The Role of Personalization, Engagement, and Feedback in Digital Health Interventions for Anxiety

Leonie Reick

Master Thesis

Positive Clinical Psychology and Technology
Department of Behavioral, Management and Social Sciences, University of Twente

1st supervisor: Dr. S.M. Kelders
2nd supervisor: Dr. M. Schotanus-Dijkstra

June 26th, 2023

Abstract

The prevalence of anxiety symptoms in university students is increasing, however, many do not receive conventional treatment due to barriers in health care. While digital mental health interventions (DMHIs) are a prospective treatment option, they often remain ineffective, which may be due to a low user-engagement. Literature indicates that personalizing interventions may lead to an increased engagement. Hence, this research aimed to investigate the effect of a personalized DMHI on user's engagement and anxiety, and if engagement mediates the link between personalization and anxiety. It was also explored if prescripted video feedback of a counselor is related to higher engagement and lower anxiety scores than other feedback types. A randomized controlled trial (RCT) was conducted. With a 3x3x3 factorial design, 27 different intervention versions were created. The experimental group ($N = 117$) received a personalized, and the control group ($N = 113$) a random 2-week-micro-intervention version. Engagement (TWEETS) and anxiety symptoms (GAD-7) were measured at multiple time points. Personalization was not linked to anxiety, nor to overall-, behavioral-, cognitive-, and affective engagement scores at post-intervention. Engagement did not mediate the association between personalization and anxiety. While engagement decreased throughout the intervention, anxiety scores decreased by post-intervention, but increased again by follow-up. Participants with a higher engagement had significantly lower anxiety scores at post-intervention and follow-up. Lastly, results indicated that prescripted councelor video feedback is not associated with lower anxiety, nor higher engagement scores. Future research is provided with recommendations on how to imporve the personalization approach. With the novel finding that a higher engagement may predict lower anxiety scores several weeks after completing the treatment, future research is advised to further investigate how users' engagement can be maintained to increase intervention effectiveness.

Keywords: digital mental health intervention (DMHI), randomized control trial (RCT), personalization, engagement, anxiety, video feedback, university students

The Role of Personalization, Engagement, and Feedback in DHIs for Anxiety

Anxiety is one of the most widespread mental disorders, with the prevalence having increased by 25% worldwide during the COVID-19 pandemic (WHO, 2022). University students have been found to be at high risk to develop mental health problems, and specifically with the pandemic the prevalence of anxiety symptoms in this population was found to have increased worldwide to 31% (Osborn et al., 2022; Chang et al., 2021). However, a considerable number of the concerned students do not receive professional help due to current barriers in mental health care. This highlights the importance to improve conventional treatment.

Multiple prominent barriers hinder the access to face-to-face treatment. Firstly, there is an increasing demand of psychotherapy, but only limited therapy spots (WHO, 2019). This leads to a treatment gap with long waiting lists, extended time between treatments, time-limited sessions, and gatekeeping, hence, those in need may not receive adequate support (Auerbach et al. 2018; Benton et al., 2016). Secondly, conventional treatment is often associated with high costs and limited financial support by health insurances. Thirdly, on the patients' side, there may be a fear of social stigma or a limited knowledge of available resources which prevent the search for help (Osborn et al., 2022; Priestley et al., 2021).

Over the last years, the usage and development of eHealth has become increasingly prevalent and holds much potential in the improvement of health care. EHealth can be understood as "the use of technology to improve health, well-being and healthcare" (van Gemert-Pijnen et al., 2018) and the specification on mental well-being is referred to as digital mental health interventions (DMHIs) which may be a prospective treatment option for anxiety. DMHIs are typically defined as technologies where "information, support, and therapy for mental health conditions [are] delivered through an electronic medium with the aim of treating, alleviating, or managing symptoms" (Lehtimaki et al., 2021). Electronic mediums may be mobile phones, including apps and short message services, as well as VR programs and offline computer-based programs, with web-based technologies being most widely spread (Lattie et al., 2019). According to the literature reviews by Bolinski et al. (2020) and Lattie et al. (2019), interventions usually range between two to eight weeks and may be based on psychotherapies like cognitive behavioral therapy or mindfulness-based interventions. Furthermore, interventions often include psychoeducation where participants receive information about their illness to help them comprehend and cope, exercises that usually complement the psychoeducation and aim to reduce anxiety, as well as guidance, support, and feedback (Fleming et al., 2018; Bry et al., 2017; Higa-McMillan et al., 2016).

DMHIs are seen as a promising treatment option through overcoming barriers of conventional treatment. Firstly, DMHIs can generally be used by more people, which decreases costs while increasing therapy spots, and enables those with a low symptom-severity to receive help as well (Lattie et al., 2019; WHO, 2019). Secondly, treatment is made more accessible. As digital natives, most young adults own considerable means of technology, making DMHIs a convenient treatment option and empowering users with an easier access to information (Lehtimaki et al., 2021; Bolinski et al., 2020; Lal, 2019). Thirdly, DMHIs present an alternative to people who do not seek face-to-face treatment out of fear of stigma, discomfort to leave the house, or difficulties to open up to a stranger, as DMHIs can be accessed anonymously from home (van Orden et al., 2022; Rodríguez-Rivas et al., 2021).

Multiple recent systematic reviews found significant effects of DMHIs in reducing anxiety symptoms amongst university students (Riboldi et al., 2022; Bolinski et al., 2020; Lattie et al., 2019). DMHIs have shown to be more effective than waiting lists (Hall et al., 2018) and some studies found them to be as effective as face-to-face treatment in decreasing anxiety symptoms (Bendtsen et al., 2020; Olthuis et al., 2016). Yet, in the systematic review by Lattie et al. (2019), which assessed improving anxiety, depression, and psychological wellbeing of college students with DMHIs, only 42.5% of the 89 included studies were fully, while 30% were partially effective in enhancing mental health outcomes. Especially in studies with partial or not effective results, attrition rates were high and sustained program use was low (Lattie et al., 2019), thus, a low adherence may have been associated with insignificant findings.

Hence, while literature reflects the potential of DMHIs to overcome health care barriers and effectiveness in tackling anxiety symptoms, they often remain ineffective due to user's non-adherence to the intervention. Adherence is found to be crucial, with higher adherence being associated with increased therapeutic gains (Fleming et al., 2018). Compared to face-to-face therapy, non-adherence is specifically prevalent in self-help DMHIs without guidance or personal support (Fleming et al., 2018). Here, users may disengage in using the technology or do not perform intervention features as intended by developers (Kelders et al., 2020a). In the search for factors to increase adherence, hence, the objective usage of DMHIs, a link to engagement has been found (Lattie et al., 2019), which focuses on users' subjective reasons for working with a technology and how involved they feel (Kelders et al., 2020a).

Previous definitions of engagement have mainly focused on the behavioral context of using a DMHI, however, the systematic review by Kelders et al. (2020b) found engagement in DMHIs to be a multidimensional construct consisting of a cognitive, behavioural, and affective component. Furthermore, engagement is not to be seen as a mere state, but process made up of becoming and maintaining engaged, disengaging, as well as re-engaging (Kelders et al., 2020b). Torous et al. (2018) found low user-engagement in mental health applications as they are often not designed in respect to the user's needs, do not solve their concerns, lack in privacy, and are not perceived as trustworthy and helpful in emergencies. In line with this, studies have shown a low user engagement to come with a lower efficacy of the technological intervention (Fleming et al., 2018; Benton et al., 2016). This demonstrates the importance to focus on designing more engaging applications, but there is limited research on how DMHIs could enhance engagement.

First, personalization has received increased attention regarding its potential to increase engagement. In digital healthcare, personalization can be defined as the "process that changes the functionality, interface, information access and content, or distinctiveness of a system to increase its personal relevance to an individual" (Fan et al., 2006). Hereby, "characteristics, preferences, interests, and needs of users" are used to tailor systems to the individual (Kocaballi et al., 2019). By adapting the content and design of an intervention to users' needs, the aim is to maximize adherence, engagement, and therefore also the efficacy of the treatment (Burley et al., 2020; Miloff et al., 2015). For instance, Jones et al. (2022) conducted a randomized control trial, in which 98 adolescents were randomly assigned to a web-based cognitive-behavioural or interpersonal prevention program, with half receiving the treatment which matched their needs. The personalized treatment group indicated a decrease in anxiety symptoms, while those who were mismatched actually showed increased symptom. Hence, a personalized treatment seems to be vital for the intervention to help.

To the writer's knowledge, only Burley et al. (2020) investigated the relationship between personalization and engagement for mobile-based interventions targeting anxiety. In their study, an intervention was personalized though the content, as well as implementation intentions and goal setting. However, while in their pilot study five participants indicated to feel engaged, the personalization procedure was not evaluated quantitatively and in a larger sample. Hence, the effect of a personalized treatment on anxiety symptoms has not been investigated in the form of mobile-based interventions in university students. Furthermore, it is not yet known if personalization works through increasing engagement.

Second, feedback from a human therapist may play a vital role in engaging users. A systematic review by Lattie et al. (2019) found that compared to self-guided interventions and automated support, such as prescripted messages, interventions with an in-person element, like a professional or peer, achieved higher efficacy and treatment adherence, and lower dropout rates from study surveys. This is supported by other studies, which found in-person elements with the therapist to be linked to higher effectiveness (Välimäki et al., 2017), better completion and outcomes (Clarke et al., 2014), and increased adherence (Hollis et al., 2016) amongst young people with anxiety. In addition, in a study by Pretorius et al. (2019) 84% of young participants in the web-based intervention indicated to find human contact important. However, there is also conflicting evidence. For instance, Harrer et al. (2018), did not detect an effect of supervision on intervention results (as in Lattie et al., 2019). Furthermore, Kelders (2015) did not find significant differences in user-involvement between automated and human support in a behavior change support system (BCSS) for treating depression.

While human therapist's feedback may increase DMHI efficacy, past research focused on therapists' support through video calls and written messages, and it has not been investigated if prerecorded video messages from a counselor may have comparable effects. These may be of great value, as therapists only need to record the messages once, making it a cost and time effective strategy to reach users. While little is known about how therapists' guidance may enhance DMHI effectiveness, the proposed mechanisms are suggested to also be present in prerecorded messages. For instance, it may be because participants feel that someone cares about them (Pretorius et al., 2019). Furthermore, there is the suggestion of surveillance by Benton et al. (2016), who proposed that participants who received a low-intensity treatment, with a weekly 15 to 20-minute therapist contact, may have shown increased engagement and adherence as they knew their therapist could see their progress and activities. Hence, it is of added value to investigate if prerecorded video messages by counselors are effective in increasing user engagement, as well as reducing anxiety symptoms amongst university students.

Research Questions:

1. Does a personalized 2-week digital health intervention for anxiety have a larger effect on anxiety scores compared to a non-personalized intervention?
2. Does a personalized 2-week digital health intervention for anxiety have a larger effect on users' engagement (overall, behavioral, cognitive, and affective engagement) compared to a non-personalized intervention?
3. Is the effect of a personalized 2-week digital health intervention on the measures of anxiety mediated by users' engagement (overall, behavioral, cognitive, and affective engagement)?
4. Is the feedback in form of spoken words by a human counselor related to lower anxiety levels compared to plain text and virtual avatar feedback?
5. Is the user's overall engagement higher for those receiving spoken feedback from a therapist compared to plain text and virtual avatar feedback?

Methods

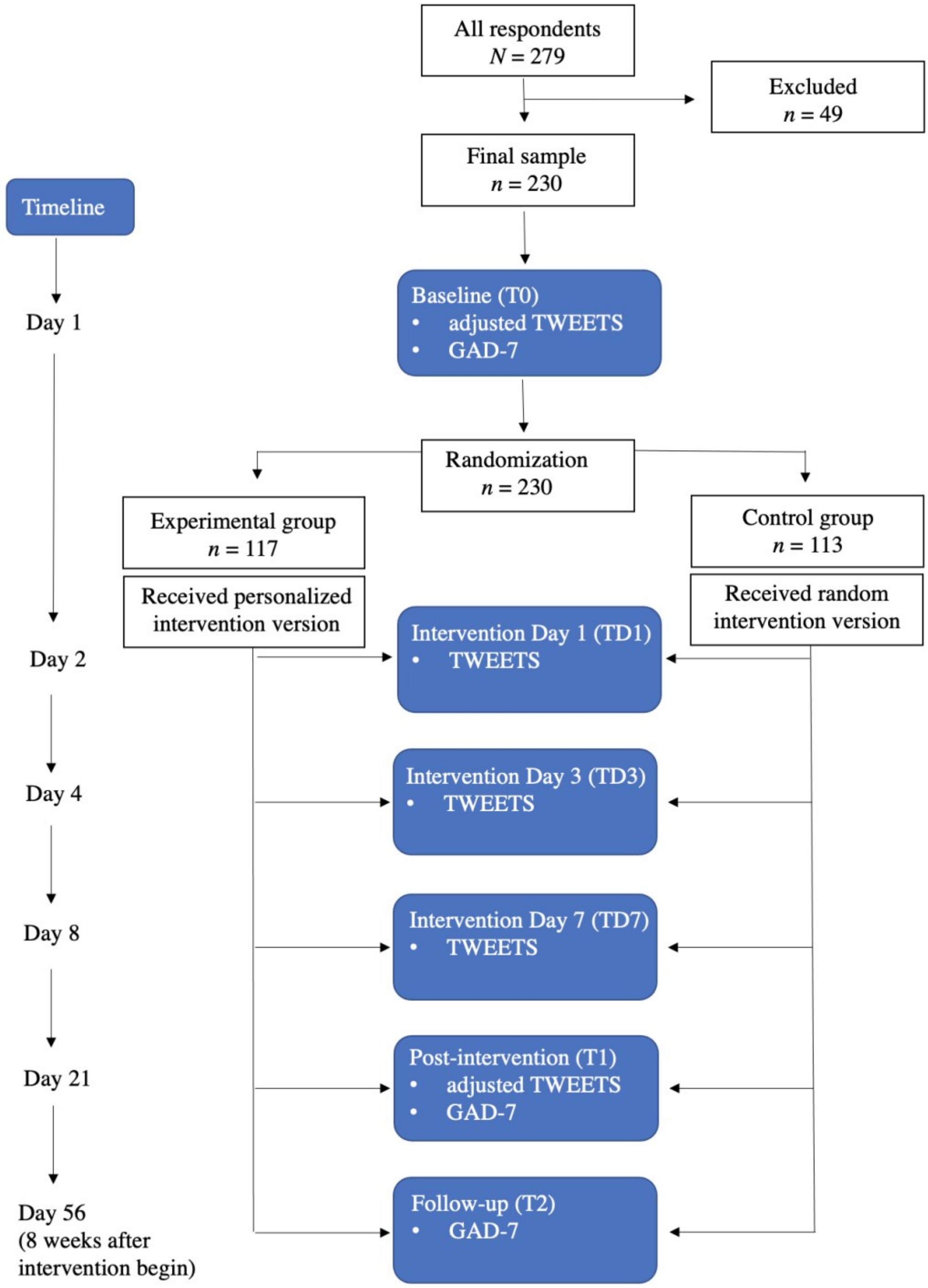
Participants

From an initial total of 279 participants, 49 were excluded as they did not fulfill the inclusion criteria. This left a final sample of 230 participants (see Figure 1). Inclusion criteria were that participants had a smartphone to download the Incredible Intervention Machine application (TIIM), spoke English, and were at least 18 years old. Participants were excluded if they did not have both a participant ID and TIIM code. In accordance with the intention-to-treat approach (Gupta, 2011), cases of noncompliance and missing outcomes were kept, as they may have been due to low engagement, which was of interest for the research.

Of the 230 participants, 117 (50.9%) were in the experimental and 113 (49.1%) in the control group. Ages ranged between 18 and 30, with a mean age of 20.47 (*SD* = 1.90). All participants were students and the majority identified as being female (*N* = 163, 70.8%). Most participants were German (57%), while 25.2% were Dutch, and 17.8% had another nationality, such as Namibian, Brazilian, and Vietnamese. Table 1 depicts the background information for each of the conditions and the total sample.

Figure 1

Participant Flow Chart with Timeline of Intervention and Measurements.



Note. Allocation of participants into experimental and control groups, timeline with intervention time points and measurements.

Table 1

Descriptive Statistics of the Background Information of each Condition and the Total Sample.

Characteristics		Experimental group		Control Group		Total	
		n	%	*n*	%	*n*	%
Participants		117	50.9	113	49.1	230	100
Gender	Female	83	36.1	80	34.8	163	70.8
	Male	33	14.3	33	14.3	66	28.7
	Non-binary	-	-	-	-	-	-
	Prefer not to say	1	0.4	-	-	1	0.4
Age	18 - 20	63	27.4	75	32.6	138	60 15.7
	21-25	51	22.2	36	87	37.8	
	26-30	1	0.4	2	0.9	3	1.3
	31	1	0.4	-	-	1	0.4
	Missing	-	-	-	-	1	0.4
Nationality	Dutch	26	11.3	32	13.9	58	25.2
	German	71	30.9	60	26.1	131	57
	Other	20	8.7	21	91.3	41	17.8
Feedback	Plain Text	29	42.6	39	57.4	68	29.6
	Virtual Avatar	26	40	39	60	65	28.3
	Human Counselor Video	62	63.9	35	36.1	97	42.2

Design and Procedure

The experimental randomized controlled trial (RCT) was part of a larger research project performed by Kelders at the University of Twente. In the RCT, an experimental group, who received a personalized intervention version was compared to a control group, who got a random intervention version. Participants were recruited through convenience sampling via the SONA system of the University of Twente and were rewarded with SONA credits, which students need to complete their studies. After signing the informed consent form, participants received the baseline questionnaire, which was administered via Qualtrics. They were asked to create a participant ID, fill out demographic questions, the GAD-7, and the adjusted TWEETS questionnaire. After the baseline questionnaire, participants were assessed for eligibility and those who met the inclusion criteria could register in the TIIM app and received a TIIM code.

There was one module for each day of the two-week intervention. Each module included an exercise that was to be completed by the end of the day, for which participants were reminded by the app. On intervention days 1, 3, and 7, participants additionally received the TWEETS questionnaire in the app. Participants were given 21 days to complete the intervention, after which they were sent a link to Qualtrics for the post-intervention survey, containing an adjusted TWEETS and GAD-7 questionnaire. Lastly, a follow-up survey was sent eight weeks after the beginning of the intervention with a GAD-7 questionnaire. Figure 1 gives an overview of the timeline and measurement points of the intervention. The study was approved by the ethics committee of the faculty of Behavioral, Management and Social Sciences (BMS) at the University of Twente (number: 220083) and data was collected from February to July 2022.

The Intervention

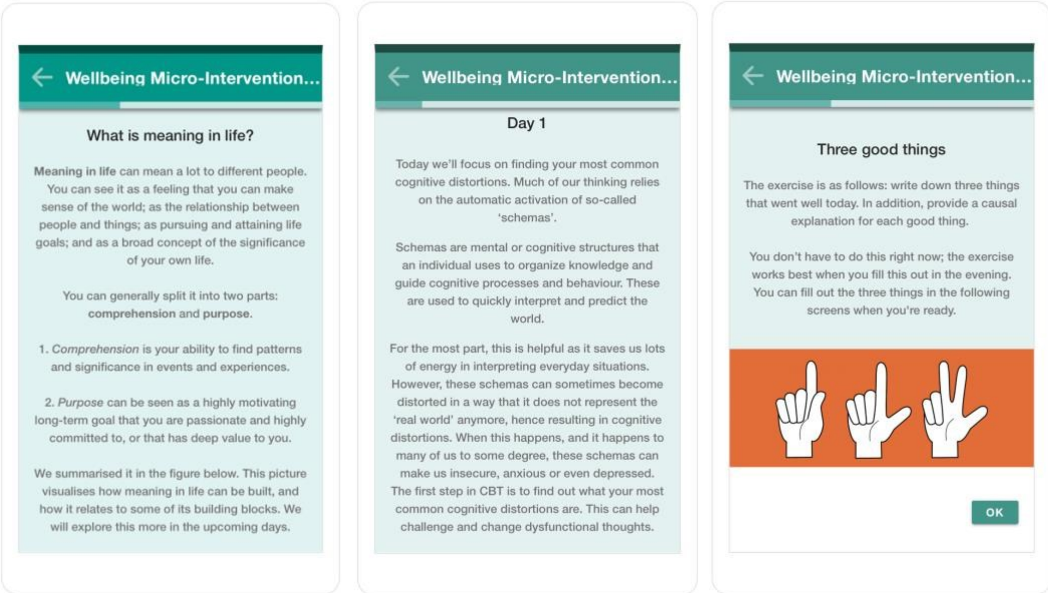
The Incredible Intervention Machine (TIIM App) was used to deliver one of 27 different intervention versions to each participant. These were created through the 27 possible combinations of the three intervention and technological *factors* (ITFs): *content, feedback,* and *design*, each of which disposed of three diverse *options* (3x3x3 research design). Hence, through the conjuncture of the three *options* across the three *factors,* 27 intervention *versions* were established. All versions had 14 modules and the same build-up.

Through the factor *content*, interventions could be based on three different evidence-based psychological approaches: the *Meaning Intervention*, *Cognitive Behavioral Therapy*, or *Positive Psychology*. In line with the therapeutic approach, each intervention included an explanation of the approach, psychoeducation for the content covered in the module, followed by instructions for the exercises (see Figure 2). For instance, CBT based interventions included the explanation that “CBT focuses on challenging and changing dysfunctional thoughts, emotions, and/or related behaviors […]”. While the content of the first week focused on reflecting, identifying and challenging dysfunctional thoughts, the second week explored ways to change these. To illustrate, psychoeducation on day 1 explained schemas and cognitive distortions. For the exercise, cognitive distortions were to be identified until the end of the day.

Feedback was delivered in three different forms after the completion of a module, with all entailing the same content (see Figure 3). First, participants may receive feedback as a *plain text*. Second, feedback was delivered by a *virtual agent*, where a drawn character was illustrated on the screen with the same text as for the first group. Lastly, participants may have received the text in form of spoken words by a *human counselor as a prerecorded video*.

Figure 2

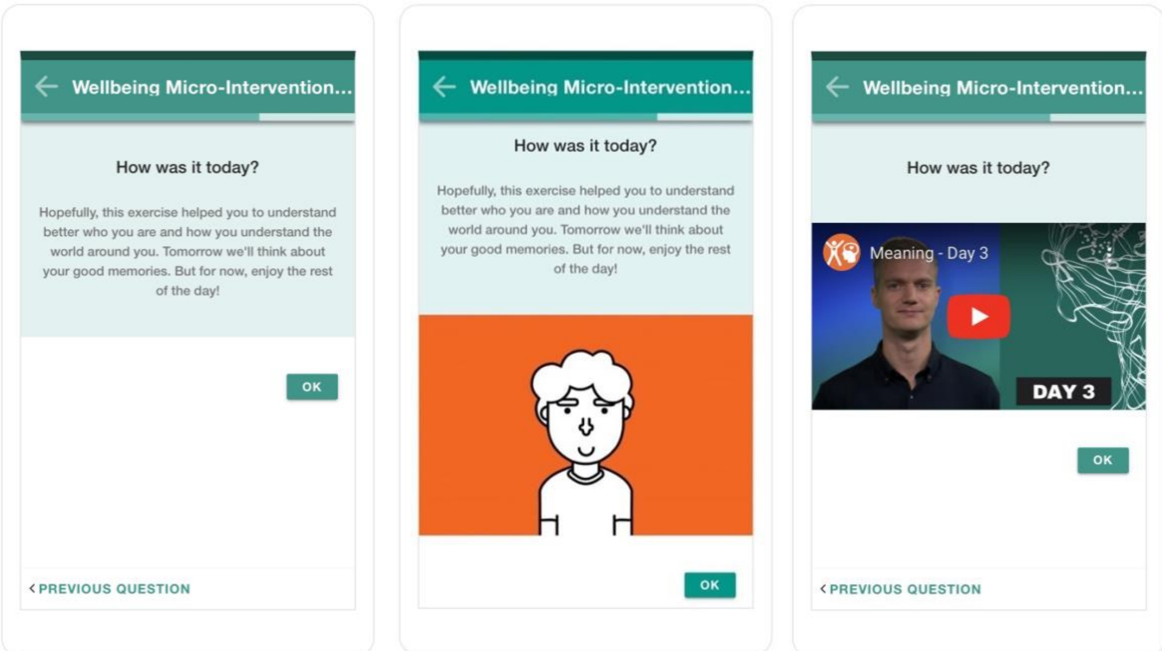
Illustration examples of psychoeducation and exercises for the psychological approaches.



Note. From left to right: explanation of the Meaning Intervention, psychoeducation for intervention day 1 of Cognitive Behavioral Therapy, instruction for intervention day 1 exercise of Positive Psychology.

Figure 3

Illustration examples of feedback possibilities after an exercise.

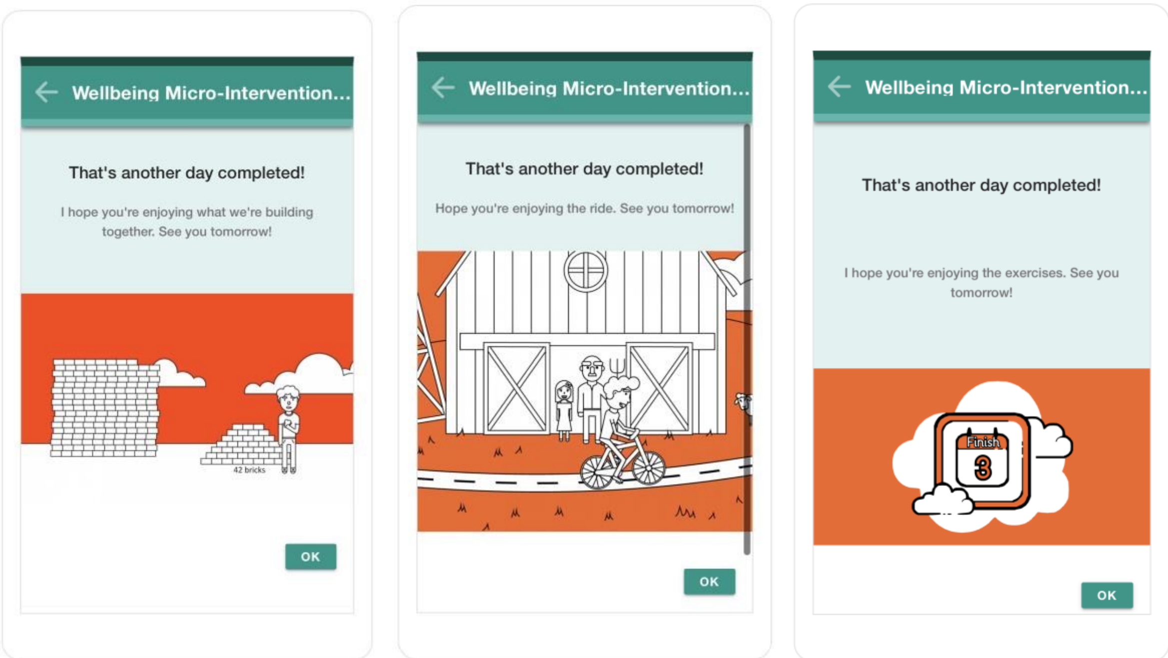


Note. From left to right: feedback in form of plain text, feedback by a virtual agent, and feedback by a human counselor as a prerecorded video.

Each module ended with one of three possible *designs* (see Figure 4). In the *competitive gamification* element, earned bricks were added from the left to the right stack, which symbolized the participant's progress and included the bricks that were collected on the days before as well. The *non-competitive gamification* showed a storyline and illustrated the progress with a person cycling to the next location of the intervention course. Lastly, if the intervention entailed *no gamification element*, a calendar was depicted which showed the day of the intervention.

Figure 4

Illustration examples of design possibilities.



Note. From left to right: competitive gamification, non-competitive gamification, no gamification element.

Personalization

As part of the personalization procedure, participants filled out a baseline survey entailing an adjusted TWEETS questionnaire by Kelders et al. (2020) (see Table 2) which asked them to indicate their expected engagement for each of the nine intervention *options*. The meaning intervention, and cognitive behavioural therapy, and positive psychology were defined and explained, including a corresponding screenshot of the intervention display with an example exercise. Participants rated their expected engagement if an app used this specific *content*. The same process was repeated for the *feedback* and *design options* (see Table 2).

The total sum score was calculated for each *option*, ranging from 0 (= strongly disagree) to 36 (= strongly agree), with higher sum scores indicating a participant's assumption to be more highly engaged with the intervention if the according *option* was entailed in the app. For each of the ITFs (*content, feedback, and design*), participants in the experimental group received the *option* which they rated they would be most engaged with, and therefore received one of the 27 intervention versions that was personalized to them. When scoring equally on two or more *options* of one factor, random.org was used to choose one option. The control group was randomly assigned to one of the 27 intervention versions.

Table 2

TWEETS items adjusted for the baseline questionnaire to assess expected engagement for content and app-specific options.

Item	Content TWEETS items	Feedback and design TWEETS items
1	Using an app with this *content* can become part of my daily routine.	Using this *version of the app* can become part of my daily routine.
2	The *content* of this app is easy to use.	This *version of the app* is easy to use.
3	I will be able to use an app with this *content* as often as needed to improve my anxiety.	I will be able to use this *version of the app* as often as needed to improve my anxiety.
4	An app with this *content* will make it easier for me to work on decreasing my anxiety.	This *version of the app* will make it easier for me to work on decreasing my anxiety.
5	This *content* motivates me to decrease my anxiety.	This *version of the app* motivates me to decrease my anxiety.
6	This *content* will help me to get more insight into my anxiety.	This *version of the app* will help me to get more insight into my anxiety.
7	I will enjoy using an app with this *content*.	I will enjoy using this *version of the app*.
8	I will enjoy seeing the progress I make by using an app with this *content*.	I will enjoy seeing the progress I make with this *version of the app.*
9	An app with this *content* will fit me as a person.	This *version of the app* will fit me as a person.

Note. The TWEETS scale by Kelders et al. (2020a) was adjusted for the personalization procedure. Content TWEETS items focus on the factor *content* (Positive Psychology, Meaning Intervention, and Cognitive Behavioural Therapy). Feedback and design TWEETS items on factors *feedback* (plain written text, virtual agent, and human counsellor) and *design* (competitive gamification, non-competitive gamification, and no gamification). Rated on a 5-point Likert scale ranging from 0 (= strongly disagree) to 4 (= strongly agree).

Materials

Engagement

To measure the level of engagement, the TWente Engagement with Ehealth Technologies Scale (TWEETS) questionnaire by Kelders et al. (2020a) was utilized. It contains 9 items, measured on a 5-point Likert scale, ranging from 0 (= strongly disagree) to 4 (= strongly agree), with overall engagement sum scores ranging from 0 (= strongly disagree) to 36 (= strongly agree). Three areas of engagement are measured, namely the behavioral, cognitive, and affective engagement, with each being determined through 3 items. Items 1 to 3 (e.g. “this app is part of my daily routine”) measure behavioral engagement, items 4 to 6 (“this app motivates me to each my goal”) assess cognitive engagement, and items 7 to 9 (e.g. “I enjoy using this app”) were used to indicate affective engagement. Behavioral, cognitive, and affective engagement sum scores ranged from 0 (= strongly disagree) to 12 (= strongly agree). Mean scores were calculated for overall, behavioral, cognitive, and affective engagement for intervention days 1, 3, 7, and post-intervention. The scale possessed an excellent reliability for overall engagement ($\alpha = .93$), acceptable reliability for the behavioral engagement subscale ($\alpha = .76$), and a good reliability for the cognitive ($\alpha = .87$) and affective subscales ($\alpha = .88$), which is in accordance with Kelders et al. (2020) who found a high internal consistency. While the TWEETS was adjusted for baseline (see Table 2) and post-intervention (see Appendix A), the regular scale was used on intervention days 1, 3, and 7 (see Appendix B).

Anxiety

The Generalized Anxiety Disorder screening (GAD-7) is a 7-item questionnaire, which was developed by Williams (2014) and measures the extent to which participants experienced anxiety symptoms within the last two weeks. Items include for instance “feeling nervous, anxious, or on edge” and were scored on a 4-point Likert scale ranging from 1 (= not at all) to 4 (= nearly every day). Total sum scores were calculated for each participant, which span from 7 to 28. The cutoff score of 12 (adjusted to the scale of the study) indicates mild anxiety (Williams, 2014). Sum scores were calculated for baseline, post-intervention, and follow-up. The scales’ reliability has been found to be reliable ($\alpha = .79$ - $.91$) (Williams, 2014). Cronbach’s Alpha was calculated for this study, indicating a good reliability ($\alpha = .88$).

Data analysis

The data was analyzed with IBM SPSS Statistics, Version 28. Data from the diverse measurement points were merged into one dataset based on the participant IDs. If participants filled out a questionnaire multiple times, the most complete was kept, and if all had the same level of completion, the first response was chosen. The data set was converted into long format.

First, a Shapiro-Wilk test was used to investigate whether the continuous variables were normally distributed amongst the participants (see Appendix C). Overall-, behavioral-, cognitive-, and affective engagement measures at intervention day 1, day 3, day 7, and at post-intervention, as well as anxiety measures at baseline, post-intervention, and follow-up were non-normally distributed. While the mean (SD) was reported for variables with a normal distribution, the median (IQR) and non-parametric tests were chosen for variables with a non-normal distribution. Second, descriptive statistics were used to explore the demographics, anxiety, and engagement scores separately for the experimental and control group. Third, bivariate analyses were computed with Pearson's correlations between the personalization group, engagement (at TD1, TD3, TD7, T1), and anxiety (T0, T1, T2) to investigate possible relationships between the variables.

For research questions one and two, Mann-Whitney U tests were used to compare anxiety mean scores and engagement median scores of the experimental and control groups. For research question four, differences in anxiety scores of the three feedback groups were investigated with the Kruskal-Wallis H tests. Followingly, generalized linear mixed models were computed with participants indicated as the random factor, anxiety and engagement scores as dependent variables, and group, time, and the interaction effect group*time as fixed factors.

Research question three was explored with a mediation analysis using IBM SPSS Statistics extension PROCESS analysis, version 3.5, by Hayes (2017), with bootstrapping set to 5000 resamples (Alfons et al., 2019). Analyses were run with general-, behavioral-, cognitive-, and affective engagement mean scores as mediating variables. Group was the independent variable and anxiety at post-intervention the dependent variable. To answer research question five, Kruskal-Wallis one-way ANOVA tests were conducted to investigate if there were significant differences between engagement scores of the feedback groups. Findings had to possess a significance level of $p < .05$ to be interpreted as significant.

Results

Pearson's Correlations Table

Table 3 depicts bivariate correlations, which indicate that people who were already more engaged on the first day of the intervention also had the tendency to be more engaged at all following measurement points of intervention day 3 ($r = .72, p < .001$), day 7 ($r = .62, p < .001$), and post-intervention ($r = .55$, $p < .001$). Furthermore, those who were more engaged (at all measurement points) depicted lower anxiety scores at post-intervention. Lastly, lower anxiety scores at follow-up were associated with a higher engagement at intervention day 1 ($r = -.22$, $p = .036$) and post-intervention ($r = -.19$, $p = .049$).

Table 3

Pearson's Correlations between variables Group (personalization), Overall Engagement (at Intervention Day 1, 3, 7, and Postintervention (T1)), Anxiety (at Baseline, Postintervention (T1), and Follow-Up (T2)).

	Group	Engagement day 1	Engagement day 3	Engagement day 7	Engagement post-intervention	Anxiety baseline	Anxiety post-intervention	Anxiety follow-up
Group	1							
Engagement day 1	.1	1						
Engagement day 3	.03	.72**	1					
Engagement day 7	-.01	.62**	.72**	1				
Engagement post-intervention	.01	.55**	.66**	.76**	1			
Anxiety baseline	-.01	.06	.02	.01	.01	1		
Anxiety post-intervention	-.06	-.23**	-.24**	-.21**	-.24**	.48**	1	
Anxiety follow-up	.01	-.22*	-.12	-.11	-.19*	.28**	.68**	1

Note. Correlations significant at *p < .05, ** p < .01 level (2-tailed).

Effect of Personalization on Anxiety (RQ1)

Results indicated that there are no significant differences in anxiety scores between the experimental and control group at baseline (U = 6581.5, p = .954, r = -0.004), which supports the comparability of the two groups. The generalized linear mixed model showed time to have a significant effect on anxiety scores $F(2, 532) = 24.58$, $p < .001$. In line with descriptive statistics (see Table 4 and Figure 2), anxiety scores were significantly lower at post-intervention compared to baseline t(532) = 3.67, p = .005. However, anxiety scores significantly increased again from post-intervention to follow-up t(532) = -2.77, p = .006 and were statistically as high again as at baseline t(532) = 1.5, p = .131. Additionally, personalization did not have a significant effect on anxiety scores $F(1, 532) = .013$, $p = .908$ and no significant interaction effect of time*group was found $F(2, 532) = .084$, $p = .920$. Hence, while there were significant changes in anxiety scores, no differences between the experimental and control groups were found.

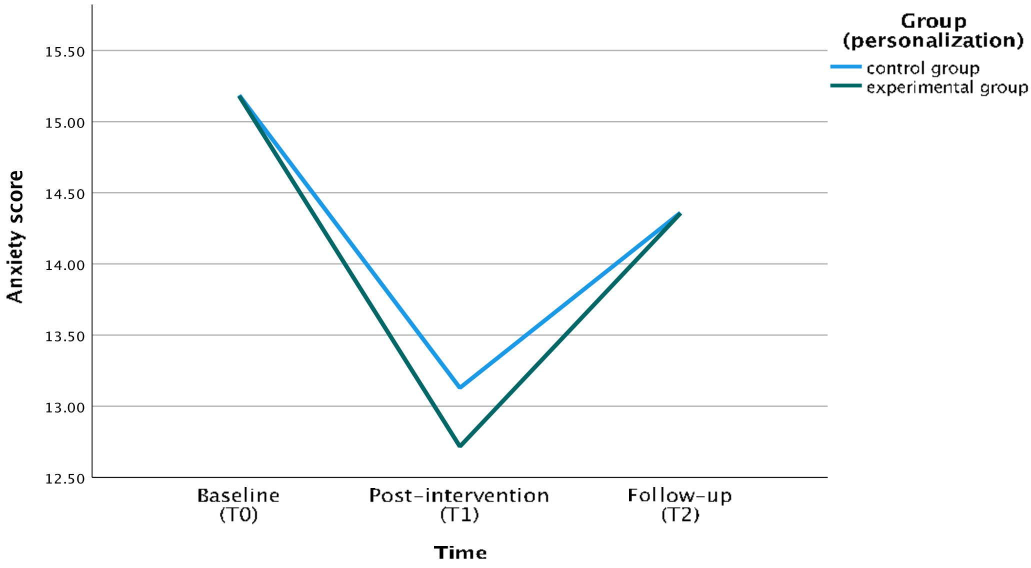
Table 4

Descriptive Statistics and Mann-Whitney U test of Anxiety and Engagement measures for both intervention groups.

Variable	Measurement	Group (personalization)					
		Experimental		Control		Mann-Whitney-U	
		Mdn	*IQR*	*Mdn*	*IQR*	*U*	*p*
Anxiety	Baseline	14	6	14	7.5	6581.5	.954
	Post-intervention	12	4	12.5	3.75	4525.5	.312
	Follow-up	13	8	13	7.5	1468	.966
Overall engagement	Day 1	26	4	25	5	3628	.057
	Day 3	25	6	24	6	4237.5	.544
	Day 7	24	6	25	5	4072.5	.751
	Post-intervention	24	9	25	6.5	4895	.933
Behavioral engagement	Day 1	8.5	2	9	2	4197	.725
	Day 3	8	2	8	2	4434	.933
	Day 7	8	2	8	2	4021	.640
	Post-intervention	8	3	8	2	4925	.992
Cognitive engagement	Day 1	9	1.25	9	1	3937.5	.228
	Day 3	9	2	8	2.75	4298	.565
	Day 7	8.5	2.25	9	2	4145.5	.808
	Post-intervention	8	3.25	9	2	4649.5	.485
Affective engagement	Day 1	9	1.25	8	2	3678.5	.056
	Day 3	8	2	8	2	4193.5	.394
	Day 7	8	3	8	2.75	4107	.727
	Post-intervention	8	4	7.5	4	4554.5	.351

Figure 2

Mean Anxiety Scores for Groups (Personalization) throughout the Intervention.

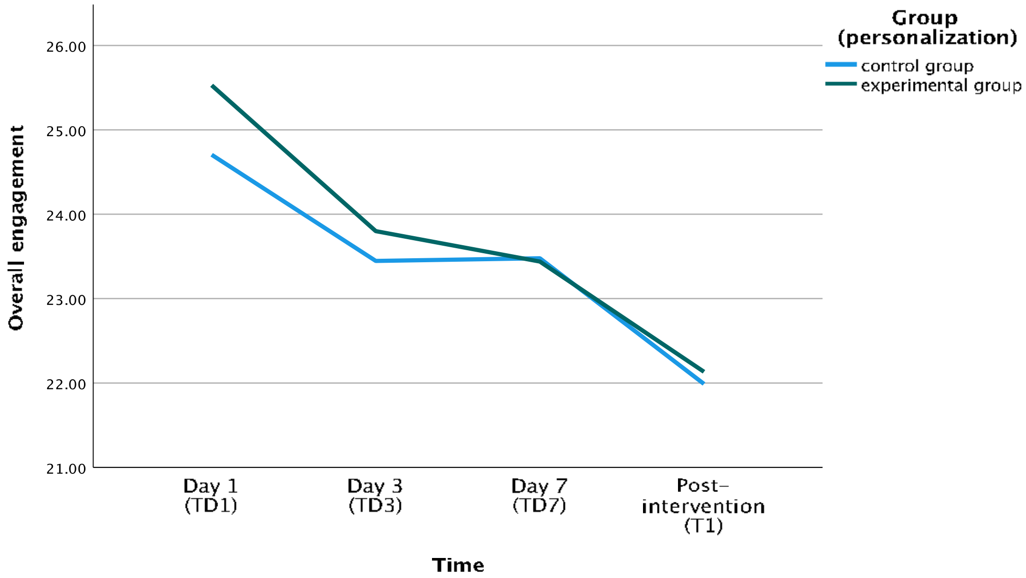


Effect of Personalization on Engagement (RQ2)

According to the generalized linear mixed model, there is no significant effect of personalization on overall $F(1, 749) = .554$, $p = .457$, behavioural $F(1, 749) = .026$, $p = .872$, cognitive $F(1, 749) = .269$, $p = .604$, nor affective engagement $F(1, 749) = 1.38$, $p = .240$. Nonetheless, there was an approach to significance regarding differences in overall engagement on intervention day 1 between the two groups ($U = 3628$, $p = .057$, $r = -1.9$), which may be due to the marginally significant differences in affective engagement ($U = 3678.5$, $p = .056$, $r = -1.91$), with the experimental group having scored higher (see Table 4). A significant effect of time was found, with a progression in the intervention being linked with a decreased overall $F(3, 749) = 23.15$, $p < .001$, behavioural $F(3, 749) = 9.77$, $p < .001$, cognitive $F(3, 749) = 11.41$, $p < .001$, and affective engagement $F(3, 749) = 27.31$, $p < .001$ (see Figure 3).

Figure 3

Mean Overall Engagement Scores for Groups (Personalization) throughout the Intervention.



Engagement Mediating the Relationship between Personalization and Anxiety (RQ3)

The 95% confidence interval entailed a 0 (-.49, .21), indicating overall engagement not to mediate the relationship between personalization and anxiety (B = .19, SE = .58, p = .745) (path c'1). In line with indications from bivariate correlations, high overall engagement was associated with lower anxiety scores at post-intervention (B = -.19, SE = .06, p = .001) (path b1). Comparable results were found for the other mediators, namely behavioral, congitive, and affective engagement (see Table 5 and Figure 4). Hence, engagement was not found to mediate the link between receiving personalized treatment and anxiety symptoms after the intervention.

Table 5

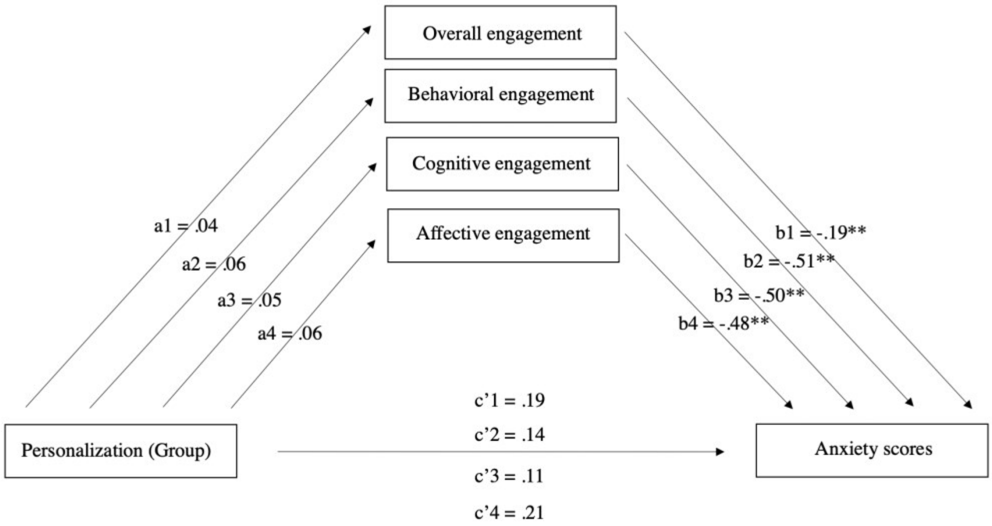
Mediation Analysis with General, Behavioral, Cognitive, and Affective Engagement mediating the relationship between Personalization (Group) and Anxiety at Post-Intervention.

Mediator	*a*	*b*	total effect *c*	indirect effect *a x b* (CI 95%)	direct effect *c′*
General engagement	.04	-.19**	.11	-.08 (-.49, .21)	.19
Behavioural engagement	.06	-.51**	.11	-.03 (-.42, .24)	.14
Cognitive engagement	.05	-.50**	.11	-.04 (-.44, .25)	.11
Affective engagement	.29	-.48**	.11	-.15 (-.58, -.14)	.21

Note. Bootstrap correction for indirect effect a x b (5000 resamples). *p < .05, **p <.01

Figure 4

Mediation Model.



Feedback by Human Counselor and Anxiety (RQ4)

The generalized linear mixed model indicated a significant effect of feedback groups $F(2, 529) = 5.278$, $p = .005$. While participants receiving human counselor video feedback had the lowest anxiety scores and significant differences between the feedback groups' anxiety scores were found at post intervention ($\chi^2(2) = 11.34$, $p = .003$) and follow up ($\chi^2(2) = 4.81$, $p = .091$), findings may have resulted from already significant differences at baseline ($\chi^2(2) = 6.73$, $p = .035$) (see Table 6 and Figure 5). In line with findings from RQ1, time had a significant effect on anxiety levels $F(2, 529) = 24.474$, $p < .001$, with all three groups showing decreased anxiety at post-intervention, followed by an increase at follow-up (see Figure 5). No interaction effect of feedback group*time $F(4, 529) = 1.681$, $p = .153$ was found.

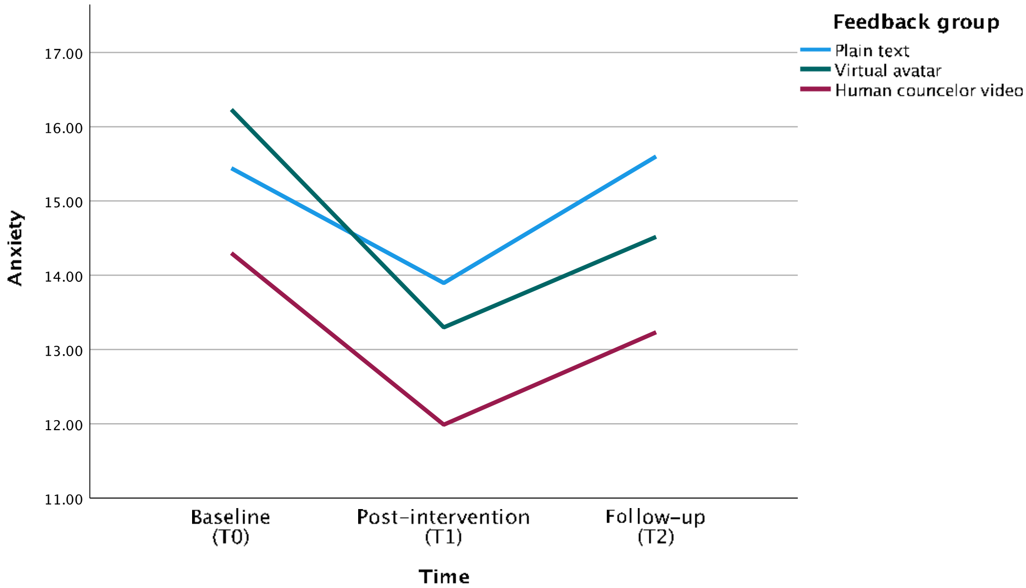
Table 6

Median Post-Intervention Anxiety Scores for Feedback groups, Kruskall-Wallis H test.

Variable	Measurement point	Feedback group							
		Plain Text		Virtual Avatar		Human Counsellor Video		Kruskal-Wallis H	
		Mdn	*IQR*	*Mdn*	*IQR*	*Mdn*	*IQR*	χ^2	*p*
Anxiety	Baseline (T0)	14.5	5.25	17	7.25	13	5	6.73	.035
	Post-intervention (T1)	13	4	13	3.5	12	4	11.34	.003
	Follow-up (T2)	14.5	8.5	14	6.25	12	7	4.81	.091

Figure 5

Mean Anxiety Scores for Feedback Groups throughout the Intervention.



Prescripted Video Feedback by Human Counselor and Overall Engagement (RQ 5)

Descriptive statistics illustrate that all three feedback groups experienced a decrease in engagement throughout the intervention (see Table 7, Figure 6). While the human counsellor feedback group's overall engagement score had slightly increased again on intervention day 7, it decreased again thereafter. Nonetheless, the Kruskal-Wallis H test showed no significant differences in overall engagement scores between the three feedback groups (see Table 7).

Table 7

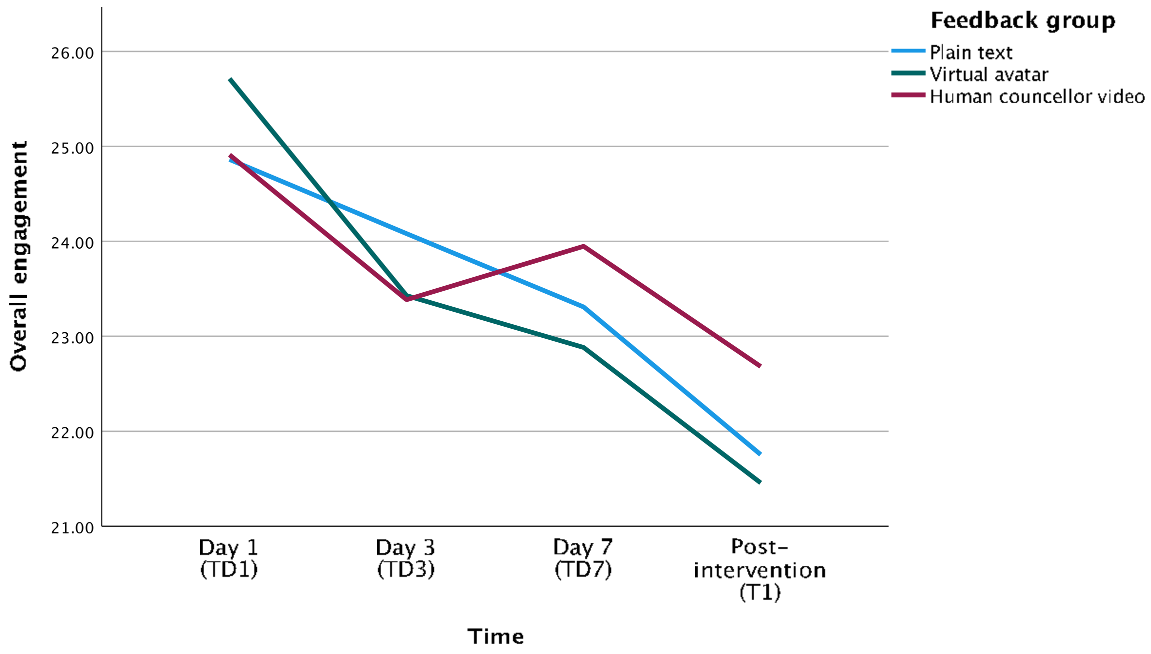
Median Overall Engagement Scores for Feedback groups.

Variable	Measurement point	Feedback group							
		Plain Text		Virtual Avatar		Human Counsellor Video		Kruskal-Wallis H	
		Mdn	*IQR*	*Mdn*	*IQR*	*Mdn*	*IQR*	χ^2	*P*
Engagement	Day 1 (TD1)	25.5	4	26	3.75	25	4	.699	.705
	Day 3 (TD3)	25	5.5	24	6	25	6.5	1.016	.602
	Day 7 (TD7)	25	6	24.5	6	24	6.5	.528	.768
	Post- Intervention (T1)	25	10.75	24	7.5	24	8	.919	.632

Note. For Plain Text N = 68, Virtual Avatar N = 65, Human Counsellor Video N = 97.

Figure 6

Mean Overall Engagement Scores for Feedback Groups throughout the Intervention.



Discussion

Results indicate that people with a personalized 2-week DHI do not experience higher reductions in anxiety symptoms, nor increased engagement than those with a non-personalized intervention. In line with this, engagement did not mediate the relationship between personalization and anxiety symptoms at post-intervention. Meanwhile, results suggest that participants with a higher general, behavioural, cognitive, and affective engagement experience lower anxiety levels post-intervention compared to those with low engagement scores. Overall, participants' anxiety scores decreased significantly by post-intervention, but increased again by follow-up. While those with human councelor video feedback had lower anxiety scores at all time points, findings are most likely due to already significant group differences at baseline. Lastly, engagement scores decreased throughout the intervention, and were not different between the three feedback groups.

Principal Results and Practical Implications

The main finding was that there were no significant differences in engagement scores between the personalized and non-personalized intervention. This was unexpected as a personalized intervention was assumed to be linked with higher engagement levels based on previous indications in literature (Burley et al., 2020; Kelders et al., 2020a, Kelders et al., 2020b). For instance, in a study by Couper et al. (2010), participants who received a tailored

DMHI to promote the consumption of fruit and vegetables were more engaged than those without a tailored intervention. To the writer's knowledge, this was the first quantitative study regarding the relationship of personalization and DMHI engagement (Burley et al., 2020).

However, multiple study findings point to the suggestion that insignificant findings may be due to a lacking personalization approach. First, near significant differences in engagement scores on intervention day 1 were found, with the experimental group scoring higher for overall and affective engagement than the control group. Significant results would suggest users to enjoy a DMHI, which is personalized to their wishes and preferences, more and feel like it fits them more as a person. These results support suggestions of engagement being a mechanism through which personalization works (Burley et al., 2020) and that the personalization was insufficient. Secondly, engagement decreased for both groups throughout the intervention. This substantiates the indication by Yardley et al. (2016) that maintaining engagement is difficult in self-help DMHIs without human support, but also implies that the personalization approach could not counter their absence. Meanwhile, the approach could still be useful if improved, as based on study results, a high user engagement throughout a 2-week-intervention may be predictive of lower anxiety levels post-intervention, and even 5-6 weeks after the end of the intervention. This matches indications of a positive link between engagement and DMHI's effectiveness (Kelders et al., 2020; Alkhaldi et al., 2017). Thirdly, the experimental group did not show significantly larger decreases in anxiety scores, which was unexpected. For instance, Fisher and Boswell (2016) found personalization to increase the intervention's efficacy in improving anxiety. Yet, the personalization approach substantially deviated, as Fisher and Boswell (2016) collected data for 30 days to design tailored treatment plans.

Considering these findings and literature, engagement may be an indispensable aspect for intervention efficacy and mechanism through which personalization works. The inability to increase the experimental group's engagement suggests that the personalization approach did not work, which limits the ability to make conclusions on the efficacy of personalization.

Multiple factors may explain why the used personalization approach was lacking. First, though the experimental groups' ITF options matched their expected engagement, they were fixed and displayed in the same manner to all. Hence, the system was not truly tailored to the "characteristics, preferences, interests, and needs of users", which is the aim of personalization, and may have led to a low interest and engagement of the users (Kocaballi et al., 2019). Connected to that, while CBT was a *content* option and is a promising treatment for anxiety (Lehtimaki et al., 2020), there are many sub-diagnoses, like social and general anxiety, requiring different treatment approaches and exercises (Bertie & Hudson, 2021). For instance, in the Challenger App by Miloff et al. (2015), users specified their anxiety triggers and names of family and friends, which were used to personalize challenges. Hence, the rather general factors may not have accounted for individual user needs, making the intervention unpersonal (Lattie et al., 2019). This may have led the experimental group not to be significantly more engaged and not to experience larger improvements in anxiety.

Second, the creation of the control group's intervention versions may have been flawed. While several participants randomly received an intervention which matched their expected engagement at baseline, a considerable number got at least one or two of their suiting *options*. This may have skewed results as it decreased the difference in the degree of personalization between the two groups. Lastly, while participants experienced anxiety scores above the cutoff, indicating mild anxiety (Williams, 2014), respondents may nonetheless have participated in the study not because they suffered under their anxiety and had planned to start therapy, but to receive SONA points. Hence, the target group may not have been intrinsically motivated to change, and with motivation being proposed as a main criterion for engagement in mental health treatment (Jochems et al., 2012), no intervention may have been able to engage them.

Another finding was that participants' anxiety scores had significantly decreased after the intervention. While the systematic review by Grist et al. (2017) did not find apps to improve anxiety in young people, study results support findings of a RCT by Bendtsen et al. (2020), who concluded their DMHI to have a protective effect on university students' anxiety symptoms. However, the samples' anxiety scores increased again between post-intervention and follow-up and were even statistically equivalent to baseline levels. This adds to the systematic review by Välimäki et al. (2017). Two studies performed follow-ups after 3-5 months, which, however, also no longer yielded improvements in anxiety symptoms (Välimäki et al., 2017). Regarding the little research of DMHIs' long-term effectiveness (Lehtimaki et al., 2021; Välimäki et al., 2017), the present study adds to literature with the practical implication that while DMHIs may decrease anxiety, improvements may not be sustainable. However, as the control group also received an intervention, study findings cannot indicate whether observed changes in anxiety scores were due to the received intervention or confounding variables.

Lastly, results illustrate that participants with human counselor video feedback had significantly lower anxiety scores at post-intervention and follow-up than those receiving virtual avatar and plain text feedback. However, significant differences were already present at baseline, hence, the groups were not suitable for comparison. Nonetheless, findings may also have been skewed due to uneven group sizes, with most users receiving human counselor feedback. This has a practical implication, as over half of participants in the experimental group chose this feedback method, reflecting it to be the option users expected to be engaged with the most. This matches indications in literature regarding users' preference to have some sort of personal contact with their therapist (Pretorius et al., 2019). If this suggestion is verified by future research, it would imply to incorporate users' wish of obtaining human counselor feedback in the *value specification* and *design* stages of the CeHRes roadmap to improve the match between the technology features and users' needs (Kip & van Gemert-Pijnen, 2018).

Strengths and Limitations

A strength of this design was the usage of a randomized control trial (RCT), as well as a factorial design. First, literature uses the term "black box" of eHealth to refer to limited insights into the technology-user-interaction (Black et al., 2011). This RCT helped to "open the black box" by depicting time-varying effects through measuring engagement and anxiety over a time span (Chen et al., 2022). Second, the factorial design enabled the creation of 27 different interventions and numerous possibilities to tailor the intervention to the users, increasing the chance of a good intervention-user-match (Gunst & Mason, 2009).

One methodological shortcoming concerns the participant IDs, on which responses from different time points were matched. First, the system allowed different participants to create the same code at baseline. Second, some participants made spelling mistakes when reporting their IDs in post-, and follow-up questionnaires. While IDs were matched manually in case of letter case discrepancies, several participants made more severe mistakes, where responses could not be matched and were lost. Next, with a rather healthy sample, the generalizability to clinical populations of university students is limited. Connected to that, as convenience sampling was used and participation was awarded with SONA points, this may not have recruited suitable participants who felt the need to improve their anxiety symptoms.

Future Research

First, regarding research limitations, future studies are suggested to automatically generate participant IDs to avoid dublicates, and to include a safety system which only allows proceeding to a questionnaire if participants indicated a valid ID. Secondly, it is advised to recruit participants who experience their anxiety as clinically relevant, such as students who are on the waiting list for psychological counselling at universities. Third, regarding the "black box", future research is advised to investigate how, why, and specifically for whom (Coughlin et al., 2017) a higher engagement may be associated with lower anxiety symptoms, which will followingly improve tailoring DMHIs to users and increase intervention efficacy.

Fourth, to improve personalization, participants' characteristics may be included to increase personal relevance (Kocaballi et al., 2019). Next, tailoring DMHIs based on algorithms has also been related to better intervention outcomes (Lustria et al., 2013), and could be used to adjust the content and exercises of upcoming modules based on individual progress. Furthermore, flexible factors are advised to be developed, for instance by using AI to customize feedback and to choose exercises based on users' anxiety type and triggers (Miloff et al., 2015).

Lastly, there was no indication to why engagement decreased throughout the intervention. There is much uncertainty how to stimulate and promote engagement (Kelders, 2019; Yardley et al., 2016). With engagement being a process (Kelders, 2020b), and this study having succeeded in getting users engaged, it vital to investigate its' maintenance and re-engagement. Exploratory research is suggested, such as a mixed study design with open questions in the TIIM app. This may give an explanation to why users felt less engaged (Chen et al., 2022) and what to incorporate in DMHIs to stabilize engagement. Future research is vital, as study findings showed higher engagement to be linked to lower anxiety scores at follow-up. Hence, anxiety scores may have stayed low at follow-up if engagement levels had been maintained throughout the intervention.

Conclusion

As DMHIs receive increasing attention with their potential to add to contemporary therapy, this study contributes to research through investigating the effect of personalization. While personalization was not found to be associated to engagement and anxiety symptoms, this may be due to methodological shortcomings, which indicate recommendations for future studies to improve personalization. Moreover, human counselor video feedback is suggested to be the feedback method users' expect to be most engaged with. Additionally, this research adds to literature with the novel finding that engagement may be linked to anxiety symptoms several weeks after completing the treatment. It is suggested to further investigate how to maintain engagement throughout DMHIs. By stabilizing engagement, intervention efficacy may be increased, leading to reduced anxiety levels amongst university students and lowering the overload of professional care.

References

Alkhaldi, G., Modrow, K., Hamilton, F. L., Davie, G., Ross, J., & Murray, E. (2017). Promoting Engagement With a Digital Health Intervention (HeLP-Diabetes) Using Email and Text Message Prompts: Mixed-Methods Study. *Interactive Journal of Medical Research*, *6*(2), e14. https://doi.org/10.2196/ijmr.6952

Alfons, A., Ates, N., & Groenen, P. J. F. (2019). A Robust Bootstrap Test for Mediation Analysis. *Organizational Research Methods*, *25*(3), 591–617. https://doi.org/10.1177/1094428121999096

Auerbach, R. P., Mortier, P., Bruffaerts, R., Alonso, J., Benjet, C., Cuijpers, P., Demyttenaere, K., Ebert, D. D., Green, J. G., Hasking, P., Murray, E., Nock, M. K., Pinder-Amaker, S., Sampson, N. A., Stein, D. J., Vilagut, G., Zaslavsky, A. M., & Kessler, R. C. (2018). WHO World Mental Health Surveys International College Student Project: Prevalence and distribution of mental disorders. *Journal of Abnormal Psychology*, *127*(7), 623–638. https://doi.org/10.1037/abn0000362

Bendtsen, M., Müssener, U., Linderoth, C., & Thomas, K. (2020). A Mobile Health Intervention for Mental Health Promotion Among University Students: Randomized Controlled Trial. *JMIR Mhealth Uhealth*, *8*(3), e17208. https://doi.org/10.2196/17208

Benton, S. A., Heesacker, M., Snowden, S. L., & Lee, G. (2016). Therapist-assisted, online (TAO) intervention for anxiety in college students: TAO outperformed treatment as usual. *Professional Psychology: Research and Practice*, *47*(5), 363–371. https://doi.org/10.1037/pro0000097

Bertie, L. A., & Hudson, J. L. (2021). CBT for Childhood Anxiety: Reviewing the State of Personalised Intervention Research. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.722546

Black, A., Car, J., Pagliari, C., Anandan, C., Cresswell, K., Bokun, T., McKinstry, B., Procter, R., Majeed, A., & Sheikh, A. (2011). The Impact of eHealth on the Quality and Safety of Health Care: A Systematic Overview. *PLOS Medicine*, *8*(1), e1000387. https://doi.org/10.1371/journal.pmed.1000387

Bolinski, F., Boumparis, N., Kleiboer, A., Cuijpers, P., Ebert, D., & Riper, H. (2020). The effect of e-mental health interventions on academic performance in university and college students: A meta-analysis of randomized controlled trials. *Internet Interventions*, *20*, 100321. https://doi.org/10.1016/j.invent.2020.100321

Bry, L. J., Chou, T., Miguel, E. C., & Comer, J. S. (2017). Consumer Smartphone Apps Marketed for Child and Adolescent Anxiety: A Systematic Review and Content Analysis. *Behavior Therapy*, *49*(2), 249–261. https://doi.org/10.1016/j.beth.2017.07.008

Burley, C., Anderson, D., Brownlee, A., Lafer, G., Luong, T., McGowan, M., Nguyen, J., Trotter, W. T., Wine, H., Baglione, A., & Barnes, L. E. (2020). Increasing Engagement in eHealth Interventions Using Personalization and Implementation Intentions. https://doi.org/10.1109/sieds49339.2020.9106640

Chen, Y., Duku, E., & Georgiades, S. (2022). Rethinking Autism Intervention Science: A Dynamic Perspective. *Frontiers in Psychiatry*, *13*. https://doi.org/10.3389/fpsyt.2022.827406

Clarke, A. M., Kuosmanen, T., & Barry, M. M. (2014). A Systematic Review of Online Youth Mental Health Promotion and Prevention Interventions. *Journal of Youth and Adolescence*, *44*(1), 90–113. https://doi.org/10.1007/s10964-014-0165-0

Coughlin, S. S., Williams, L. B., & Hatzigeorgiou, C. (2017). A systematic review of studies of web portals for patients with diabetes mellitus. *mHealth*, *3*, 23. https://doi.org/10.21037/mhealth.2017.05.05

Couper, M. P., Alexander, G. L., Zhang, N., Little, R. J. A., Maddy, N., Nowak, M. A., McClure, J. B., Calvi, J., Rolnick, S. J., Stopponi, M. A., & Johnson, C. C. (2010). Engagement and Retention: Measuring Breadth and Depth of Participant Use of an Online Intervention. *Journal of Medical Internet Research*, *12*(4), e52. https://doi.org/10.2196/jmir.1430

Chang, J. J., Ji, Y., Li, Y. H., Pan, H. F., & Su, P. Y. (2021). Prevalence of anxiety symptom and depressive symptom among college students during COVID-19 pandemic: A meta-analysis. *Journal of Affective Disorders*, *292*, 242–254. https://doi.org/10.1016/j.jad.2021.05.109

Fan, H., & Poole, M. S. (2006). What Is Personalization? Perspectives on the Design and Implementation of Personalization in Information Systems. *Journal of Organizational Computing and Electronic Commerce*, *16*(3), 179–202. https://doi.org/10.1207/s15327744joce1603

Fisher, A. J., & Boswell, J. (2016). Enhancing the Personalization of Psychotherapy With Dynamic Assessment and Modelling. *Assessment*, *23*(4), 496–506. https://doi.org/10.1177/1073191116638735

Fleming, T., Bavin, L. M., Lucassen, M., Stasiak, K., Hopkins, S., & Merry, S. E. (2018). Beyond the Trial: Systematic Review of Real-World Uptake and Engagement With Digital Self-Help Interventions for Depression, Low Mood, or Anxiety. *Journal of Medical Internet Research*, *20*(6), e199. https://doi.org/10.2196/jmir.9275

Grist, R., Porter, J. C., & Stallard, P. (2017). Mental Health Mobile Apps for Preadolescents and Adolescents: A Systematic Review. *Journal of Medical Internet Research*, *19*(5), e176. https://doi.org/10.2196/jmir.7332

Gunst, R. F., & Mason, R. L. (2009). Fractional factorial design. *Wiley Interdisciplinary Reviews: Computational Statistics*, *1*(2), 234–244. https://doi.org/10.1002/wics.27

Gupta, S. K. S. (2011). Intention-to-treat concept: A review. *Perspectives in Clinical Research*, *2*(3), 109. https://doi.org/10.4103/2229-3485.83221

Hall, B. J., Xiong, P., Guo, X., Sou, E. K. L., Chou, U. I., & Shen, Z. (2018). An evaluation of a low intensity mHealth enhanced mindfulness intervention for Chinese university students: A randomized controlled trial. *Psychiatry Research-neuroimaging*, *270*, 394–403. https://doi.org/10.1016/j.psychres.2018.09.060

Harrer, M., Adam, S. H., Baumeister, H., Cuijpers, P., Karyotaki, E., Auerbach, R. P., Kessler, R. C., Bruffaerts, R., Berking, M., & Ebert, D. D. (2018). Internet interventions for mental health in university students: A systematic review and meta-analysis. *International Journal of Methods in Psychiatric Research*, *28*(2), e1759. https://doi.org/10.1002/mpr.1759

Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford publications.

Higa-McMillan, C. K., Francis, S. L., Rith-Najarian, L. R., & Chorpita, B. F. (2016). Evidence Base Update: 50 Years of Research on Treatment for Child and Adolescent Anxiety. *Journal of Clinical Child & Adolescent Psychology*, *45*(2), 91–113. https://doi.org/10.1080/15374416.2015.1046177

Hollis, C., Falconer, C. J., Martin, J. L., Whittington, C., Stockton, S., Glazebrook, C., & Davies, E. B. (2016). Annual Research Review: Digital health interventions for children and young people with mental health problems - a systematic and meta-review. *Journal of Child Psychology and Psychiatry*, *58*(4), 474 503. https://doi.org/10.1111/jcpp.12663

Jochems, E. C., Mulder, C. L., Van Dam, A., Duivenvoorden, H. J., Scheffer, S. C. M., Van Der Spek, W., & Van Der Feltz-Cornelis, C. M. (2012). Motivation and treatment engagement intervention trial (MotivaTe-IT): the effects of motivation feedback to clinicians on treatment engagement in patients with severe mental illness. *BMC Psychiatry*, *12*(1). https://doi.org/10.1186/1471-244x-12-209

Jones, J. D., Hankin, B. L., Gallop, R., Haraden, D., Sbrilli, M. D., Garber, J., & Young, J. F. (2022). Effects of personalized depression prevention on anxiety through 18-month follow-up: A randomized controlled trial. *Behaviour Research and Therapy*, *156*, 104156. https://doi.org/10.1016/j.brat.2022.104156

Kelders, S. M. (2015). Involvement as a Working Mechanism for Persuasive Technology. In *Lecture Notes in Computer Science* (pp. 3–14). Springer Science Business Media. https://doi.org/10.1007/978-3-319-20306-5_1

Kelders, S. M., Kip, H., & Greeff, J. (2020a). Psychometric evaluation of the TWente Engagement with Ehealth Technologies Scale (TWEETS): evaluation study. *Journal of medical internet research*, *22*(10), e17757. https://doi.org/10.2196/17757

Kelders, S. M., Van Zyl, L. E., & Ludden, G. D. (2020b). The concept and components of engagement in different domains applied to ehealth: a systematic scoping review. *Frontiers in psychology*, *11*, 926. https://doi.org/10.3389/fpsyg.2020.00926

Kip, H., & van Gemert-Pijnen, L. J. (2018). Holistic development of eHealth technology. In Gemert Pijnen, V. L., Kelders, S. M., Kip, H., & Sanderman, R. (editors) *eHealth Research, Theory and Development: A Multi-Disciplinary Approach.* Abingdon-on Thames, UK: Routledge, 152-187.

Kocaballi, A. B., Berkovsky, S., Quiroz, J. C., Laranjo, L., Tong, H. L., Rezazadegan, D., Briatore, A., & Coiera, E. (2019). The Personalization of Conversational Agents in Health Care: Systematic Review. *Journal of Medical Internet Research*, *21*(11), e15360. https://doi.org/10.2196/15360

Lattie, E. G., Adkins, E. C., Winquist, N., Stiles-Shields, C., Wafford, Q. E., & Graham, A. K. (2019). Digital Mental Health Interventions for Depression, Anxiety, and

Enhancement of Psychological Well-Being Among College Students: Systematic Review. *Journal of Medical Internet Research*, *21*(7), e12869. https://doi.org/10.2196/12869

Lal, S. (2019). E-mental health: Promising advancements in policy, research, and practice. *Healthcare Management Forum*, *32*(2), 56–62. https://doi.org/10.1177/0840470418818583

Lehtimaki, S., Martic, J., Wahl, B., Foster, K. T., & Schwalbe, N. (2021). Evidence on Digital Mental Health Interventions for Adolescents and Young People: Systematic Overview. *JMIR Mental Health*, *8*(4), e25847. https://doi.org/10.2196/25847

Lustria, M. L. A., Noar, S. M., Cortese, J., Van Stee, S. K., Glueckauf, R. L., & Lee, J. (2013). A Meta-Analysis of Web-Delivered Tailored Health Behavior Change Interventions. *Journal of Health Communication*, *18*(9), 1039–1069. https://doi.org/10.1080/10810730.2013.768727

Miloff, A., Marklund, A., & Carlbring, P. (2015). The challenger app for social anxiety disorder: New advances in mobile psychological treatment. *Internet Interventions*, *2*(4), 382–391. https://doi.org/10.1016/j.invent.2015.08.001

Olthuis, J. V., Watt, M. C., Bailey, K. M., Hayden, J. A., & Stewart, S. H. (2016). Therapist-supported Internet cognitive behavioural therapy for anxiety disorders in adults. *The Cochrane Library*, *2016*(3). https://doi.org/10.1002/14651858.cd011565.pub2

Osborn, T. G., Li, S., Saunders, R., & Fonagy, P. (2022). University students' use of mental health services: a systematic review and meta-analysis. *International Journal of Mental Health Systems*, *16*(1). https://doi.org/10.1186/s13033-022-00569-0

Pretorius, C., Chambers, D., & Coyle, D. (2019). Young People's Online Help-Seeking and Mental Health Difficulties: Systematic Narrative Review. *Journal of Medical Internet Research*, *21*(11), e13873. https://doi.org/10.2196/13873

Priestley, M., Broglia, E., Hughes, G., & Spanner, L. (2021). Student Perspectives on improving mental health support Services at university. *Counselling and Psychotherapy Research*, *22*(1). https://doi.org/10.1002/capr.12391

Riboldi, I., Cavaleri, D., Calabrese, A., Capogrosso, C. A., Piacenti, S., Bartoli, F., Crocamo, C., & Carrà, G. (2022). Digital mental health interventions for anxiety and depressive symptoms in university students during the COVID-19 pandemic: A systematic review of randomized controlled trials. *Revista De Psiquiatría Y Salud Mental*. https://doi.org/10.1016/j.rpsm.2022.04.005

Rodríguez-Rivas, M. E., Cangas, A. J., Cariola, L. A., Varela, J. J., & Valdebenito, S. (2021). Innovative Technology–Based Interventions to Reduce Stigma Toward People With Mental Illness: Systematic Review and Meta-analysis. *JMIR Serious Games*, *10*(2), e35099. https://doi.org/10.2196/35099

Torous, J., Nicholas, J., Larsen, M. E., Firth, J., & Christensen, H. (2018). Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements. *Evidence Based Mental Health*, *21*(3), 116–119. https://doi.org/10.1136/eb-2018-102891

Van Gemert-Pijnen, L. J., Kip, H., Kelders, S. M., & Sanderman, R. (2018). Introducing ehealth. In *eHealth research, theory, and development* (pp. 3-26). Routledge.

Van Gemert-Pijnen, J. E., Nijland, N., Van Limburg, M., Ossebaard, H. C., Kelders, S. M., Eysenbach, G., & Seydel, E. R. (2011). A Holistic Framework to Improve the Uptake and Impact of eHealth Technologies. *Journal of Medical Internet Research*, *13*(4), e111. https://doi.org/10.2196/jmir.1672

Van Orden, M., Kraaijeveld, J. C., Spijker, A., Silven, A. V., Bonten, T., Chavannes, N. H., & Van Dijke, A. (2022). Preliminary effects of a digital mental health intervention for depression and anxiety. *Clinical eHealth*, *5*, 44–51. https://doi.org/10.1016/j.ceh.2022.06.002

Välimäki, M., Anttila, K., Anttila, M., & Lahti, M. (2017). Web-Based Interventions Supporting Adolescents and Young People With Depressive Symptoms: Systematic Review and Meta-Analysis. *JMIR MHealth and UHealth*, *5*(12), e180. https://doi.org/10.2196/mhealth.8624

Williams, N. (2014). The GAD-7 questionnaire. *Occupational medicine*, *64*(3), 224- 224. https://doi.org/10.1093/occmed/kqt161

World Health Organization: WHO. (2022, March 2). COVID-19 pandemic triggers 25% increase in prevalence of anxiety and depression worldwide. *World Health Organization*. https://www.who.int/news/item/02-03-2022-covid-19-pandemic-triggers-25-increase-in-prevalence-of-anxiety-and-depression-worldwide

WHO. (2019). WHO Guideline: Recommendations on Digital Interventions for Health System Strengthening. Geneva.

Yardley, L., Spring, B., Riper, H., Morrison, L., Crane, D., Curtis, K., Merchant, G., Naughton, F., & Blandford, A. (2016). Understanding and Promoting Effective Engagement With Digital Behavior Change Interventions. *American Journal of Preventive Medicine*, *51*(5), 833–842. https://doi.org/10.1016/j.amepre.2016.06.015

Appendix A

The Adjusted TWEETS to measure Engagement at Post-Intervention.

Item	Thinking about using this app the last week, I feel that:
1	using this app did become part of my daily routine
2	the app took me little effort to use
3	I was able to use the app as often as needed to improve my anxiety
4	this app made it easier for me to work on decreasing my anxiety
5	this app motivated me to decrease my anxiety
6	this app helped me to get more insight into my anxiety
7	I enjoyed using this app
8	I enjoyed seeing the progress I made in this app
9	this app fit me as a person

Appendix B

The TWEETS by Kelders et al. (2020a).

Item	Thinking about using this app the last week, I feel that:	Construct
1	this app is part of my daily routine	Behavior
2	this app takes me little effort to use	Behavior
3	I'm able to use this app as often as needed to improve my anxiety	Behavior
4	this app makes it easier for me to work on decreasing my anxiety	Cognition
5	this app motivates me to decrease my anxiety	Cognition
6	this app helps me to get more insight into my anxiety	Cognition
7	I enjoy using this app	Affect
8	I enjoy seeing the progress I make in this app	Affect
9	this app fits me as a person	Affect

Appendix C

Variable	Shapiro-Wilk Test		
	Statistic	*df*	*Sig.*
Age	.960	141	
General engagement Day 1	.960	141	<.001
General engagement Day 3	.957	141	<.001
General engagement Day 7	.951	141	<.001
General engagement Post-intervention (T1)	.956	141	<.001
Behavioral engagement Day 1	.935	141	<.001
Behavioral engagement Day 3	.942	141	<.001
Behavioral engagement Day 7	.940	141	<.001
Behavioral engagement Post-intervention (T1)	.939	141	<.001
Cognitive engagement Day 1	.932	141	<.001
Cognitive engagement Day 3	.945	141	<.001
Cognitive engagement Day 7	.935	141	<.001
Cognitive engagement Post-intervention (T1)	.916	141	<.001
Affective engagement Day 1	.939	141	<.001
Affective engagement Day 3	.948	141	<.001
Affective engagement Day 7	.949	141	<.001
Affective engagement Post-intervention (T1)	.961	141	<.001
Anxiety Baseline (T0)	.946	107	.001
Anxiety Post-intervention (T1)	.968	107	.000
Anxiety Follow-up (T2)	.941	107	.001