Understanding E-Mental Health for People with Depression: An Evaluation Study

Kim Janine Blankenhagel[1](✉), Johannes Werner[1], Gwendolyn Mayer[2], Jobst-Hendrik Schultz[2], and Rüdiger Zarnekow[1]

[1] Technical University Berlin, Straße des 17. Juni 135, 10623 Berlin, Germany
k.blankenhagel@tu-berlin.de

[2] Heidelberg University Hospital, Im Neuenheimer Feld, 69120 Heidelberg, Germany

Abstract. Depression is widespread and, despite a wide range of treatment options, causes considerable suffering and disease burden. Digital health interventions, including self-monitoring and self-management, are becoming increasingly important to offer e-mental health treatment and to support the recovery of people affected. SELFPASS is such an application designed for the individual therapy of patients suffering from depression. To gain more insights, this study aims to examine e-mental health treatment using the example of SELFPASS with two groups: healthy people and patients suffering from depression. The analysis includes the measurement of the constructs Usability, Trust, Task-Technology Fit, Attitude and Intention-to-use, the causal relationships between them and the differences between healthy and depressive participants as well as differences between participants' evaluations at the beginning and at the end of the usage period. The results show that the Usability has the biggest influence on the Attitude and the Intention-to-use. Moreover, the study reveals clear differences between healthy and depressive participants and indicates the need for more efforts to improve compliance.

Keywords: eHealth · Digital mental health · Depression · Individual therapy · Self-management · Structural equation modelling · ANOVA

1 Introduction

Depression is a severe and widespread disease with considerable effects on people's well-being and quality of life [1]. Although evidence-based treatments such as psychotherapy are available for depressive disorders, a significant portion of people afflicted with such disorders do not receive treatment [1] or wait a long time for treatment to begin [2]. At the same time, a large portion of the world's population uses the Internet, with much of this usage being focused on health [3]. Consequently, self-monitoring and self-management are becoming increasingly important [1], and digital health interventions have proven to be a promising way of supporting people with depression [4]. These can be an effective complement to personal psychotherapy or pharmacological treatment and are particularly suitable for people who have insufficient access to psychological

J. Ye et al. (Eds.): MobiHealth 2020, LNICST 362, pp. 34–51, 2021.
https://doi.org/10.1007/978-3-030-70569-5_3

treatment or do not wish to get in personal contact with a psychotherapist [5]. To seize this potential, a web-based therapeutic platform for patient-individualized therapy and self-management called SELFPASS has been developed for people suffering from depression. So far, web intervention research has mainly focused on the effectiveness of therapy and changes in symptom severity [6], but little research (quantitative or qualitative) has been conducted on gathering more insights about those systems in terms of their acceptance or attitudes towards them. To close this gap, we analyze the underlying factors that affect people's decision to use e-mental health applications to manage depression. The aim is to examine the Usability, Trust, Task-Technology Fit, Attitude and Intention-to-use of those applications using the example of SELFPASS. Of particular interest are the causal relationships between the above-mentioned constructs and the differences between healthy and depressive participants as well as differences between the beginning and the end of the usage period.

2 Background

The background section consists of a practical and a theoretical part. Section *SELFPASS* outlines the SELFPASS application, and section *Theoretical Background* presents the theoretical foundation of the study.

2.1 SELFPASS

SELFPASS is a therapeutic platform designed for the individual therapy and self-management of patients suffering from depression. It is based on a combination of algorithms for a daily self-assessment and analysis of the patient's biosignal data and environmental information. SELFPASS enables therapy by offering self-assessment of the severity of the patient's mental distress and by suggesting practical steps for self-management. This takes into account the integrated biosignal data (for example, heart rate) and current environmental information (for example, weather). Depending on his/her individual situation, the patient receives individualized guidance for self-management and therapeutic interventions. The structure of the therapy sessions varies according to the degree of severity indicated. SELFPASS is designed for depressive people, who have received a diagnosis by a medical institution and are now waiting for personal therapy. We conducted this study with a SELFPASS prototype, which did not contain a link to biosignal data or environmental information. Instead, the participants in the study were able to use the self-assessment and various interventions (diary, activity plan, relaxation exercises and so on).

2.2 Theoretical Background

We have focused on the measurement of five constructs (Usability, Trust, Task-Technology Fit, Attitude, and Intention-to-use), which are of crucial importance for the evaluation of eHealth technologies. The term "Usability" describes the degree to which a product can be used by a particular user in a certain context with effectiveness, efficiency and satisfaction [7]. The usability of technology plays a significant role in increasing its acceptance and creating user loyalty, which is especially important in the healthcare sector [8]. Usability factors remain one of the major obstacles to the adoption of health technologies, emphasizing the necessity of usability evaluations [9]. Trust "indicates a positive belief about the perceived reliability of, dependability of, and confidence in a person, object or process" [10]. In addition, trust in technology is a key factor in establishing a satisfactory relationship between the user and product in any interactive situation [11]. Trust is a component of Trust and Mistrust, where Mistrust is the complement of Trust [12]. We decided to integrate the positive part of Trust into our study and excluded Mistrust. Goodhue and Thompson proposed the Task-Technology Fit theory to highlight the importance of an adequate correspondence between the characteristics of technologies and user tasks for achieving the desired effects in terms of individual performance [13]. Therefore, the technology must be a good fit with the tasks it supports in order to have a positive impact. We have used the two very established constructs Attitude and Intention-to-use, known from the Technology Acceptance Model (TAM) [14], to assess technology usage.

The five constructs presented above are not only an evaluation standard in themselves but are also linked to each other in certain relationships. TAM, as a widespread innovation adoption model, explains the use of new technology and outlines the Attitude construct having a positive effect on the Intention-to-use construct [14]. Furthermore, the literature shows that the Usability, Trust and Task-Technology Fit constructs have a positive effect on the Attitude and Intention-to-use constructs [15–17].

Due to the high importance of the above-mentioned constructs for the success of eHealth interventions, it is crucial for research and practice to thoroughly analyze them with respect to new platforms such as SELFPASS.

3 Method

Between February and May 2019, study participants were recruited in Berlin and Heidelberg (Germany), and the survey took place in the same period. Participants were acquired offline at the University Hospital in Heidelberg and at the Technical University in Berlin through personal information sessions pertaining to SELFPASS and to participation in the study. Furthermore, the study was made public through online forums and in order to attract more participants we used snowball sampling. Those deemed to be eligible included adult people with German language skills and, with respect to the participants in Berlin, those who had access to the Internet and owned an Internet-enabled device. In Heidelberg, patients were provided with an Internet connection and tablets to access SELFPASS. For each participant, the survey took place over a period of five consecutive days, during which the participants used SELFPASS daily for a period of around 30 min. Each day, participants were asked to log in, complete the self-assessment and try at least one intervention. They were also encouraged to test SELFPASS critically by skipping some of the interventions or stopping them altogether and noticing anything conspicuous as a result. At the end of the first day (point in time - T1) and the fifth day (point in time - T2), the participants completed a questionnaire to assess SELFPASS. All participants participated on a voluntary basis, and the procedures of the study were approved by the Ethics Committee of the Heidelberg University. Furthermore, the study is listed in the registry of clinical trials.

3.1 Research Model and Hypotheses

We have conducted an empirical study on SELFPASS in order to first examine the causal relationships between the following five constructs Usability (USA), Trust (TR), Task-Technology Fit (TTF), Attitude (ATT) and Intention-to-use (INT).

The hypotheses regarding the causal relationships are derived from the literature and shown in Table 1:

Table 1. Research hypotheses

Research hypotheses	Path
H1: Usability relates positively to Attitude	USA → ATT
H2: Usability relates positively to Intention-to-use	USA → INT
H3: Trust relates positively to Attitude	TR → ATT
H4: Trust relates positively to Intention-to-use	TR → INT
H5: Task-Technology Fit relates positively to Attitude	TTF → ATT
H6: Task-Technology Fit relates positively to Intention-to-use	TTF → INT
H7: Attitude relates positively to Intention-to-use	ATT → INT

Second, we determined differences between the scores of the constructs on the first and on the last day of the trial period (T1 and T2) as well as differences between the scores obtained by healthy participants and by those suffering from depression.

3.2 Questionnaire Design and Data Collection

The survey scheduled for T1 comprised relevant socio-demographic and demographic questions as well as questions related to the participants' experiences with digital technologies. Depression symptoms were measured using the Patient Health Questionnaire 9 (PHQ-9) [18] to verify whether the participant was suffering from depression. All participants with low PHQ-9-scores (smaller than ten) were classified as not depressive while the rest (PHQ-9 greater than or equal to 10) were classified into the comparison group whose members suffered from depression. Additionally, the five constructs Usability, Trust, Task-Technology Fit, Attitude and Intention-to-use were first assessed at T1 and a second time at T2. All constructs have a reflective measurement, because the measured variables do not construct their respective latent variables, instead they measure or manifest them. All used instruments had been previously validated. Usability was evaluated by the SUS (System Usability Scale) [19], which consists of ten items answered on a 5-point Likert rating scale, ranging from "strongly disagree" to "strongly agree". Among them, five are positive statements, and the rest are negative. SUS can provide a single score that ranges from 0 to 100, with higher scores denoting higher usability (scores were manipulated to a 0 to 4 rating and multiplied by 2.5). For the measurement of the Trust construct [12] the 7 Items of Jian et al. were used. They were arranged into a 7-point Likert rating scale and consisted of positive statements. The Task-Technology

Fit [20] comprised of 8 items answered on a 7-point scale, just like Attitude [21] (5 items) and Intention-to-use [22] (3 Items). The answer alternatives to the questions were all formed with unweighted scores. The questionnaires were delivered in German after being translated from the English original. In order to ensure that the content did not lose its original meaning, one of the other authors translated it back from German into English and compared it with the original.

Figure 1 summarizes the methodological procedure.

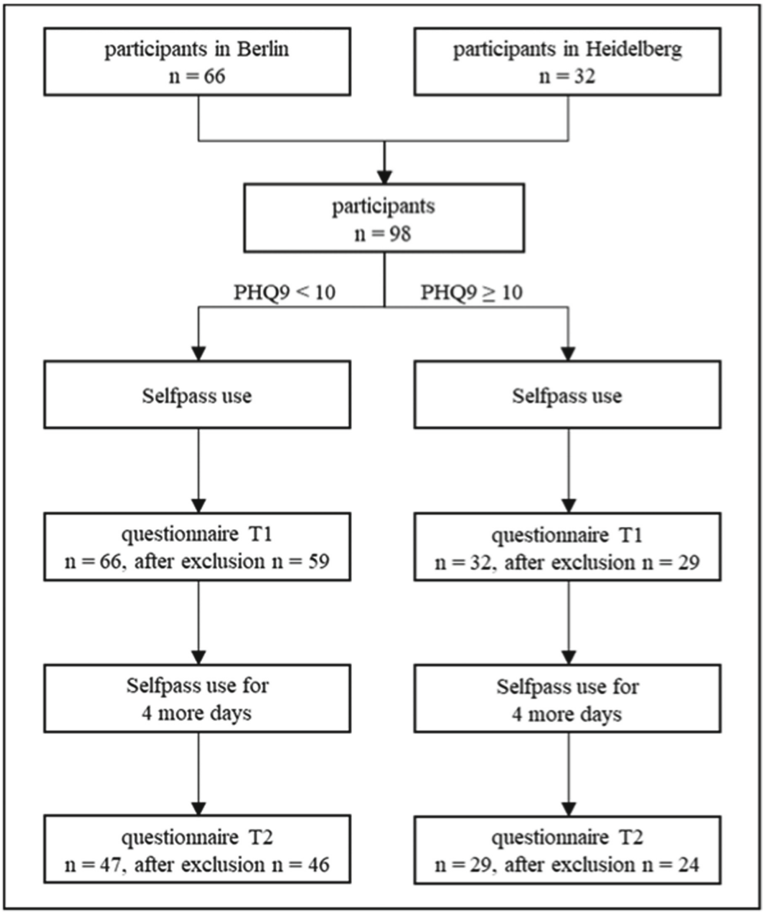

Fig. 1. Methodological procedure

The online questionnaire service SoSciSurvey was used to create and distribute the questionnaires of the study in Berlin. The participants received an email with a web link to the survey. At the clinic in Heidelberg, participants received a paper-pencil-version of the questionnaire, which was handed out to them by the investigator-in-charge. They received 50 Euro for completing the questionnaires. Each questionnaire was anonymous and identified by a unique identification number. A cover letter presenting the study's objectives and a brief overview of the key characteristics of SELFPASS was attached to the questionnaire. The collected data were organized in Microsoft Excel. We excluded questionnaires with incomplete information (more than 20% missing data) and those with an obviously distorted response behavior. In case of missing values in the underlying sample after the exclusion, we applied the method of medium value replacement.

3.3 Statistical Analyses

The data were analyzed in three different ways: descriptive statistics, structural equation modelling (SEM) and analysis of variance (ANOVA) with repeated measures. Descriptive analyses were used for obtaining the summary statistics of all measures and for the study of general characteristics.

We applied partial least squares structural equation modeling (PLS-SEM) to examine the causal relationships between the individual constructs. This approach was deemed suitable due to the complexity of the model and the high number of constructs and indicators involved. SmartPLS 3 was used to validate the measures and to test the research hypotheses. The quality measures factor loadings, composite reliability, displayed average variance and heterotrait-monotrait (HTMT) ratio were used as a basis for the evaluation of the reflective measurement model. To assess the structural model, we used R2, path coefficients significance and the effect size.

During the second phase ANOVA was used to compare changes in the constructs at T1 and T2 as well as between the healthy and depressive participants. Mean scores were calculated for all subscales, and the significance level for the tests was alpha = 0.05. Before performing data analysis, we used the Shapiro-Wilk test to assess normality and calculated the Cronbach alpha coefficients to assess the internal consistency of the theoretical constructs. Therefore, SPSS Version 25 was used.

4 Results

A total of 98 participants completed the questionnaire at T1, and 76 completed it at T2. 66 were classified as having no symptoms of depression and 32 as suffering from depression. The demographic data of the participants in this study largely corresponded to the demographics of the population that uses health apps (more females, young and with high education levels) [23]. Therefore, we conclude that we have a representative and relevant sample for this study in terms of early adopters of eHealth applications, but not with respect to the general population. The resulting samples formed the basis for subsequent statistical analysis. Table 2 shows the demographic statistics of the sample, subdivided into healthy participants and those with depression. These include the characteristics of all participants whose questionnaire responses at T1 and/or T2 were included in the analysis.

Table 2. Demographics

		Healthy participants n (%)	Participants suffering from depression n (%)
Gender	Male	31 (51)	11 (37)
	Female	26 (43)	19 (63)
	Not specified	4 (7)	0 (0)
Age	<25	18 (30)	18 (60)
	25–35	29 (48)	6 (20)
	35–45	0 (0)	2 (7)
	>45	8 (13)	4 (13)
	Not specified	6 (10)	0 (0)
Marital status	Single	45 (74)	21 (70)
	Married	10 (16)	4 (13)
	Separated/divorced/widowed	2 (3)	5 (17)
	Not specified	4 (7)	0 (0)
Highest degree	No/lower education	6 (10)	11 (37)
	Moderate/high education	51 (84)	18 (60)
	Not specified	4 (6)	1 (3)
Computer skills	Sufficient	2 (3)	2 (7)
	Moderate	3 (5)	4 (13)
	Good	25 (41)	17 (57)
	Excellent	30 (49)	7 (23)
	Not specified	1 (2)	0 (0)
Job situation	Self-employed	2 (3)	0 (0)
	Apprentice	0 (0)	1 (3)
	University/school	38 (62)	15 (50)
	Employee	16 (26)	11 (37)
	Unemployed	1 (2)	2 (7)
	Pensioners	0 (0)	0 (0)
	Other/not specified	4 (7)	1 (3)

4.1 Structural Equation Modelling – Measurement Model

In order to perform structural equation modelling, we started with a validation of our measurement model. First, we examined the factor loadings and eliminated items if their factor loadings were smaller than 0.7 and if eliminating the item resulted in an increase in the internal consistency reliability [24]. This method led to an elimination of a total of 6 items (all eliminated items pertained to Usability). Thereafter, the considered items had values greater than the minimum value of 0.4 (smallest value being 0.63) and were regarded suitable [24]. Subsequently, we assessed the construct reliability by determining the composite reliability. A construct reliability greater than 0.7 was deemed an acceptable reliability coefficient [25], and Table 3 shows that all the constructs met this criterion and demonstrated their internal consistency. All constructs showed an average variance extracted above 0.5, meaning that on average, the construct explains more than 50% of the variance of its indicators [25]. In discriminant analysis, the results of the HTMT ratio met the discriminatory criterion (being below 0.9) [26]. Thus, the measurement model had acceptable reliability and convergent validity, leading to a viable structural analysis of the model.

Table 3. Validation of the measurement model

Construct	Number of items	Composite reliability	Average variance extracted
USA	4	0.8	0.5
TR	7	0.9	0.6
TTF	8	0.9	0.8
ATT	5	0.9	0.7
INT	3	0.9	0.9

4.2 Structural Equation Modelling – Structural Model Assessment

We estimated the structural model paths and tested the research hypotheses with the entire sample (at T1 and T2, all participants) to identify the main determinants in the usage of SELFPASS. The multi-group analysis did not show significant differences between the causal relationships of T1 and T2 and revealed a significant difference between healthy and depressive participants in only one causal relationship (TTF → INT). All other relationships showed no significant differences. Therefore, it is reasonable to calculate a structural equation model based on all subgroups.

The evaluation of the structural model included the execution of SmartPLS under default settings with 5.000 samples, with a bootstrap of 5.000 resampling iterations and with mean replacement of missing data. All constructs had variance inflation factor (VIF) values less than 5, indicating that there was no multicollinearity problem.

The PLS-SEM model and its loadings are depicted in Fig. 2. The value in parentheses is the p value, the result of calculating the significance of quality to success relationships using the bootstrapping approach.

The model explains 44% of variance for the Attitude construct and 29% of variance for the Intention-to-use construct. The results show that all hypothetical relationships except for H4 and H6 are supported. The path coefficients of these hypotheses are very close to zero. As predicted by H1 and H2, the study finds significant positive impacts of Usability on Attitude and on Intention-to-use. The effect size of both relationships proves to be moderate. Our findings confirm the favorable effect of Trust on Attitude (H3); H5, which predicted a positive relationship between Task-Technology Fit and Attitude, is also confirmed. Regarding H7, Attitude is found to be positively related to Intention-to-use. The last three causal relationships mentioned have small effect sizes. Table 4 shows the results of the modeling.

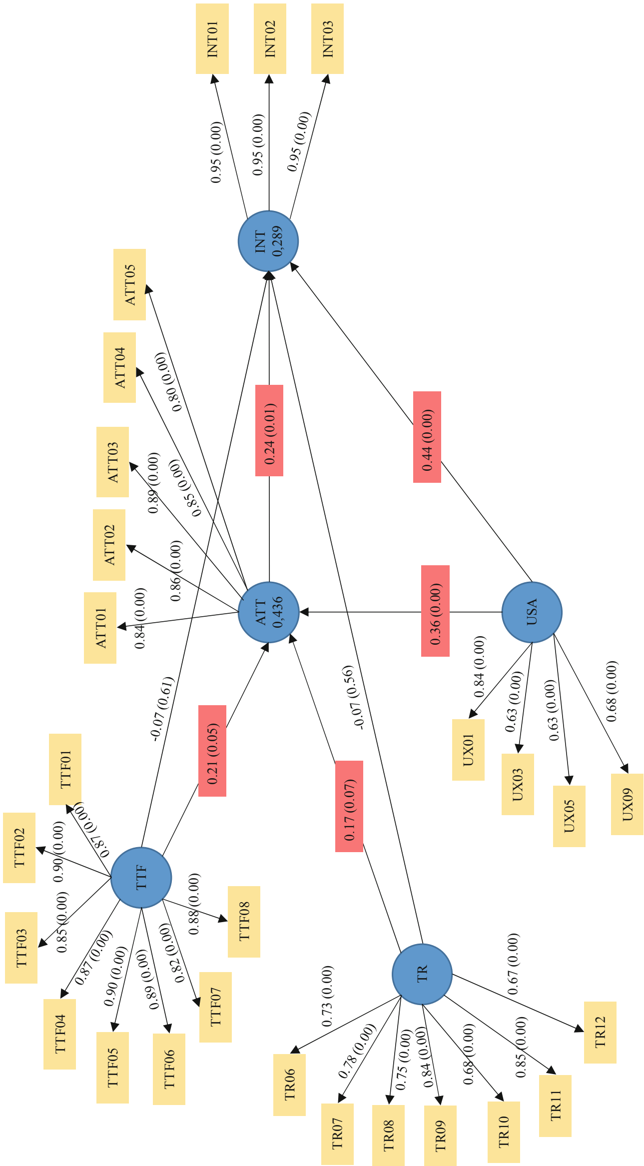

Fig. 2. PLS-SEM model

Table 4. PLS-SEM modelling results

Research hypotheses	Coefficient	P value	Outcome	f^2
H1: USA → ATT	0.357	$<0.000^a$	Supported	0.12^c
H2: USA → INT	0.444	$<0.000^a$	Supported	0.13^c
H3: TR → ATT	0.173	0.069^b	Supported	0.02^d
H4: TR → INT	−0.069	0.562	Not supported	0.00
H5: TTF → ATT	0.212	0.052^b	Supported	0.03^d
H6: TTF → INT	−0.066	0.612	Not supported	0.00
H7: ATT → INT	0.244	0.006^a	Supported	0.05^d

a: $p \leq 0.05$ b: $0,05 < p \leq 0.10$ c: moderate d: small

4.3 ANOVA

The descriptive analysis of the chosen constructs shows an overall good evaluation of the Usability of SELFPASS (SUS approx. 79). Among the four other constructs, the Task-Technology Fit construct receives the best rating on the 7-point Likert scale with a value of approximately 5, closely followed by Attitude and then Trust with an overall rating of approximately 4.7. The Intention-to-use construct was rated worst with an overall rating of approximately 4.0.

Table 5 illustrates the mean value and standard deviation of all constructs and distinguishes between T1 and T2 and the healthy and depressive groups.

Table 5. Descriptive analysis results

	Healthy T1 Mean (SD)	Depressive T1 Mean (SD)	Healthy T2 Mean (SD)	Depressive T2 Mean (SD)
Usability – SUS*	80.83 (12.30)	73.94 (14.23)	79.49 (14.87)	80.30 (9.07)
Trust**	4.81 (1.19)	4.20 (0.86)	5.03 (1.27)	4.69 (0.86)
Task-Technology-Fit**	5.32 (1.18)	4.55 (1.19)	5.18 (1.36)	5.25 (0.92)
Attitude**	5.19 (1.12)	4.38 (1.29)	4.99 (1.31)	4.81 (1.24)
Intention-to-use**	3.84 (2.15)	5.32 (1.37)	2.96 (2.04)	4.74 (1.57)

*Score from 0 to 100. **Score on a 7-point Likert scale

We calculated Cronbach's Alpha at both points in time (see Table 6) in order to determine sufficient reliability for the following analyses. Table 6 indicates, that the reliability of Usability (T1 and T2), Trust (T1 and T2) and Attitude (T1 and T2) can be rated as excellent [27]. The Task-Technology Fit and Intention-to-use construct have high reliability measures, indicating them as redundant items. Since the elimination of individual items did not lead to any significant improvement in reliability, we refrained from doing so.

Table 6. Cronbach alpha coefficients

	Usability (10 items)	Trust (7 items)	Task-technology Fit (8 items)	Attitude (5 items)	Intention-to-use (3 items)
T1	0.82	0.89	0.97	0.91	0.94
T2	0.76	0.87	0.94	0.90	0.97

As this sample is relatively small, we used the Shapiro-Wilk test to verify normal distribution. Turns out, not all constructs are normally distributed. However, we performed an ANOVA because there is no non-parametric equivalent and studies have shown that ANOVA is robust against violations of normality [28]. Table 7 presents the results of the ANOVA.

Table 7. ANOVA results

Construct	Group	P-value	Partial eta-squared
Usability - SUS	Time of measurement (T1 vs. T2)	0.070	0.05
	Condition (healthy vs. depressive)	0.006[a]	0.111[c]
Trust	Time of measurement (T1 vs. T2)	<0.000[a]	0.175[d]
	Condition (healthy vs. depressive)	0.158	0.030
Task-Technology Fit	Time of measurement (T1 vs. T2)	0.012[a]	0.095[c]
	Condition (healthy vs. depressive)	<0.000[a]	0.190[d]
Attitude	Time of measurement (T1 vs. T2)	0.383	0.012
	Condition (healthy vs. depressive)	0.018[a]	0.085[c]
Intention-to-use	Time of measurement (T1 vs. T2)	0.002[a]	0.148[d]
	Condition (healthy vs. depressive)	0.513	0.007

a: $p \leq 0.05$ c: moderate effect d: strong effect

The analysis of the differences between T1 and T2 and among healthy and depressive participants revealed a total of 6 significant differences from the 10 analyzed ones.

The Usability does not change significantly during the five days of use, but the two groups differ significantly with a moderate effect. Figure 3 depicts clearly that at the beginning the participants suffering from depression rate the Usability of SELFPASS significantly worse than the healthy participants. However, the depressive participants give a considerably better rating after the five days of usage, and their evaluation even exceed that of the healthy participants.

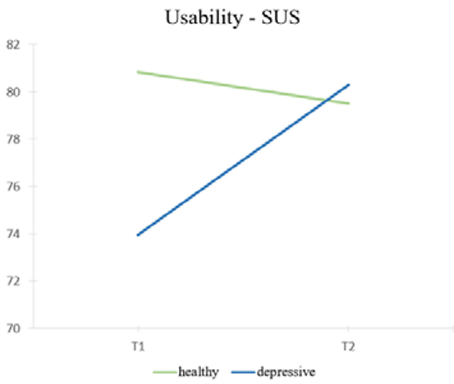

Fig. 3. Usability results

Trust in SELFPASS does not differ significantly between healthy and depressive participants, but there is a noticeable change in the assessment of trust during use, in the sense that trust in SELFPASS increases remarkably (strong effect, see Fig. 4). The Task-Technology Fit shows significant differences between T1 and T2 as well as between healthy and depressive participants. This construct also improves during application with moderate effect, with the increase being observed within the group of depressive participants. At the end of the usage period, the Task-Technology Fit is rated higher by depressive participants than by healthy ones (Fig. 5).

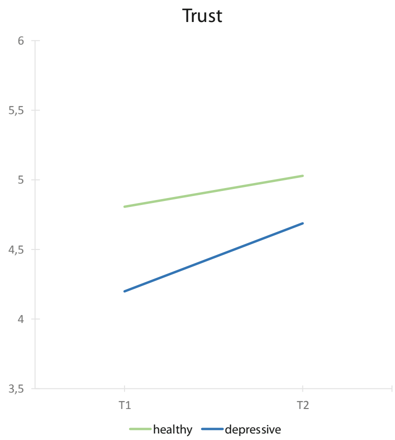

Fig. 4. Trust results

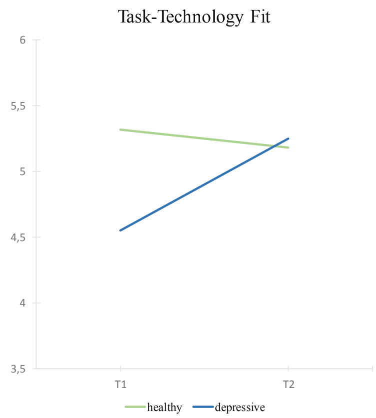

Fig. 5. Task-Technology Fit results

The Attitude construct shows a moderate, significant difference with respect to the existing health condition, whereby healthy participants demonstrate higher Attitude values than the depressive ones. The Attitude of the healthy participants decreases over the 5-day usage, while the Attitude of the depressive participants increases. Intention-to-use decreases significantly throughout the five days of usage with a strong effect (Figs. 6 and 7).

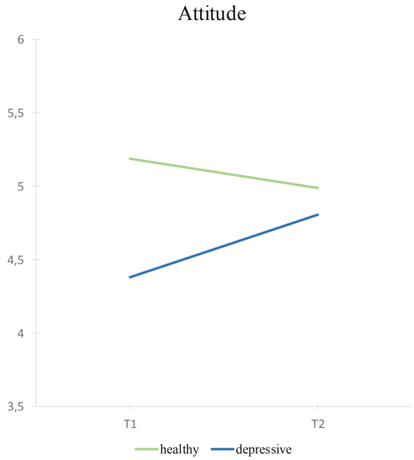

Fig. 6. Attitude results

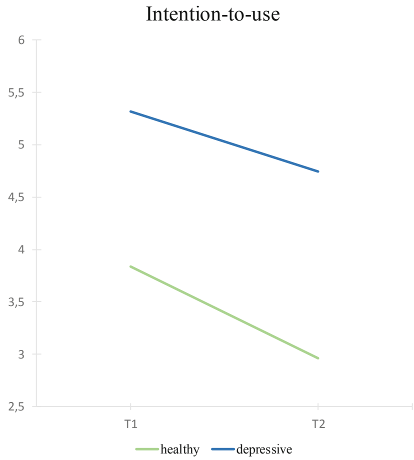

Fig. 7. Intention-to-use results

5 Discussion

The causal relationships described in the PLS-SEM model clearly show that Usability has the greatest influence on Attitude (path coefficient 0.36/f2 = 0.12) and on Intention-to-use (path co-efficient 0.44/f2 = 0.13). Therefore, it is fundamentally important to place a high significance on Usability when developing such therapeutic systems for people suffering from depression. It is striking that the factor loadings of the items of the Usability construct differ a lot and therefore do not show a satisfactory reliability before elimination. The ambiguity of the points or the participants' insufficient vocabulary seems to be unlikely causes due to the previous validation of the questionnaires. One reason could be the high complexity of the Usability construct which comprises various aspects. Furthermore, it is noticeable that the elimination of precisely those items led to an increase in reliability that were formulated with a negation (UX02,04,06,08,10) and caused strongly distorted distributions. Altogether, SELFPASS already achieves a good overall Usability evaluation (SUS ca. 79). On the first day of usage, the depressive participants rated the Usability significantly worse than after the five-day usage period, although this change is not observed among the healthy participants. This development indicates that the depressive participants became accustomed to SELFPASS and to the handling of the system. The SUS of approximately 73 at T1 is improvable and shows that SELFPASS could not be operated easily enough at the beginning. Because depressive people frequently suffer from lack of motivation and digital self-management systems show high dropout rates [29, 30], an improvement of compliance could be achieved through a specific Usability adapted to the target group. We should aim to enable effortless and intuitive usage in digital self-management systems for people suffering from depression. Thus, the period of familiarization with the system could be reduced, thereby preventing premature dropout. This finding is in accordance with the literature, which shows that for example guidance regarding key functionalities [31], clearly structured content and overviews [32], warning notices [33] and confirmation or congratulation messages after completed activities [34] are particularly important for people suffering from depression to improve orientation and ease of use.

Trust (path coefficient 0.17/f2 = 0.02) and Task-Technology Fit (path coefficient 0,21/f2 = 0.03) have a positive influence on Attitude. The Trust score improves greatly in the course of the five-day usage. This shows that trust in health-related digital platforms is only built up over time and does not come about immediately. Other studies have proven that trust in medical technology empirically differs from the general trust in technology [35]. Therefore, trust seems to be more indispensable if health-related aspects are conveyed technologically [36]. Although the literature attaches a very critical importance to trust in the field of eHealth, SELFPASS achieves an overall moderate to good evaluation with a mean of approximately 4.7. The evaluation of the Task-Technology Fit improves among depressive participants during usage. This suggests that users recognize an added value of SELFPASS while using the platform and shows that SELFPASS fulfils its purpose as a self-management tool for depression. This goes hand in hand with the observation, that this trend is not discernible among the healthy participants. They are healthy and have no psychological strain; therefore, they naturally recognize less benefit and improvement with SELFPASS, which is why the Task-Technology Fit barely changes for them in the course of five days. On the fifth day of us-age, they rate the fit slightly worse than the depressive group; hence, they assess SELFPASS as being less helpful and suitable.

The Attitude construct shows a similar curve progression. Over the period of use, the participants suffering from depression improve their attitude towards SELFPASS while that of healthy participants diminishes slightly. From this, we conclude that people suffering from depression generally have a positive estimation regarding e-mental health applications such as SELFPASS. The coefficient of determination (R2) of Attitude is approximately 0.44 and therefore, explains 44% of the variance. Compared to other studies in the field of eHealth this is a good result [3, 37, 38]. This study is one in which we measure human behavior and naturally in this area, numerous and of-ten not directly measurable, influences come into play. Therefore, smaller R2 values are to be expected here than in other disciplines, such as physics, with exactly measurable variables and low disturbances.

The Attitude construct shows a positive influence on Intention-to-use (path coefficient 0.24/f2 = 0.05). We had expected this effect, and it is congruent with the literature [14]. The small R2 of Intention-to-use (approximately 0.29) could be explained by the fact that Intention-to-use strongly depends on the subgroup, and depressive and healthy participants have generally different motivations for usage. Strikingly, in this study, healthy participants quit the study rather earlier than the depressive ones. One reason could be the non-existing psychological strain. Intention-to-use decreases over the five-day period, highlighting the necessity to integrate strong elements into e-mental health applications that increase motivation and compliance. Poor compliance is also discussed in the literature as a common obstacle to the use of eHealth applications [5] and gamification is addressed among other things. There is promising evidence that suggests gamification works. Innovative ways need to be found to make digital health interventions entertaining and appealing; these may include, for example, providing meaningful rewards or making the system more social [39]. Undoubtedly combining gamification and the special needs of depressive people in a meaningful way would be a challenging task.

6 Limitations

This study has some limitations. As conducting studies anonymously is a sensitive process, especially in the health-related area, we used an identification code to maintain confidentiality in the collection of the survey data. This in turn did not allow us to confirm whether the participants did in fact use SELFPASS daily in the manner required. Furthermore, we assume that due to the iterative nature of Internet interventions and the varying intensity and duration of time for which the users tested SELFPASS, the intervention exposure was likely to be different for each user. The research used participants' self-reports, and we can't guarantee that the participants correctly articulated their assessments. Since patients of the Heidelberg University Hospital received a fee for participating in the study, but the participants in Berlin did not, distortions cannot be ruled out. The SELFPASS version used was only a prototype and did not have the full range of the functions. A repeated measurement with the completed SELFPASS version could lead to different results, especially in terms of the Task-Technology Fit. Due to feasibility constraints, the resulting sample size was relatively small for such a complex investigation, leading to limited generalizability of the findings to the population as a whole. The small sample size of the group consisting of participants with depression could be a reason why the multi-group analysis of the structural equation modelling did not reveal significant group differences.

7 Conclusion and Future Work

The study contributes to the literature by pinpointing significant effects to help understand the usage of e-mental health applications to manage depression. PLS-SEM structural equation modelling proves that the Usability, Trust and Task-Technology Fit constructs have a positive effect on Attitude towards SELFPASS and that Attitude has a positive influence on Intention-to-use. The Usability has the biggest influence and should therefore be given special consideration in the development of self-management systems for people suffering from depression.

Overall, the ANOVA results reveal clear differences between healthy and depressive participants. The trend observed is that depressive participants generally rate SELFPASS better on the fifth day than on the first, therefore showing that they require a longer period for familiarization with the system compared to the healthy participants. The Intention-to-use decreases in both subgroups during the five-day usage, showing the necessity of further research projects to improve compliance to digital self-management systems for people suffering from depression. Furthermore, an effectiveness study of SELFPASS compared to a waiting list group could be a topic of interest for future research and practice. Whether self-management systems such as SELFPASS will also be suitable for patients suffering from severe depression, and under which conditions, remains largely unknown and also requires further research.

References

1. Hartmann, R., Sander, C., Lorenz, N., Böttger, D., Hegerl, U.: Utilization of patient-generated data collected through mobile devices: insights from a survey on attitudes toward mobile self-monitoring and self-management apps for depression. JMIR Mental Health **6**(4), e11671 (2019). https://doi.org/10.2196/11671
2. Bundespsychotherapeutenkammer BPtK: Ein Jahr nach der Reform der Psychotherapie-Richtlinie, Berlin (2018)
3. Ahadzadeh, A.S., Pahlevan Sharif, S., Ong, F.S., Khong, K.W.: Integrating health belief model and technology acceptance model: an investigation of health-related internet use. J. Med. Internet Res. **17**(2), e45 (2015). https://doi.org/10.2196/jmir.3564
4. Radovic, A., Gmelin, T., Hua, J., Long, C., Stein, B.D., Miller, E.: Supporting our valued adolescents (SOVA), a social media website for adolescents with depression and/or anxiety: technological feasibility, usability, and acceptability study. JMIR Mental Health **5**(1), e17 (2018). https://doi.org/10.2196/mental.9441
5. Lutz, W., et al.: Defining and predicting patterns of early response in a web-based intervention for depression. J. Med. Internet Res. **19**(6), e206 (2017). https://doi.org/10.2196/jmir.7367
6. Andersson, G., Cuijpers, P., Carlbring, P., Riper, H., Hedman, E.: Guided Internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: a systematic review and meta-analysis. World Psychiatry Official J. World Psychiatr. Assoc. (WPA) **13**(3), 288–295 (2014). https://doi.org/10.1002/wps.20151
7. Harrison, R., Flood, D., Duce, D.: Usability of mobile applications: literature review and rationale for a new usability model. J. Interact. Sci. **1**(1), 1 (2013). https://doi.org/10.1186/2194-0827-1-1
8. Kortum, P., Peres, S.C.: Evaluation of home health care devices: remote usability assessment. JMIR Hum. Factors **2**(1), e10 (2015). https://doi.org/10.2196/humanfactors.4570
9. Cho, H., et al.: A mobile health intervention for HIV prevention among racially and ethnically diverse young men: usability evaluation. JMIR mHealth uHealth **6**(9), e11450 (2018). https://doi.org/10.2196/11450
10. Tseng, S., Fogg, B.J.: Credibility and computing technology. Commun. ACM **42**(5), 39–44 (1999). https://doi.org/10.1145/301353.301402
11. Mcknight, D.H., Carter, M., Thatcher, J.B., Clay, P.F.: Trust in a specific technology. ACM Trans. Manage. Inf. Syst. **2**(2), 1–25 (2011). https://doi.org/10.1145/1985347.1985353
12. Jian, J.-Y., Bisantz, A.M., Drury, C.G.: Foundations for an empirically determined scale of trust in automated systems. Int. J. Cogn. Ergon. **4**(1), 53–71 (2000). https://doi.org/10.1207/S15327566IJCE0401_04
13. Goodhue, D.L., Thompson, R.L.: Task-technology fit and individual performance. MIS Q. **19**(2), 213 (1995). https://doi.org/10.2307/249689
14. Davis, F.D.: A Technology Acceptance Model for Empirically Testing New End-User Information Systems: Theory and Results (1986)
15. Burney, S.M.A., Ali, S.A., Ejaz, A., Siddiqui, F.A.: Discovering the correlation between technology acceptance model and usability. IJCSNS Int. J. Comput. Sci. Netw. Secur. **17**(11), 53–61 (2017)
16. Dishaw, M.T., Strong, D.M.: Extending the technology acceptance model with task–technology fit constructs. Inf. Manage. **36**(1), 9–21 (1999). https://doi.org/10.1016/S0378-7206(98)00101-3
17. Zhao, J., Fang, S., Jin, P.: Modeling and quantifying user acceptance of personalized business modes based on TAM, trust and attitude. Sustainability **10**(2), 356 (2018). https://doi.org/10.3390/su10020356

18. Löwe, B., Kroenke, K., Herzog, W., Gräfe, K.: Measuring depression outcome with a brief self-report instrument: sensitivity to change of the Patient Health Questionnaire (PHQ-9). J. Affect. Disord. **81**(1), 61–66 (2004)
19. Brooke, J.: SUS-A quick and dirty usability scale. In: Usability Evaluation in Industry, vol. 189, pp. 4–7 (1996)
20. Lin, T.-C., Huang, C.-C.: Understanding knowledge management system usage antecedents: an integration of social cognitive theory and task technology fit. Inf. Manage. **45**(6), 410–417 (2008). https://doi.org/10.1016/j.im.2008.06.004
21. Ajzen, I.: The theory of planned behavior. Organ. Behav. Hum. Decis. Process. **50**(2), 179–211 (1991). https://doi.org/10.1016/0749-5978(91)90020-T
22. Ajzen, I.: Perceived behavioral control, self-efficacy, locus of control, and the theory of planned behavior 1. J. Appl. Soc. Psychol. **32**(4), 665–683 (2002). https://doi.org/10.1111/j.1559-1816.2002.tb00236.x
23. Zhang, X., Yu, P., Yan, J., Spil, I.T.A.M.: Using diffusion of innovation theory to understand the factors impacting patient acceptance and use of consumer e-health innovations: a case study in a primary care clinic. BMC Health Serv. Res. **15**, 71 (2015). https://doi.org/10.1186/s12913-015-0726-2
24. Hair, J.F., Hult, G.T.M., Ringle, C.M., Sarstedt, M., Richter, N.F., Hauff, S.: Partial Least Squares Strukturgleichungsmodellierung. Eine anwendungsorientierte Einführung. Franz Vahlen, München (2017)
25. Hair, J.F., Sarstedt, M., Ringle, C.M., Mena, J.A.: An assessment of the use of partial least squares structural equation modeling in marketing research. J. Acad. Mark. Sci. **40**(3), 414–433 (2012). https://doi.org/10.1007/s11747-011-0261-6
26. Henseler, J., Ringle, C.M., Sarstedt, M.: A new criterion for assessing discriminant validity in variance-based structural equation modeling. J. Acad. Mark. Sci. **43**(1), 115–135 (2014). https://doi.org/10.1007/s11747-014-0403-8
27. Tavakol, M., Dennick, R.: Making sense of Cronbach's alpha. Int. J. Med. Educ. **2**, 53–55 (2011). https://doi.org/10.5116/ijme.4dfb.8dfd
28. Blanca, M.J., Alarcón, R., Arnau, J., Bono, R., Bendayan, R.: Non-normal data: Is ANOVA still a valid option? Psicothema **29**(4), 552–557 (2017). https://doi.org/10.7334/psicothema2016.383
29. Ryan, C., Bergin, M., Wells, J.S.G.: Theoretical perspectives of adherence to web-based interventions: a scoping review. Int. J. Behav. Med. **25**(1), 17–29 (2017). https://doi.org/10.1007/s12529-017-9678-8
30. Sherdell, L., Waugh, C.E., Gotlib, I.H.: Anticipatory pleasure predicts motivation for reward in major depression. J. Abnorm. Psychol. **121**(1), 51–60 (2012). https://doi.org/10.1037/a0024945
31. Fuller-Tyszkiewicz, M., et al.: A mobile app-based intervention for depression: end-user and expert usability testing study. JMIR Mental Health **5**(3), e54 (2018). https://doi.org/10.2196/mental.9445
32. Good, A., Sambhanthan, A.: Accessing web based health care and resources for mental health: interface design considerations for people experiencing mental illness. In: Marcus, A. (ed.) DUXU 2014. LNCS, vol. 8519, pp. 25–33. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-07635-5_3
33. Stiles-Shields, C., Montague, E., Lattie, E.G., Schueller, S.M., Kwasny, M.J., Mohr, D.C.: Exploring user learnability and learning performance in an app for depression: usability study. JMIR Hum. Factors **4**(3), e18 (2017). https://doi.org/10.2196/humanfactors.7951
34. Tiburcio, M., Lara, M.A., Aguilar Abrego, A., Fernández, M., Martínez Vélez, N., Sánchez, A.: Web-based intervention to reduce substance abuse and depressive symptoms in mexico: development and usability test. JMIR Mental Health **3**(3), e47 (2016). https://doi.org/10.2196/mental.6001

35. Montague, E.N.H., Kleiner, B.M., Winchester, W.W.: Empirically understanding trust in medical technology. Int. J. Ind. Ergon. **39**(4), 628–634 (2009). https://doi.org/10.1016/j.ergon.2009.01.004
36. Wilkowska, W., Ziefle, M.: Understanding trust in medical technologies. In: Proceedings of the 4th International Conference on Information and Communication Technologies for Ageing Well and e-Health : Funchal, Madeira, Portugal, 22–23 March 2018. SCITEPRESS - Science and Technology Publications Lda, Setúbal (2018)
37. Eivazzadeh, S., Berglund, J.S., Larsson, T.C., Fiedler, M., Anderberg, P.: Most influential qualities in creating satisfaction among the users of health information systems: study in seven European union countries. JMIR Med. Inform. **6**(4), e11252 (2018). https://doi.org/10.2196/11252
38. Laugesen, J., Hassanein, K., Yuan, Y.: The impact of internet health information on patient compliance: a research model and an empirical study. J. Med. Internet Res. **17**(6), e143 (2015). https://doi.org/10.2196/jmir.4333
39. Cugelman, B.: Gamification: what it is and why it matters to digital health behavior change developers. JMIR Serious Games **1**(1), e3 (2013). https://doi.org/10.2196/games.3139