion criteria

The exclusion criteria were designed to exclude duplicate reporting of earlier versions of studies fully reported later. We excluded papers with the following characteristics:

1. Extended abstracts, work-in-progress, workshops.
2. Study protocols or conceptual designs.
3. Studies that only cover serious games.
4. Studies that do not report an effect.
5. Systematic reviews.
6. Non-scholarly books.
7. Papers not written in English.

Criteria 1-2 exclude early and incomplete versions of studies. Criteria 3 excludes studies that mislabel serious games as gamification. Criteria 4 makes sure that the effect on adherence can be compared. Criteria 5 excludes studies that did not focus on one particular study. Criteria 6 excludes books that do not have a scholarly focus. Finally, we only included original research written in English.

3.5 Classification

Effect: Studies were classified as ‘significantly positive’ when they explicitly mention a significant positive effect. When they only mention a trend, they are classified as ‘positive trend’. Studies reporting no effect are classified as ‘no effect’. Studies that report a negative effect are classified as ‘negative’.

Adherence: To research how adherence is defined and measured in the gamification domain, we built on [26]’s findings to classify the papers: attrition, dropout, noncompleters, lost to follow-up, participant withdrawal, nonresponse, completion rate, did not complete, retention, loss, and compliance. Effectiveness was also added as a classification term as two papers explicitly used the term effectiveness. Additionally, drawing on Sieverink et al. [7], we developed an Adherence Rationale Index (ARI), to classify studies as follows:

- A study specifies intended use and provides a theoretical justification,
- B study specifies intended use but lacks a theoretical justification,
- C study neither specifies intended use nor theoretical justification.

Gamification techniques: In this systematic review, we used the classification proposed by Hamari et al. [3] to classify gamification techniques as shown in Table 2. The following list was used: points, leaderboards, achievements/badges, levels, story/theme, clear goals, feedback, rewards, progress, and challenge. However, as multiple studies included social motivational affordances into their gamification research, we augmented Hamari’s set with ‘social affordances’ which grouped guilds/teams, social network, social status, opponents, and direct communication. Finally, although serious game studies are excluded by the exclusion criteria, ‘serious game’ was added to the coding table as four studies [10, 29–31] each apply three conditions in their study design: the control condition, a gamified intervention, and a serious game as a third condition.

Study design and criteria: Each paper was classified as either a randomized control study (RCT) or as a baseline study. Studies were classified as an RCT when they had a randomized control group in parallel with the intervention group. Studies were classified as a baseline study when they could compare their results to a baseline value. Additionally, the sample size, demographics, and duration of each study were listed.

Scientific fields: A scoping review by O’Donnel et al. [32] shows that gamification became a multidisciplinary research topic, applied and used in several domains. To remain consistent, we used O’Donnel et al. [32]’s ten categories to classify the primary scientific fields of the papers: (1) Sciences; (2) Information & Computing Science & Technology; (3) Medical & Health Sciences; (4) Education; (5) Economics, Commerce, Management, Tourism & Services; (6) Psychology & Cognitive Sciences; (7) Law & Legal Studies; (8) Engineering, Built Environment & Design; (9) Arts, Humanities, & Social Sciences; and (10) Games, Digital Entertainment Media. This is a condensed format of the 22 top-level divisions of the Australian and New Zealand Standard Research Classification.

3.6 Intercoder reliability

All studies were coded and calculated by two independent coders (RDC and JG). Intercoder reliability was calculated using Cohen's kappa statistic. The mean value was 0.78 (± 0.13) and all values were significant $P < 0.001$. Overall, all intercoder reliability values were at an acceptable level, i.e., > 0.60 [33, 34].

All the adherence terms, rationales (Table 1), as well as the scientific fields were consistently coded with a kappa agreement of 0.82 ($P < 0.001$), 1.0, and 0.80 ($P < 0.001$) respectively. The gamification techniques were also reliably coded and values were found to be between 0.62 and 1.0 kappa agreement depending on the technique. The lowest rated principle in terms of intercoder reliability was 'story/theme' (kappa agreement: 0.62). The lower agreement is due to the blurry line between graphic additions and a theme. For example, is the addition of a penguin [35] a graphical asset, a story, or even an avatar?

4 Results

As shown in Figure 1, a total of 1122 papers were retrieved from the database searches by using the search terms described in Section 3.1. After removing duplicates and filtering papers based on the inclusion and exclusion criteria, 99 papers were evaluated by considering their full texts. Twenty-seven papers focusing on both gamification and adherence met the criteria and were thus included in this systematic review.

4.1 Effect of gamification on adherence

The results as reported by the authors are summarized in Tables 1, 2, and 3: of the 27 studies, 33% reported a scientifically significant positive effect of gamification techniques on adherence. Additionally, 37% reported positive trends, but could not provide significant effects. Finally, 30% reported no effects at all, while no studies reported a negative effect. However, a decrease in adherence over time was reported by 26% [29, 30, 36–40], either to the intervention or to the gamification techniques themselves. For example, Bodduluri et al. [29]'s results suggest "*that the motivating effects of gamification 'wear-off' and become boring as participants continue in a session that is lengthy, unless there is greater variety or progression of challenge in the task.*" This is similar to Dugas et al. [36] who also state that their participants' motivation diminished as the study continued, which resulted in decreasing adherence over time. Fotaris et al. [30] noticed the demotivating aspect of leaderboards as "*students [...] began to lose interest once they trailed behind in the leaderboard.*"

Table 1: Overview of the terms used: ARI (*A*: *studies specify intended use and provide a theoretical justification*, *B*: *studies specify intended use but do not provide a theoretical justification*, *C*: *studies neither specify intended use nor provide a theoretical justification*), measurement variables, study design, sample size, duration, and reported outcome (++ significantly positive, + positive trend, = no effect).

<i>Study Term</i>	<i>ARI</i>	<i>Measurement variable</i>	<i>Design</i>	<i>Sample</i>	<i>Days</i>	<i>Effect</i>
[41] completion rate	B	average percentage of students that submitted a solution to an assignment	RCT	37	28	++
[29] completion rate	C	absolute completion rate between three interventions	baseline	144	1	=
[42] adherence	C	number of self-reported sessions	RCT	78	30	=
[43] adherence	A	average value of daily average frequency of blood glucose	baseline	20	84	++
[44] compliance	C	one inspection per day by a project manager whether work tools are returned and employees are compliant with safety regulations	baseline	28	31	+
[45] completion rate	C	class attendance & attitudinal survey	RCT	371	365	=
[36] adherence	C	points for achieving daily goals related to reporting and reaching target levels of glucose, exercise, nutrition, and medication adherence	baseline	27	91	++
[30] completion rate	C	average class attendance and exercise completion rate	baseline	106	84	+
[37] compliance	A	average daily steps and active minutes	RCT	146	60	+
[46] retention rate	C	total number of students in class (anonymous) and number of students who registered and participated at the final exam	baseline	1375	70	++
[47] effectiveness	C	time spent, the number of exercises completed, and the number of words used	RCT	94	84	=
[48] retention rate	C	completed at least one quiz per week	baseline	1763	56	+
[11] retention rate	C	number of videos watched	RCT	206	28	++
[49] dropout	C	time spent in application	baseline	?	93	++
[10] compliance	A	number of times the user logged into the application, physical activity level based on wearable measurements	RCT	498	186	=
[50] attrition	B	mean number of sessions completed per participant and percentage of users that completed at least X sessions	RCT	482	6	=
[51] attrition	C	completed questions and number of players that keeps on playing each day	baseline	92	97	++
[52] dropout	C	number of participants that did not complete the study	RCT	41	20	=
[12] retention rate	C	percentage of students retained	baseline	394	84	+
[53] retention rate	C	the extent to which children retain their willingness in completing skill-builder modules between sessions	baseline	30	8	=
[13] compliance	B	coding convention compliance as assessed by automatic Checkstyle script	RCT	17	47	+
[38] compliance	B	number of participants who continued to log steps at least once per week and compliance to step goal (10,000)	RCT	110	50	+
[31] adherence	B	number of application sessions logged	baseline	24	17	+
[40] retention rate	C	user retention curve as defined by the proportion of users who revisited the tool and the frequency and duration of their sessions	RCT	206	60	+
[35] attrition	C	number of usable trials, i.e., correctly labeling the target image	RCT	16	1	++
[39] retention rate	C	completed workout sessions	RCT	30	70	++
[54] effectiveness	C	daily number of times the application was accessed and daily walking time	RCT	99	14	++

4.2 Adherence Measurement Variables

As illustrated in Table 1, not adherence (15%) but retention (26%) was the term used most frequently. Other terms were compliance (19%), completion rate (15%), attrition (11%), effectiveness (7%), and dropout (4%). Scrutinizing the ARI, the majority of studies (70%) neither specified intended use, nor provided a theoretical justification. Instead, they followed a “*the more, the better*” approach. Just 19% specified intended use but lacked theoretical foundation. Only three studies both specified predefined use and provided *some* theoretical foundations. Cafazzo et al. [43] assess treatment adherence at baseline and post-intervention using the validated 14-item Self-Care Inventory [55]. A participant is adherent when they have three or more measurements, as “*frequent self-monitoring of blood glucose (\geq times daily) is associated with better glycemic control among patients with type 1 diabetes.*” Gremaud et al. [37] selected a 1250 steps per day threshold as a conservative estimate based on previous research that found adding 1385 steps per day resulted in significant reductions in multiple cardiometabolic risk factors. Finally, Leinonen et al. [10] followed the Finnish national recommendations for those in the age group of 13 to 18 years as at least 1.5 hours of daily physical activity [56]. In the end, most studies simply calculated adherent behavior by measuring and comparing a quantitative, behavioral measure. This ranges from the daily average frequency of blood glucose measurements [43], the number of exercises completed [47], to the number of app sessions [40]. Even studies that conceptualized engagement as (part of) adherence still relied on behavioral measurements exclusively. For example, Stanculescu et al. [40] claim that “*The average session length falls into the online behavior metrics and is a good indicative of user engagement.*” Finally, some studies, such as the one from Dugas et al. [36] use a combination of variables to calculate adherence: “*points were used to assess treatment adherence during the intervention. Points were allocated for achieving daily goals related to reporting and reaching target levels of glucose, exercise, nutrition, and medication adherence.*”

4.3 Gamification techniques

A multitude of gamification techniques were being integrated to improve adherence as shown in Table 2. All studies used at least three different gamification techniques with on average 5.7 (± 1.8) techniques. The majority of the studies implemented points (85%) and feedback (67%). These points were often simple numerical values for an action or a combination of actions and were used in distinct forms. For example, Ryan et al. [38] use the step count as points, while Dugas et al. [36] use traditional points. Points are typically displayed on a leaderboard (63%) or used to calculate badges/achievements (52%). All four studies [31, 45, 47, 48] that did not use points focused on progress in a certain theme/story.

Feedback was also implemented in a great variety, either immediate feedback with pop-up messages, e.g., [54], or in the form of reports, e.g., [13]. Additionally,

information visualizations [43], summary screens [11], or the ability to monitor results [50] were classified as feedback. Rewards were implemented with a similar variety, such as virtual rewards [31, 38, 54], candy [30], iTunes music [43], actual money [13, 50], physical trophies [51] or grades [13].

Another highly popular technique was the addition of social affordances (56%). Scase et al. [31] reported that the bonding aspect between participants helped encourage their participants to adhere. Like all other gamification techniques, social affordances were implemented differently: social networking [43, 10, 31, 39, 46, 54], teams [13, 30, 38, 51] opponents [11], social status [30], and communication features [36].

Challenges (30%) and levels (30%) were found to be the least commonly used gamification technique in the selected studies.

Table 2: Frequency of the gamification techniques used in the included studies grouped by reported outcome. Note that all studies use at least three different gamification techniques.

<i>Gamification Techn.</i>	<i>Significantly positive study</i>	<i>Positive trend study</i>	<i>No effect study</i>	<i>Total</i>
points	[43, 11, 36, 39, 41, 46, 49, 51, 35, 54]	[12, 13, 30, 37, 38, 40, 44]	[10, 29, 42, 50, 52, 53]	85%
feedback	[43, 11, 39, 41, 51, 35, 54]	[13, 30, 38, 40, 44, 48]	[10, 29, 42, 50, 52]	67%
leaderboards	[11, 39, 41, 46, 49, 51, 54]	[12, 13, 30, 37, 38, 40, 44]	[10, 45, 52]	63%
social affordances	[43, 11, 36, 39, 46, 51, 54]	[13, 30, 31, 37, 38, 40]	[10, 45]	56%
progress	[11, 36, 39, 41, 51, 54]	[44]	[31, 38, 42, 45, 47, 48, 50, 53]	56%
achievements/badges	[43, 11, 39, 41, 49, 51]	[12, 30, 38, 40, 44]	[45, 47, 53]	52%
rewards	[43, 41, 51, 54]	[13, 30, 31, 38, 40, 44]	[45, 50, 53]	48%
story/theme	[39, 35]	[30, 37]	[10, 31, 42, 47, 48, 50, 52]	41%
clear goals	[43, 36, 35, 54]	[13, 38, 40]	[42, 50]	33%
levels	[39, 51]		[29, 42, 50, 45, 47, 53]	30%
challenge	[11, 39, 46]	[30, 37]	[45, 52, 53]	30%
serious game		[30]	[29, 31, 53]	15%

4.4 Study design criteria

An RCT approach was used by 56% of the studies to study the effect of gamification on adherence, while 44% of the studies compared the gamified version to some baseline (see Table 1). A difference was most noticeable in the papers that reported no effect of gamification on adherence: six RCTs reported no effect, while only two baseline studies reported no effect.

Sample sizes varied greatly, ranging from 16 [35] to 1763 (284 versus 1479 control) [48]. The median sample size is 97 participants, while the lower and upper quartile lie between 30 and 200 participants. The duration also varied ranging from one day [35] (adherence measured as more usable trials) to one year [45] as shown in Table 1.

4.5 Scientific fields and gamification studies on adherence

As shown in Table 3, the Education, Medical & Health Sciences, and Psychology & Cognitive Sciences fields are well represented in the selected studies with ten, nine, and four studies respectively. Humanities & Social Sciences, Economics, Engineering, and ICT were also represented with one study each.

Table 3: The representative scientific fields of study of the included studies grouped by reported outcome.

Field	Sign. positive (33%)	Pos. trend (37%)	No effect (30%)	Total
Education	6 [11, 41, 46, 49, 51, 35]	3 [12, 30, 48]	1 [45]	10 (37%)
Medical & Health Sciences	3 [43, 39, 54]	4 [31, 36–38]	2 [10, 29]	9 (33%)
Psychology & Cognitive Sciences	0	0	4 [42, 50, 47, 53]	4 (15%)
Arts, Humanities, & Social Sciences	0	0	1 [52]	1 (4%)
Economics; Commerce, Management, Tourism & Services	0	1 [40]	0	1 (4%)
Engineering, Built Environment & Design	0	1 [44]	0	1 (4%)
Information & Computing Science & Technology	0	1 [13]	0	1 (4%)

5 Discussion

We first discuss the results obtained with respect to the impact of gamification on adherence across disciplines. Next, we reflect on the extent to which the recommendations by Sieverink et al. [7] were found. We end the discussion section with a reflection on the specific gamification strategies used.

5.1 Impact of gamification on adherence

The results of our systematic review lend support to the hypothesis that gamification has a positive effect on adherence. Nineteen out of 27 studies report a significantly positive effect or a positive trend. We emphasize that we applied strict standards with respect to scientific quality, and only included peer-reviewed conference or journal papers, and limited ourselves to studies that (1) conducted an experimental design and (2) studied adherence as a research goal and not as a consequence of the methodology used. Moreover, sample sizes and intervention duration were high among studies, suggesting that research studies have been conducted in a scientifically adequate manner.

Nevertheless, we must remain cautious. Although none of the papers reported a negative effect, 26% mentioned some form of a decrease in adherence over time. This is in line with the work of Koivisto et al. [57] who found that the perceived usefulness of gamification declines with use. This novelty effect should be considered when evaluating gamified systems and services as it can skew the results [58] with respect to adherence studies. This also highlights the importance of specifying how adherence is measured.

5.2 Pre-specifying and justifying adherence measurements

Our study also foregrounds that further debate is necessary on how to conceptualize and measure adherent behavior in gamification studies. As shown in Table 1, all authors used distinct terms and distinct measurement variables. This is of course a natural consequence of the different disciplines and research objectives. Yet, transcending the different disciplines, we found both intended use and empirical, theoretical, or rational justifications of intended use were mostly lacking. Only five studies did specify intended use [13, 31, 38, 41, 50], and only three [10, 43, 37] did provide a additional justification of the intended use [7].

Unfortunately, the lack in specifying intended use *a priori*, and the lack of providing an accompanying justification might introduce researcher bias, as a researcher is currently completely unconstrained in defining what and how to measure. Moreover, it is hard to compare effects of gamification on adherence when measurements variables and thresholds differ between studies. The lack of a pre-specified intended use and the lack of justification in the current studies urges us to remain cautious in making bold claims about the impact of gamification on adherence.

5.3 Gamification techniques

We found that all papers used a combination of minimum three distinct gamification techniques, with on average 5.7 (± 1.8) techniques. At first sight, this large average number may indicate the maturity of the implementations. On the other hand, the frequent occurrence of points and leaderboards might also suggest that the included studies focus largely on the PBL (points, badges, leaderboards) triad [59]. Limiting gamification to the PBL triad may result in failing to capture what makes games engaging, which in its turn leads to ineffective systems, and this is fundamental criticism of the field [60]. Due to the large distribution of gamification techniques and the large variety in their implementation (and perhaps the quality), no conclusions can be made about adherence and its relation to specific gamification techniques used in the study.

6 Limitations

This study has a number of limitations that affect the contribution. First, the terms used in the search string might have impacted the results, as we did not include domain-specific constructs of adherence, such as customer loyalty and learner conversion. Such studies may in fact implicitly report adherence measurements when they report impact. In these studies, however, improving adherence is often not the main goal, and therefore were excluded from this study. Moreover, we based ourselves for our search string on the findings of Brown et al. [26]. We acknowledge that these terms could have induce a bias towards papers in the health domain. However, 37% of the studies were from the education domain, 33% of the health domain, and 15% from psychology

and cognitive sciences. This suggest that we were able to include studies across disciplines.

Second, we did not score for ‘quality’ of the implemented gamification techniques or the gamified application. Gamification designers often critique the approach of using individual gamification techniques without acknowledging the quality of the implementation [61]. Moreover, gamified applications are perceived as ‘gestalts’ by their users, they are perceived as one whole, rather than a mere atomistic addition of gamification elements [60]. Hence, future research may attempt to include measures of quality of implemented gamification techniques as perceived by end-users.

7 Conclusions and future work

This paper first explored the concept of adherence and presented a tailored definition for the gamification domain as “*the extent to which individuals use a gamified service or system and experience its content, as defined or implied by its creators, in order to derive a certain benefit.*” Next, it reports on a systematic literature review summarizing the published research on the effect of gamification on adherence, across disciplines. Twenty-seven papers focusing on both gamification and adherence, and including empirical measurements, met the criteria. The results of our systematic review lend support to the hypothesis that gamification has a positive effect on adherence. However, our results also suggest the need for a more refined conceptualization, as well as a set of standard elements to be present in any gamification study on adherence: (1) the ability to measure the usage behavior of individuals, (2) an operationalization of intended use, (3) an empirical, theoretical, or rational justification of the intended use.

Acknowledgments

This work is part of the research projects PANACEA Gaming Platform with project number HBC.2016.0177, and Personal Health Empowerment with project number HBC.2018.2012, which are financed by Flanders Innovation & Entrepreneurship.

References

1. K. Seaborn and D. I. Fels, “Gamification in theory and action: A survey,” *International Journal of human-computer studies*, vol. 74, pp. 14–31, 2015.
2. Y.-K. Chou, *Actionable gamification: Beyond points, badges, and leaderboards*. Octalysis Group, 2016.
3. J. Hamari, J. Koivisto, and H. Sarsa, “Does gamification work? – a literature review of empirical studies on gamification,” in *2014 47th Hawaii international conference on system sciences*, (Hawai, USA), pp. 3025–3034, IEEE, 2014.
4. L. E. Nacke and S. Deterding, “The maturing of gamification research,” *Computers in Human Behaviour*, vol. 71, pp. 450–454, 2017.

5. A. Rapp, F. Hopfgartner, J. Hamari, C. Linehan, and F. Cena, "Strengthening gamification studies: Current trends and future opportunities of gamification research," *International Journal of human-computer studies*, vol. 127, pp. 1–6, 2019.
6. S. M. Kelders, R. N. Kok, H. C. Ossebaard, and J. E. W. C. Van Gemert-Pijnen, "Persuasive system design does matter: a systematic review of adherence to web-based interventions," *Journal of medical Internet research*, vol. 14, no. 6, p. e152, 2012.
7. F. Sieverink, S. M. Kelders, and J. E. van Gemert-Pijnen, "Clarifying the concept of adherence to ehealth technology: Systematic review on when usage becomes adherence," *Journal of medical Internet research*, vol. 19, no. 12, p. e402, 2017.
8. H. Christensen, K. M. Griffiths, A. E. Korten, K. Brittliffe, and C. Groves, "A comparison of changes in anxiety and depression symptoms of spontaneous users and trial participants of a cognitive behavior therapy website," *Journal of medical Internet research*, vol. 6, p. e46, Dec. 2004.
9. R. De Croon, A. M. Bezabih, J. Geuens, D. Wildemeersch, D. Oeyen, K. Verbert, and V. Vanden Abeele, "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)*, pp. 1–12, 2019.
10. A.-M. Leinonen, R. Pyky, R. Ahola, and et al., "Feasibility of Gamified Mobile Service Aimed at Physical Activation in Young Men: Population-Based Randomized Controlled Study (MOPO)," *JMIR mHealth and uHealth*, vol. 5, no. 10, p. e146, 2017.
11. M. Krause, M. Mogalle, H. Pohl, and J. J. Williams, "A playful game changer: Fostering student retention in online education with social gamification," in *Proceedings of the Second (2015) ACM Conference on Learning @ Scale, L@S '15*, (New York, NY, USA), pp. 95–102, ACM, 2015.
12. E. Pechenkina, D. Laurence, G. Oates, D. Eldridge, and D. Hunter, "Using a gamified mobile app to increase student engagement, retention and academic achievement," *International Journal of Educational Technology in Higher Education*, vol. 14, no. 1, pp. 1–12, 2017.
13. C. R. Prause and M. Jarke, "Gamification for enforcing coding conventions," in *Proceedings of the 2015 10th Joint Meeting on Foundations of Software Engineering*, (New York, NY, USA), pp. 649–660, ACM, 2015.
14. E. Sabaté, "Adherence to long-term therapies," tech. rep., {World Health Organization}, 2013.
15. H. Christensen, K. M. Griffiths, and L. Farrer, "Adherence in internet interventions for anxiety and depression," *Journal of medical Internet research*, vol. 11, p. e13, Apr. 2009.
16. E. Brown and P. Cairns, "A grounded investigation of game immersion," in *CHI '04 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '04, (New York, NY, USA), p. 1297–1300, Association for Computing Machinery, 2004.
17. J. H. Brockmyer, C. M. Fox, K. A. Curtiss, E. McBroom, K. M. Burkhardt, and J. N. Pidruzny, "The development of the game engagement questionnaire: A measure of engagement in video game-playing," *Journal of Experimental Social Psychology*, vol. 45, no. 4, pp. 624 – 634, 2009.
18. A. Denisova, A. I. Nordin, and P. Cairns, "The convergence of player experience questionnaires," in *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play*, CHI PLAY '16, (New York, NY, USA), pp. 33–37, ACM, 2016.

19. L. Sardi, A. Idri, and J. L. Fernández-Alemán, “A systematic review of gamification in e-health,” *Journal of Biomedical Informatics*, vol. 71, pp. 31 – 48, 2017.
20. D. Guertler, C. Vandelanotte, M. Kirwan, and M. J. Duncan, “Engagement and nonusage attrition with a free physical activity promotion program: The case of 10,000 steps australia,” *Journal of medical Internet research*, vol. 17, p. e176, July 2015.
21. J. E. Saul, M. S. Amato, S. Cha, and A. L. Graham, “Engagement and attrition in internet smoking cessation interventions: Insights from a cross-sectional survey of “one-hit-wonders”,” *Internet Interventions*, vol. 5, pp. 23–29, Sept. 2016.
22. L. Hassan and J. Hamari, “Gameful civic engagement: A review of the literature on gamification of e-participation,” *Government Information Quarterly*, p. 101461, 2020.
23. J. Koivisto and J. Hamari, “The rise of motivational information systems: A review of gamification research,” *International Journal of Information Management*, vol. 45, pp. 191 – 210, 2019.
24. J. Koivisto and A. Malik, “Gamification for older adults: A systematic literature review,” *The Gerontologist*, 2020.
25. J. Looyestyn, J. Kernot, K. Boshoff, J. Ryan, S. Edney, and C. Maher, “Does gamification increase engagement with online programs? A systematic review,” *PLoS One*, vol. 12, no. 3, p. e0173403, 2017.
26. M. Brown, N. O'Neill, H. van Woerden, P. Eslambolchilar, M. Jones, and A. John, “Gamification and Adherence to Web-Based Mental Health Interventions: A Systematic Review,” *JMIR Mental Health*, vol. 3, p. e39, Aug. 2016.
27. D. Moher, A. Liberati, J. Tetzlaff, D. G. Altman, and Prisma Group, “Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement,” *PLoS Med.*, vol. 6, no. 7, p. e1000097, 2009.
28. E. Grönvall, N. Verdezoto, N. Bagalkot, and T. Sokoler, “Concordance: A critical participatory alternative in healthcare IT,” in *Aarhus Series on Human Centered Computing*, vol. 1, p. 4, Oct. 2015.
29. L. Bodduluri, M. Y. Boon, M. Ryan, and S. J. Dain, “Impact of Gamification of Vision Tests on the UX,” *Games for Health Journal*, vol. 6, no. 4, pp. 229–236, 2017.
30. P. Fotaris, T. Mastoras, R. Leinfellner, and Y. Rosunally, “From hiscore to high marks: Empirical study of teaching programming through gamification,” in *Proceedings of the European Conference on Games-based Learning*, (Bristol, UK), pp. 186–194, UWE Bristol, 2015.
31. M. Scase, B. Marandure, J. Hancox, K. Kreiner, S. Hanke, and J. Kropf, “Development of and adherence to a Computer-Based gamified environment designed to promote health and wellbeing in older people with mild cognitive impairment,” *Studies in Health Technology and Informatics*, vol. 236, pp. 348–355, 2017.
32. N. O'Donnell, D. L. Kappen, Z. Fitz-Walter, S. Deterding, L. E. Nacke, and D. Johnson, “How multidisciplinary is gamification research?: Results from a scoping review,” in *Extended Abstracts Publication of the Annual Symposium on Computer-Human Interaction in Play*, CHI PLAY '17 Extended Abstracts, (New York, NY, USA), pp. 445–452, ACM, 2017.
33. R. J. Landis and G. G. Koch, “The measurement of observer agreement for categorical data,” *Biometrics*, vol. 33, pp. 159–174, Mar. 1977.
34. M. Lombard, J. Snyder-Duch, and C. C. Bracken, “Content analysis in mass communication: Assessment and reporting of intercoder reliability,” *Human Communication Research*, vol. 28, no. 4, pp. 587–604, 2002.

35. T. S. Zamuner, L. Kilbertus, and M. Weinhold, "Game-influenced methodology: Addressing child data attrition in language development research," *International Journal of Child-Computer Interaction*, vol. 14, pp. 15 – 22, 2017.
36. M. Dugas, K. Crowley, G. G. Gao, T. Xu, R. Agarwal, A. W. Kruglanski, and N. Steinle, "Individual differences in regulatory mode moderate the effectiveness of a pilot mHealth trial for diabetes management among older veterans," *PLoS One*, vol. 13, no. 3, p. e0192807, 2018.
37. A. L. Gremaud, L. J. Carr, J. E. Simmering, N. J. Evans, J. F. Cremer, A. M. Segre, L. A. Polgreen, and P. M. Polgreen, "Gamifying accelerometer use increases physical activity levels of sedentary office workers," *Journal of the American Heart Association*, vol. 7, no. 13, pp. 1–12, 2018.
38. J. Ryan, S. Edney, and C. Maher, "Engagement, compliance and retention with a gamified online social networking physical activity intervention," *Translational Behavioral Medicine*, vol. 7, no. 4, pp. 702–708, 2017.
39. Z. Zhao, A. Arya, A. Whitehead, G. Chan, and S. A. Etemad, "Keeping users engaged through feature updates: A long-term study of using wearable-based exergames," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, (New York, NY, USA), p. 1053–1064, Association for Computing Machinery, 2017.
40. L. C. Stanculescu, A. Bozzon, R.-J. Sips, and G.-J. Houben, "Work and play: An experiment in enterprise gamification," in *Proc. 19th CSCW*, (New York, NY, USA), pp. 346–358, ACM, 2016.
41. P. E. Anderson, T. Nash, and R. McCauley, "Facilitating programming success in data science courses through gamified scaffolding and learn2mine," in *Proceedings of the 2015 ACM Conference on Innovation and Technology in Computer Science Education*, (New York, NY, USA), pp. 99–104, ACM, 2015.
42. W. R. Boot, D. Souders, N. Charness, K. Blocker, N. Roque, and T. Vitale, "The Gamification of Cognitive Training: Older Adults' Perceptions of and Attitudes Toward Digital Game-Based Interventions," *Human Aspects of IT for the Aged Population. Design for Aging*, vol. 9754, pp. 290–300, 2016.
43. J. A. Cafazzo, M. Casselman, N. Hamming, D. K. Katzman, and M. R. Palmert, "Design of an mHealth app for the self-management of adolescent type 1 diabetes: A pilot study," *Journal of medical Internet research*, vol. 14, no. 3, p. E70, 2012.
44. R. M. Cunha Leite, D. Bastos Costa, H. Meijon Morêda Neto, and F. Araújo Durão, "Gamification technique for supporting transparency on construction sites: a case study," *Engineering, Construction and Architectural Management*, vol. 23, no. 6, pp. 801–822, 2016.
45. L. De-Marcos, A. Domínguez, J. Saenz-de Navarrete, and C. Pagés, "An empirical study comparing gamification and social networking on e-learning," *Computers & Education*, vol. 75, pp. 82–91, 2014.
46. B. Harrington and A. Chaudhry, "Tracademic: Improving participation and engagement in cs1/cs2 with gamified practicals," in *Proceedings of the 2017 ACM Conference on Innovation and Technology in Computer Science Education, ITiCSE '17*, (New York, NY, USA), p. 347–352, Association for Computing Machinery, 2017.
47. S. M. Kelders, M. Sommers-Spijkerman, and J. Goldberg, "Investigating the direct impact of a gamified versus nongamified well-being intervention: An exploratory exp.," *Journal of medical Internet research*, vol. 20, no. 7, p. e247, 2018.
48. M. Khalil, M. Ebner, and W. Admiraal, "How can gamification improve MOOC student engagement?," in *Proceedings of the European Conference on Game Based Learning*, (Sonning Common, UK), pp. 819–828, ACPI, 2017.

49. T. Lehtonen, T. Aho, E. Isohanni, and T. Mikkonen, “On the role of gamification and localization in an open online learning environment: Javala experiences,” in *Proceedings of the 15th Koli Calling Conference on Computing Education Research*, Koli Calling ’15, (New York, NY, USA), p. 50–59, Association for Computing Machinery, 2015.
50. J. Lumsden, A. Skinner, D. Coyle, N. Lawrence, and M. Munafò, “Attrition from web-based cognitive testing: A repeated measures comparison of gamification techniques,” *Journal of medical Internet research*, vol. 19, no. 11, pp. 1–18, 2017.
51. C. R. Nevin, A. O. Westfall, J. Martin Rodriguez, D. M. Dempsey, A. Cherrington, B. Roy, M. Patel, and J. H. Willig, “Gamification as a tool for enhancing graduate medical education,” *Postgraduate Medical Journal*, vol. 90, no. 1070, pp. 685–693, 2014.
52. M. V. Palacin-Silva, A. Knutas, M. A. Ferrario, J. Porras, J. Ikonen, and C. Chea, “The role of gamification in participatory environmental sensing: A study in the wild,” in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI ’18, (New York, NY, USA), p. 1–13, Association for Computing Machinery, 2018.
53. G. Pramana, B. Parmanto, J. Lomas, O. Lindhiem, P. C. Kendall, and J. Silk, “Using mobile health gamification to facilitate cognitive behavioral therapy skills practice in child anxiety treatment: Open clinical trial,” *JMIR serious games*, vol. 6, no. 2, pp. 1–15, 2018.
54. O. Zuckerman and A. Gal-Oz, “Deconstructing gamification: evaluating the effectiveness of continuous measurement, virtual rewards, and social comparison for promoting physical activity,” *Personal and Ubiquitous Computing*, vol. 18, no. 7, pp. 1705–1719, 2014.
55. A. B. Lewin, A. M. LaGreca, G. R. Geffken, L. B. Williams, D. C. Duke, E. A. Storch, and J. H. Silverstein, “Validity and reliability of an adolescent and parent rating scale of type 1 diabetes adherence behaviors: The self-care inventory (sci),” *Journal of Pediatric Psychology*, vol. 34, no. 9, pp. 999–1007, 2009.
56. ukkinstituutti.fi. Helsinki: Ministry of Education and Culture & Nuori Suomi ry, “Recommendation for the physical activity of school-aged children,” 2018.
57. J. Koivisto and J. Hamari, “Demographic differences in perceived benefits from gamification,” *Computers in Human Behavior*, vol. 35, pp. 179–188, June 2014.
58. K. R. Clark, “The effects of the flipped model of instruction on student engagement and performance in the secondary mathematics classroom,” *Journal of Educators Online*, vol. 12, no. 1, pp. 91–115, 2015.
59. K. Werbach and D. Hunter, *For the win: How game thinking can revolutionize your business*. Wharton Digital Press, 2012.
60. S. Deterding, “Eudaimonic design, or: Six invitations to rethink gamification,” in *Rethinking Gamification*, pp. 305–331, Lüneburg, Germany: Lüneburg: meson press 2014, 06 2014.
61. I. Bogost, “Why gamification is bullshit,” *The gameful world: Approaches, issues, applications*, vol. 65, pp. 65–79, 2015.