Comparing the Effects of Gamification to User Engagement in Stress Management Application

Nurhafiyah Hazwani Haris Fadzillah¹*, Nur Zuraifah Syazrah Othman¹, Masitah Ghazali¹, Nor Azman Ismail¹

¹ UTM VicubeLab, Faculty of Computing, Universiti Teknologi Malaysia, 81310 Johor Bahru, Johor, Malaysia

ARTICLE INFO

ABSTRACT

The increasing accessibility of mobile technology makes it easier for users to use health related applications to combat health issues. In addition to this, gamification has been widely used to increase the engagement between user and application. This study provides an insight into how gamification affects user engagement when using stress management application. Two versions of stress management application were developed, one with gamification element, and the other one without. These two applications were then evaluated by 20 participants using the UES-SF questionnaire, where data on the engagement level of the participants were collected and analysed. The results showed that all participants (100%) prefer the application with gamification to help them manage their stress level better. Statistical tests done revealed that the stress management application with gamification was also found to be significantly more engaging compared to stress management application without gamification (given p ≤ .05). A Wilcoxon Signed-Ranks test indicated that gamified stress management application was rated more favourably than the non-gamified stress management application, with Z score = -3.92, and p = 0.00. The results demonstrated the significance of gamification in increasing user engagement among users when using stress management application.

Keywords: Gamification; user engagement; stress management application

1. Introduction

Nowadays, with the increasing accessibility of mobile technology, gamification has been widely used to increase engagement between user and application. Gamification refers to the utilization of game designs in non-game settings [11]. Previous researchers have found that gamification element brings a lot of positive impact on user engagement due to similar excitement in playing games [21]. Gamification has the potential to boost players' motivation, engagement with the games, and learning experiences, according to data from existing studies [1]. It is a way to motivate users to constantly use an application, sometimes in the long term, as the user sees the benefit in terms of rewards. Since managing health is also a long-term process, incorporating gamification elements in

* Corresponding author.
E-mail address: hazwani1998@utm.my

https://doi.org/10.37934/araset.30.1.290302
healthcare application might be a suitable motivation factor to ensure user’s commitment to continue to use the app. A good application will be able to collect the information entered by the user and perform calculations and provide a test result. Furthermore, if they have a health problem, the consistency of the healthcare process tracking serves as a point of reference for them. This study explores the idea of incorporating gamification to enhance user’s engagement when using health management application. Statistical findings from this research can provide an insight of the relationship between gamification and user engagement, and how gamification can be an added motivation factor for users to manage their health better.

2. Related Work

In this section, the related domains of the study will be discussed, including reviews on principles and techniques used in this study.

2.1 Mental Health and Stress Management

According to the World Health Organization (WHO) [41], “mental health is a state of well-being in which individual recognizes his or her potential, can deal with the normal demands of life, can work effectively, and can contribute to his or her community”. WHO emphasizes that mental health is more than merely the absence of mental illnesses or impairments. Stress management is thus vital because stress has a powerful impact on many aspects of daily life—not only can it affect mood, energy level, relationships [30], and work performance [24], but it can also cause and aggravate a wide range of health conditions. A form of stress management is by monitoring the users or patients stress levels. One established method is by using the Perceived Stress Scale (PSS), a tool that measures the user’s own perception of stress [6]. Respondents are asked ten questions on how unexpected, unmanageable, and overloaded their lives had become in the last month. The perceived stress levels were rated on a 5-point Likert scale (range: 0 = "never" to 4 = "very often"). The PSS score ranges from 0 to 40, with higher values indicating more stress perception. Low stress is defined as a score of 0–13, while moderate stress is defined as a score of 14–26. A high amount of felt stress is indicated by a score of 27–40.

2.2 Gamification

Gamification, is by definition a goal-oriented activity [11]. Goals can be implemented in a variety of ways. It can be communicated directly, for example as quests, or they might be given implicitly as outcomes that can be pursued [34]. Game-like design aspects can be used in gamification to create clear goals. Examples include boss encounters, certificates, collections, exploratory tasks, learning, quests, unlockable or uncommon items, and obtaining access to additional features [34]. Six gamification principles have been proposed by Eisingerich et al., [13] to encourage user engagement. Using both inductive and deductive reasoning, the study compiled a list of second-order themes based on the inferred findings and data were collected and coded until it reached category saturation, at which point no unused experiences about the construct or their connections emerged [15]. The themes are then refined by integrating relevant literature and holding cooperative conversations to reflect the gamification concepts and fundamental instruments that surfaced more than once within the information [16]. The data analysis indicated six gamification principles—social interaction, sense of control, goals, progress tracking, rewards, and prompts—that serve as affordances to drive individuals to take action. These principles are shown in Table 1.
### Table 1

<table>
<thead>
<tr>
<th>Principle</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social Interaction</td>
<td>Application’s capacity to assist clients remain in touch with friends and others</td>
</tr>
<tr>
<td>Sense of Control</td>
<td>Application’s capacity to form clients feel they can control their possess fates and be dependable</td>
</tr>
<tr>
<td>Progress Tracking</td>
<td>Application’s capacity to record, report, and delineate customers' advances and competence</td>
</tr>
<tr>
<td>Rewards</td>
<td>Application’s capacity to offer modern materials when clients accomplish particular objectives</td>
</tr>
<tr>
<td>Prompt</td>
<td>Application’s capacity to remind clients of their commitment and empower them to require activity through alarms</td>
</tr>
<tr>
<td>Goals</td>
<td>Application’s capacity to indicate and set achievable goals</td>
</tr>
</tbody>
</table>

### 2.3 Gamification in Healthcare Application

Gameplay attracts and influences our attention, drawing on our inherent talents, thrilling us, and inspiring us to become more competitive by pursuing ever stronger and more successful abilities. Thus, many suggests that it can be used to change health behaviour. A recent research has looked at Pokémon Go’s effect in cultivating good exercise habit [18]. According to the data collected, the average number of steps taken by a Pokémon Go player climbed from 4256 to 5123 in the first week of play. The Gamification factor in healthcare can also track medication and the treatment of chronic conditions. Several applications for smartphones have been designed to enable patients to take their medication on time. Users set the hours when they will take medicine, and the app alerts them. This also provides prescription knowledge and warnings about reactions and side effects of the drug. Patients earn points for gift cards or charitable donations based on a record when taking medication correctly. Another example is the Bayer blood glucose meter, connected to the Nintendo DS gaming device, which helps regulate user’s diabetes by providing daily monitoring of blood glucose [40]. Users are able to unlock new game levels and options, as their points accumulate. There are also leader boards for users who have earned the most points, as well as video games and online community.

### 2.4 Measuring User Engagement

There are several methods for measuring user engagement. The questionnaire method has been found to be the most widely used techniques for evaluating engagement [12]. They are simple to use and can support extensive research [3, 26, 31, 32, 35]. The User Engagement Scale (UES) is one of the most comprehensive attempts in the literature to understand user involvement [26]. This questionnaire was inspired by an initial conceptual framework for user engagement based on a study of 17 interviews with participants about their experiences with online purchasing, searching, video games, and distance learning [25]. This concept defines engagement as a four-stage process consisting of a point of engagement, a period of sustained engagement, disengagement, and re-engagement within the context of an “effective, efficient, and satisfying” framework. An initial evaluation study of 440 participants’ responses to an online shopping experience resulted in the reduction of the scale to 33 items distributed across six major factors: Focused Attention, Perceived Usability, Aesthetics, Endurability, Novelty, and Felt Involvement [26]. Subsequent research has refined the model into four factors [26, 37]. UES, and its short form version UES-SF (its factors shown in Table 2) has since been applied by many researchers looking to measure engagement. Once data is collected, it can be analysed using suitable statistical techniques.
### Table 2
Factors in the User Engagement Scale (UES-SF)

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>FA</td>
<td>Focused attention</td>
</tr>
<tr>
<td>PU</td>
<td>Perceived usability</td>
</tr>
<tr>
<td>AE</td>
<td>Aesthetic appeal</td>
</tr>
<tr>
<td>RW</td>
<td>Reward (Endurability, Novelty and Felt involvement)</td>
</tr>
</tbody>
</table>

#### 3. Methodology

Previous works suggested that gamification can help increase user engagement when using an application. This study aims to further contribute to this knowledge by evaluating the effect of gamification in user engagement when applying it to a stress management application. Taking inspiration from the study done by Cechetti et al., [5] two version of a stress management application were developed, one with gamification elements (by applying the six gamification principles proposed by Eisingerich et al., [13]), and the other one without any gamification element. The two application were then used to evaluate user engagement by collecting user’s feedback on both apps. Evaluation was done in form of A/B testing, and statistical analysis is then used to determine which version provides a higher level of user engagement. The data were collected using the User Engagement Scale - Short Form (UES-SF) questionnaire proposed by O’Brien and Toms [26]. The quantitative data collected through the UES-SF questionnaires were then analysed using the software Statistical Package for Social Science (SPSS). Depending on the suitability (normality) of the data, statistical techniques which involved both descriptive and inferential can be used to analyse the data. In this study, the Wilcoxon signed-rank test were used to compare the results for participants’ engagement level when using different applications (gamified vs non-gamified) to accomplish the given user tasks.

#### 4. Developed Application

Both applications were developed using Flutter and Firebase. Application A was developed without any gamification principles. Figure 1 shows screenshots of the main menu and stress test while Figure 2 shows a personal diary and profile. The flow of the application is as follows. Once logged in, the user is able to either view infographics on stress managements, create personal diary or take the stress test. The stress test applied in this application is based on the Perceived Stress Scale (PSS) test discussed in Section 2.1. Users will be able to answer the test anytime they want, and can check their stress test result history from the profile page. The infographics are health-related so the users can be more informed on managing stress. The personal diary allows the user to write what they want any time of the day.
Application B on the other hand, was developed based on the six gamification principles proposed by Eisingerich et al., [13]. These principles were discussed in Section 2.2. All six principles were implemented in Application B. The main gamification principle implemented is the sense of control as this application help user to track their stress. The flow of the application is as follows. Once logged in (Figure 3), the main menu shows the main avatar purchased by the user using points located and the daily goals for the user. Users can click on the daily goals and it will navigate the user to the daily goals page. The users can also take a PSS stress test where on this page, gamification principles reward and prompt were implemented. The stress test features offered users to answer the stress tests to know their stress level and received points that can be used to buy avatars for the homepage. A prompt message will appear before the user take the stress test to explain the instructions, and another one appears after the user takes the test to inform their stress level in detail. Next, users can access the personal diary page and can then continue to the avatar page, which consists of change avatar features (Figure 4). The avatars can be bought using the points received from the stress test. There are many avatars offered to the user which acts as rewards to the user. There is also a prompt message that appears when the user changes the avatar successfully.
Next, the user can also access the daily goals page as shown in Figure 5 below. As the name suggests, this feature implemented the goals principle. Since these are daily goals, the goals will be renewed every day. The daily goals offered users deep breathing exercises to release their stress in a minute. When users completed each goal, they will receive a voucher. This is a reward for the user after release their stress by using this application.

Meanwhile Figure 6 shows the profile page and result history for both stress tests and daily exercise. The application offers users to share their results on social media such as Facebook and Instagram. These features implemented the social interaction and progress tracking principles.
5. Experiment Design

Both Application A and B discussed in the previous section were used to evaluate the effects of gamification on user engagement. Participants were asked to use each application by completing several user tasks given to them. After completing the tasks, the participants filled in UES-SF questionnaire on how engaged they feel when using the application. All tests were performed on an android phone with the lowest Android version is Android 4.4 KitKat (API 19). All interaction with each application is through the phone’s touch screen. The following subsection details the experiment design.

5.1 Measures

For both pilot test and main experiment, the UES-SF questionnaire were given after the participants have completed the tasks for each application. There are 12 items from four factors in the UES-SF. Participants gave a score of 1 to 5 to each item, where 1 stands for strongly disagree and 5 stands for strongly agree. Table 3 shows the 12 items used in this experiment.

<table>
<thead>
<tr>
<th>Table 3</th>
<th>User Engagement Scale - Short Form items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item</td>
<td>Question</td>
</tr>
<tr>
<td>FA-S.1</td>
<td>I lost myself in this experience</td>
</tr>
<tr>
<td>FA-S.2</td>
<td>The time I spent using this app just slipped away</td>
</tr>
<tr>
<td>FA-S.3</td>
<td>I was absorbed in this experience</td>
</tr>
<tr>
<td>PU-S.1</td>
<td>I felt frustrated when using this app</td>
</tr>
<tr>
<td>PU-S.2</td>
<td>I found this app confusing to use</td>
</tr>
<tr>
<td>PU-S.3</td>
<td>Using this app was taxing</td>
</tr>
<tr>
<td>AE-S.1</td>
<td>This application was attractive</td>
</tr>
<tr>
<td>AE-S.2</td>
<td>This application was aesthetically appealing</td>
</tr>
<tr>
<td>AE-S.3</td>
<td>This application appealed to my senses</td>
</tr>
<tr>
<td>RW-S.1</td>
<td>Using this application was worthwhile</td>
</tr>
<tr>
<td>RW-S.2</td>
<td>My experience was rewarding</td>
</tr>
<tr>
<td>RW-S.3</td>
<td>I felt interested in this experience</td>
</tr>
</tbody>
</table>
5.2 Test hypothesis

The engagement level of the user when using the non-gamified stress management application (Application A) is tested against the gamified stress management application (Application B). The evaluation will measure four engagement factors in UES-SF, Focused Attention (FA), Perceived Usability (PU), Aesthetic Appeal (AE) and Reward Factor (RW), as discussed in the previous section. The null hypothesis for this study is,

$H_0$: There is no significant difference in user engagement between non-gamified stress management application and gamified stress management application.

5.3 Participants

The participants acquired for this study fulfils the user target group for the developed application, which are university students. Participants responded voluntarily through advertisement in student chat group. In total, twenty-two (22) participants participated; two (2) participants for the pilot test, and twenty (20) participants for the main test. The participants age is between 22 – 25 with the average age of 23. All of them are from a single nationality (Malaysian).

5.4 Setup and Procedure

There were two tests done, pilot test and the main experiment test. The setup and procedures for both tests are discussed in the following subsections.

5.4.1 Pilot test

The pilot test was done two days before the main experiment, and took place at a student college. Two (2) undergraduate students (henceforth referred as Participant 1 and Participant 2) volunteered. Before the experiment began, each participant had to fill out a consent form and a pre-test form. Each participant was presented with the installed app and given 20 minutes to explore the app on their own. The order of the tested app differs for each participant. Participant 1 tested Application A first, followed by Application B. Meanwhile, Participant 2 tested Application B first, followed by Application A. This is to reduce bias and counteract learning effects. After the participant completed all tasks, they were asked to fill in the UES-SF form. Based on the interview result of the pilot test, it was found that 20 minutes is enough for the user to discover both applications before starting to perform the task required. In addition, it was discovered that several questions from the interview needed to be refined. These findings were used to improve on the setup and design for the main experiment.

5.4.2 Main experiment

The main experiment is a within-subject test in which each participant uses both application A and application B. 20 participants were acquired. The participants were divided into two groups at random. As with the pilot test, the first group started by using application A and followed by application B. The second group started with application B and then moved on to application A. Before the experiment began, each participant had to fill out a consent form and a pre-test form. After that, each user has 20 minutes to explore the first application. Following that, each participant
was expected to execute the tasks assigned to them. After the participant completed all of the tasks, the UES-SF was distributed to the participants for them to fill in.

6. Results Analysis

At the end of the experiment, participants also answered an open-ended question (in addition to the UES-SF) of which application they feel will help them manage their stress level better. All 20 participants felt that Application B (with gamification) will help them manage their stress level better. The following describes results from the UES-SF questionnaire.

6.1 Overall UES-SF score

The UES-SF questionnaire's responses are graded on a five-point Likert scale (1=strongly disagree, 5=strongly agree). Figure 7 shows the comparison of the mean for each factor in both applications. The overall engagement score showed that the overall mean for application A is 2.19 while the overall mean for application B is 4.2. Application B has a higher mean in all factors and in the overall mean score than application A. The highest contrast of results between the two application is for the Reward Factor (RW) with application B achieving a high mean of 4.47 compared to Application A which only achieved a mean of 2.15. As explained by O’Brien and Toms [26], RW is users’ sentiments of being drawn in, interested, and having fun amid the interaction and users’ level of intrigued within the assignment and interest evoked by the framework and its substance. Application B has more rewarding features compared to application A. One of the gamification elements implemented in application B is reward, in which a user receives rewards or prizes for completing a task in the application.

Breaking down the details further on the data scoring for Application A, FA has a total mean of 41.66 with average of 2.083, and a standard deviation of 0.7478. The statement “I lost myself in this experience” is neither agreed upon nor disagreed upon by the majority of participants. PU has a total mean of 46.32, with average of 2.316, and a standard deviation of 0.7676. The majority of participants agree that application A is difficult to use, and they are frustrated and exhausted as a result. The total mean for AE is 43.32, the average is 2.166, and the standard deviation is 0.6152. The majority of participants disagree that application A is attractive and appealing aesthetically and neither agree nor disagree that it appeals to their senses. RW has a 443.01 total mean, with average 2.150, and a 0.7056 standard deviation. The vast majority of participants totally disagree application A is worthwhile or intriguing, and did not feel their experiences were rewarding.
Meanwhile for Application B, FA has a total mean of 74.00 with average of 3.700 and a standard deviation of 0.6109. The majority of participants agree that they were immersed in their experiences and that they assimilated information from application B. The standard deviation for PU is 0.5988, and a total mean of 85.35 with average of 4.2675. The majority of participants disagree that application B is difficult to use. They are completely opposed to the idea that application B is taxing. The total mean for AE is 87.01, the average is 4.3505 and the standard deviation is 0.4768. The majority of participants totally agree that application B is attractive. Meanwhile, the majority of the participants strongly agree that application B is aesthetically appealing and appeals to their senses. RW has an 85.35 total mean, with average of 4.2675 mean, and a 0.3658 standard deviation. The vast majority of participants agree that application B is worthwhile and that their experiences were rewarding.

### 6.2 Wilcoxon Signed-Rank Test Result

The previous subsections have discussed the results of the experiment that suggested how gamification does effect user engagement when using stress management applications. To further support this finding, statistical analyses were done to see whether there is significant difference between the two versions of applications (gamified vs non gamified) in terms of user engagement. Results from the Shapiro Wilk test showed the data are not normally distributed and therefore were analysed using Wilcoxon signed-rank test, (the equivalent of non-parametric paired sample t-test). Specifically, the null hypothesis stated in Section 5.2 is tested, which is \( H_0: \) There is no significant difference in user engagement between non-gamified stress management application and gamified stress management application. The results from the Wilcoxon signed-rank test indicated a significant difference in overall score, \( z = -3.924, p < .05. \) The results indicate that application B (with gamification) has the higher level of user engagement and received significantly more favourable rankings than application A (without gamification).

Table 4 shows the summary of the experiment results depicting higher mean for Application B (with gamification) in all four factors: FA (M=3.70, SD=0.61), PU (M=4.27, SD=0.60), AE (M=4.35, SD=0.48) and RF (M=4.47, SD=0.37). In addition, the Overall Engagement factor for Application B also shows a higher mean (M=4.20, SD=0.315). Meanwhile for Wilcoxon Signed-rank test results, the Z value for FA, PU, AE, RF and overall engagement score are -3.786, -3.792, -3.929, -3.929, and -3.924 respectively, and the p-value for all comparisons is less than 0.05. Conclusively, the results shown suggest a statistically significant difference between using Application A (without gamification) and Application B (with gamification) in terms of user engagement (given \( p < .05 \)). Thus, the null hypothesis of this study is rejected, meaning that there is a significant difference between the two application in terms of user engagement. This result supports and strengthen the findings that application with gamification had better levels of user engagement than application without gamification.

<table>
<thead>
<tr>
<th>Table 4</th>
<th>Summary of the result</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor</td>
<td>Mean Score (standard deviation)</td>
</tr>
<tr>
<td>Design</td>
<td>Application A (non-gamified)</td>
</tr>
<tr>
<td>Focused Attention (FA)</td>
<td>2.08 (0.75)</td>
</tr>
<tr>
<td>Perceived Usability (PU)</td>
<td>2.32 (0.77)</td>
</tr>
<tr>
<td>Aesthetic Appeal (AE)</td>
<td>2.17 (0.62)</td>
</tr>
<tr>
<td>Reward Factor (RF)</td>
<td>2.15 (0.71)</td>
</tr>
<tr>
<td>Overall Score</td>
<td>2.19 (0.512)</td>
</tr>
</tbody>
</table>
7. Limitation and Future Work

There are a number of limitations in this study that can be improved and added for future work. In terms of the application developed in this study, an improvement to the gamified application can be done by adding more stress test modules, for example, the Depression Anxiety Stress Scale (DASS) to gain a better understanding of the emotions, reactions, and moods [20]. In terms of the experiment done in this study, due to time constraint only a small sample was able to be acquired. Moreover, only university students were recruited as participants for both pilot study and main study, meaning that the outlook given could have a potential bias due to their academic background. Future work could verify the results with more suitable sized data sets from different backgrounds.

Acknowledgement

This work was supported by the UTM Encouragement Research grant, Universiti Teknologi Malaysia, Malaysia (cost centre code: Q.J130000.2651.17J61).

References


