



A Smart Solution? Efficacy of a Self-Management Smartphone App for Panic Disorder and Agoraphobia

Justine Spies¹ · Thomas Lang^{2,3} · Sylvia Helbig-Lang¹

Accepted: 9 July 2024

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2024

Abstract

Background Cognitive behavioral therapy has proven to be effective in the treatment of panic disorder and/or agoraphobia (PD/A). However, psychotherapy is often not available in a timely manner. Hence, there is a need for alternative and easily accessible interventions. The current study, thus, aimed to evaluate a smartphone-based self-management-tool for panic disorder and/or agoraphobia.

Method 107 participants were randomly assigned to either an intervention group (app “Mindable”) or a waitlist control group. Primary outcome was panic and agoraphobic symptom severity, further outcomes included perceived quality of life, anxiety control beliefs, and functional impairment. Linear mixed models were used to analyze the data in intent-to-treat- and completer-data sets.

Results After the eight-week observation period, participants in the intervention group reported significantly stronger reductions in panic and agoraphobic symptoms ($R^2=0.20$) and a higher increase in anxiety-related control beliefs ($R^2=0.125$) than participants in the control condition. Despite this symptom reduction, only 4.3% of participants in the app group and none in the control group were considered as remitted. Quality of life and functional impairment improved in all participants over time, but no significant group x time interactions were found.

Discussion A smartphone-based self-management tool is an easily accessible intervention to reduce panic and agoraphobic symptoms and enhance anxiety control in people with PD/A. As no symptom remission could be achieved, this approach may be suitable as part of a stepped care approach for anxiety disorders.

Trial Registration Registration number: DRKS00029090, registered on 10.06.2022.

Keywords Smartphone-based · Panic disorder · Agoraphobia · Cognitive behavioral therapy

Introduction

Panic disorder and agoraphobia (PD/A) are very common and tend to become chronic if untreated (Jacobi et al., 2014). These anxiety disorders often cause considerable impairments in patients’ social, work, and private life (Skapinakis et al., 2011). Research has shown that cognitive behavioral

therapy (CBT) is effective in the treatment of PD/A (e.g. Bandelow et al., 2021; NICE, 2020). However, numerous patients remain untreated due to long waiting times for psychotherapy, fear of being stigmatized, and differences in geographical proximity to mental health care institutions (Neutens, 2015; Villatoro et al., 2022). Hence, there is a great need for alternative treatments which are easy to access. Recently, digital health applications (DHA) have gained popularity as an alternative approach to treating people with various mental disorders, including anxiety disorders. DHA are medical applications, which can be prescribed by therapists or physicians, and which are implemented by means of a computer program, a browser-based web-program, or a smartphone-based app (Maaß et al., 2022). Meta-analyses have shown that DHA are well accepted and effective in the treatment of various mental problems, including anxiety disorders (Andersson et al., 2014; Andrews et al., 2010, 2018).

✉ Justine Spies
spies@pta-hamburg.de

¹ PTA Hamburg, Psychotherapieausbildung an der Universität Hamburg, Amsinckstraße 41, 20097 Hamburg, Germany

² Christoph-Dornier-Stiftung, Institut Bremen, Grazer Straße 2b, 28359 Bremen, Germany

³ Constructor University Bremen gGmbH, Campus Ring 1, 28759 Bremen, Germany

A more detailed analysis of specific effects of such interventions for panic and agoraphobic symptoms is provided in a more recent review and meta-analysis (Stech et al., 2020). The authors summarized that DHA outperformed waitlist control groups and psychoeducational control groups, both for panic and agoraphobic symptoms with higher effect sizes for panic symptoms (Stech et al., 2020). Results from three studies even suggested that DHA might be as effective as face-to-face treatment for PD/A, and that these effects are stable. Even more encouraging, Pauley et al. (2023) calculated a Number needed to treat in DHA for PD/A with 1.8, suggesting that half of all persons treated with DHA will benefit from treatment.

These findings are somewhat limited by a high heterogeneity of study quality (Stech et al., 2020) and by collapsing different kinds of DHA and hybrid interventions combining digital and therapeutic contact in one analysis. It is, thus, still difficult to conclude which kind of DHA might work best for anxiety disorders. As these disorders are often linked to considerable situational avoidance, DHA using smartphone-based interventions might be especially promising as treatment often incorporates exercises outside, and people are used to taking their smartphones with them wherever they go (de Vries et al., 2021; Heron & Smyth, 2010).

To date, smartphone-based interventions for anxiety yielded somewhat mixed results (see Firth et al., 2017). Positive effects were shown by Ivanova and colleagues (2016) who compared a waitlist control group with an Acceptance and Commitment Therapy (ACT) for social anxiety and panic disorder, delivered by laptop and smartphone. ACT outperformed the control group with an effect size of $d=0.39$ in the reduction of general anxiety symptoms. Interestingly, effects were comparable between participants who received additional therapeutic feedback and participants that used the digital intervention without further guidance, suggesting that the app as stand-alone treatment was effective. However, it has to be noted that no significant group differences were found when PD/A symptoms were analyzed separately. Another recent study by Ebenfeld and colleagues (2021) found a desktop- and smartphone-based CBT self-management intervention for PD/A to be superior in reducing PD/A symptoms compared to a waitlist control group. Here, participants received standardized written feedback and reminders from coaches to increase motivation and adherence.

While there is some evidence that DHA might be effective in reducing symptoms of panic disorder and agoraphobia, the relevance of additional therapeutic contact is still object of debate. While some studies have indicated that guided DHA were significantly more effective in reducing symptom severity (Domhardt et al., 2019) and led to higher adherence rates (Linardon et al., 2019) than

self-management interventions, Pauley et al. (2023) and Ivanova et al. (2016) found no significant difference between guided and unguided digital interventions for anxiety. With regard to the limited resources in health care, evidence-based interventions working without any therapist guidance might further help to improve health care for individuals suffering from PD/A who cannot be treated otherwise. In addition, relying on a smartphone app rather than relying on hybrid internet- and smartphone-interventions offers many advantages, such as smartphones being used much more frequently and, therefore, offering immediate engagement in exercises and monitoring of symptoms (Linardon et al., 2020). These advantages may also keep users more engaged (Heron & Smyth, 2010).

However, specific studies examining smartphone-based DHA as a self-management treatment in panic disorder and agoraphobia are missing. Therefore, the app “Mindable” was developed for providing self-help for people suffering from PD/A. App development was informed by an evidence-based cognitive behavioral treatment manual for panic disorder and agoraphobia (Lang et al., 2012), that emphasizes exposure exercises as a core active ingredient of CBT. The aforementioned treatment manual has demonstrated its effectiveness in a large multicenter RCT, examining two variants of treatment delivery: Exposure with and without therapeutic assistance (Gloster et al., 2011). Interestingly, both treatment variants led to stable reductions in anxiety symptoms, suggesting that an instructed exposure exercise without therapist-assistance in the field might yield positive effects. This finding encourages the assumption that a well described self-help instruction to exposure treatment might hold potential as a low-threshold alternative when psychotherapy is not available.

The development of the app “Mindable” was guided by the treatment manual used in the study of Gloster and colleagues, thus, attempting to translate an evidence-based treatment to a digital intervention. As studies have shown that adding cognitive interventions to exposure does not lead to a significant improvement of the treatment, the app “Mindable” does not include cognitive therapy (Lang et al., 2009) but focuses on psychoeducation, self-monitoring, and exposure to internal and external cues.

The current study aimed to evaluate the effects of a smartphone-based self-management delivery of an exposure-based treatment manual for panic disorder and agoraphobia. Notably, the treatment was realized without additional therapeutic contacts as a stand-alone self-help intervention. Based on previous studies, it was assumed that the app “Mindable” would be effective in reducing symptoms associated with panic disorder and agoraphobia. More specifically, it was hypothesized that using the app would result in higher reductions in panic and agoraphobic symptoms

as well as stronger improvements in anxiety-related control beliefs, quality of life, and functional impairment compared to a waitlist control group.

Method

Study Design

To evaluate the app's efficacy, we conducted a prospective multicentered two-armed RCT. Participants diagnosed with panic disorder and/ or agoraphobia were randomly assigned either to an intervention group (AG) who received the app "Mindable" for eight weeks or to a waitlist control group (CG) who received general information on self-help during waiting time for psychotherapy (Hähnel et al., 2004). Ethical approval for the study was granted by the ethical committee of the German Society for Psychology (reference number: LangThomas2020-12-14VA). The study was pre-registered with the Clinical Trial Registration (Registration number: DRKS00029090) and conducted in agreement with CONSORT guidelines.

Randomization

Randomization was conducted by a staff member not involved in the diagnostic processes. The assignment was stratified for the presence or absence of agoraphobia as well as for the study center, which conducted the initial diagnostic processes. Block randomization with $n = 20$ participants per block was used to achieve a balanced distribution. The random numbers were determined by the program "research randomizer" (www.randomizer.org). The staff members involved in the diagnostic processes were blinded to the participants' assignment, and baseline- and post-assessment were carried out by different staff members.

Participants

Participants were recruited from July 2020 until September 2022 via online-advertisements, social media, press releases, and during initial presentations at the outpatient treatment centers involved in the study (Hamburg, Bremen, Münster). Because of the Covid-Pandemic, recruitment and conduction of baseline- and post-assessments were switched from face-to-face to online in September 2021. Hence, participants could then be recruited from all over Germany.

Participants had to meet the DSM-5 diagnostic criteria of panic disorder and/ or agoraphobia which were assessed using a structured clinical interview and be at least 18 years of age. Participants were excluded, if they (a) were currently undergoing psychotherapy, (b) did not have a smartphone,

(c) had had a change in medication in the last two months or were taking benzodiazepines, (d) had a comorbid substance use or psychotic disorder, (e) had comorbid chronic respiratory or cardiovascular diseases, (f) were suicidal and (g) did not have sufficient German language abilities or illiteracy.

Treatment

Participants in the treatment condition received a standardized introduction to the app "Mindable" and were guided through the installation of the app on their individual smartphone. They received an activation code for the app directly from the researcher. Usually, the app "Mindable" has to be prescribed by physicians or psychotherapists who confirm that people meet the diagnostic criteria of PD/A. With the prescription, patients may ask their health insurance for an activation code that activates the app on their smartphone. For study purposes, the app was ready to use when installed.

The app includes a module "psychoeducation", comprising nine multimedia information lessons on the etiology and maintenance of panic disorder and agoraphobia. A second module "symptom provocation" informs about aims and procedures of interoceptive exposure and provides instructions for interoceptive exercises, such as hyperventilation or spinning. The third module "exposure" provides information about exposure in vivo as well as suggestions and protocols for specific exposure exercises. In both the "symptom provocation" and the "exposure" module, participants are first instructed to select a specific exercise out of given examples (e.g. spinning or riding a bus). It is also possible to create an individual exposure hierarchy and derive exercises. Users might start an exercise within the app and can record the course of the exercise within specific protocols. The app also provides reminders for planned exercises and charts examining anxiety curves within and across exercises. The app also includes a daily "diary" concerning anticipatory anxiety, panic attacks and avoidance behavior. Once a week, users are invited to give information on the current state of their symptoms in the "checkup".

For our study, participants were instructed to use the app on a daily basis over the period of eight weeks. Participants were informed that they could contact research staff in case of technical problems, but there was no further therapeutic contact. All app modules were accessible from the beginning, and participants were free to decide how they wanted to use the app. More detailed information on all app data points can be found in Appendix D.

The waitlist control group received a standardized leaflet with non-specific information on how to deal with waiting time for psychotherapy (recommendations of self-help literature and general recommendations, such as staying in contact with others, carrying out self-care activities, and

exercising). This leaflet was developed for another study examining effects of a minimal intervention during the waiting time for psychotherapy (Helbig & Hoyer, 2007).

Assessments

The baseline-assessment (T1) comprised the informed consent procedure, and – when a written consent was provided, a structured diagnostic interview as well as all outcome measures and additional symptom measures to further examine moderator and mediator variables. After four weeks, a between-assessment (T2) was conducted online in order to be able to analyze specific mediators or moderators in the future. Eight weeks after the baseline-assessment, the post-assessment (T3) took place, which comprised a structured diagnostic interview as well as the outcome measures and the additional symptom measures. After twelve weeks (T4), participants were asked to rate the outcome measures and the additional symptom measures again in an exploratory follow-up-assessment. The between-assessments are not included in the current analysis, as no moderator or mediator analyses are conducted. Follow-up data was only partially available, and waitlist control group participants had mostly started using the app “Mindable” by the time the follow-up-assessment took place compromising treatment integrity at follow-up. Therefore, these data were not used for further group comparisons. However, we provide descriptive follow-up data in Appendices E to G.

Treatment integrity was evaluated by means of an adherence score, which was defined a priori based on clinical judgements. To generate the adherence score, the recommended app use was defined for each app module and meeting these recommendations or even exceeding them resulted in an adherence score of 100%. Falling below the recommendations was also defined and thus, categorized in lower adherence scores. Firstly, three adherence scores were calculated for (1) the module “psychoeducation”, (2) the “checkup” and the “diary” and (3) the number of overall exercises in the modules “symptom provocation” and “exposure”. To reach a 100% adherence score, participants had to (1) complete nine out of nine psychoeducation lessons, (2) have at least eight entries in the “checkup” and “diary”, and (3) have conducted at least five exercises. Secondly, an average and overall adherence score was calculated out of the three adherence scores. Participants having an adherence score of $\geq 75\%$ were considered as adherent.

Diagnostic Status

The German version of the Diagnostic Interview for Mental Disorders (DIPS, Margraf et al., 2017) was used to evaluate whether participants met any exclusion or the inclusion

criteria in the baseline-assessment. The DIPS assesses all mental disorders according to the DSM-5. The interview was delivered in the centers or via certified video software (RED Medical, 2014) by trained members of staff with a diploma or master’s degree in Psychology and who were currently training to be psychotherapists.

Outcome Measures

Severity of Panic and Agoraphobic Symptoms

The primary outcome measure was the self-assessed severity of panic and agoraphobic symptoms by means of the German version of the Panic and Agoraphobia Scale (PAS, Bandelow, 2016). The PAS is useful in measuring treatment efficacy (Bandelow et al., 1998) and comprises five subscales (panic attacks, agoraphobic avoidance, anticipatory anxiety, impairment in social relationships and work, assumptions of somatic disease). The 13 items are rated on five-point Likert-scales. Total scores range between 0 and 52 with scores from 0 to 8 indicating *no clinically relevant symptoms*, scores from 9 to 28 indicating *moderate symptoms* and scores from 29 and above indicating *severe levels of symptoms*. Cronbach α for the total score was 0.84 in this study.

Quality of Life

Quality of life was assessed by means of the German version of the World Health Organization Quality of Life Scale (WHOQOL-BREF, Angermeyer et al., 2000). Its psychometric properties are highly satisfactory (Skevington et al., 2004). Participants are asked to rate 26 items which cover the four domains *physical health*, *psychological health*, *social relationships* and *environment*. For the current study, the domain *psychological health* was chosen as outcome-relevant (Cronbach $\alpha = 0.81$ in this study). Scores can range from 0 to 100.

Anxiety-Related Control Beliefs

Anxiety-related control beliefs were assessed by means of the German version of the Anxiety Control Questionnaire (ACQ, Helbig-Lang et al., 2012b). Studies have shown that anxiety-related control beliefs serve as a mediator for symptom change in psychotherapy of anxiety disorders. The ACQ is highly reliable (Brown et al., 2004). The 30 items are rated on a scale from 0 (“strongly disagree”) to 5 (“strongly agree”) and total scores range from 0 to 150. In the current study, Cronbach α was 0.81.

Functional Impairment

Perceived functional impairment was assessed by means of the German version of the Sheehan Disability Scale (SDS, Leon et al., 1992). The SDS consists of three items. Participants are asked to rate their impairment in professional, social and family settings. Impairment scores range from 0 (“not at all”), 1–3 (“mildly”), 4–6 (“moderately”), 7–9 (“markedly”) to 10 (“extremely”). The SDS is a well evaluated measure (Hodgins, 2013). Cronbach α was 0.69 in the current study.

User Satisfaction

User satisfaction was assessed by means of the German version of the Client Satisfaction Questionnaire adapted to Internet-based interventions (CSQ-I; Boß et al., 2016). It consists of eight statements with response scales ranging from 1 (*does not apply to me*) to 4 (*does totally apply to me*) concerning participants’ satisfaction with the app. Its psychometric properties are of good quality (Boß et al., 2016).

Statistical Analyses

Sample Size

The program G*Power (Faul et al., 2007) was used to calculate the sample size needed in the study based on the assumption of finding moderate effects ($d=0.6$) in the case of the primary outcome measure. As a power of 0.8, a significance level of $\alpha=0.05$ and a dropout-rate of 30% were assumed, $N=92$ participants were the calculated sample size needed in the current study.

Test for Bias

The primary and secondary outcomes’ descriptive statistics and distributions were calculated to test whether they met the requirements for the statistical analyses. In addition, all baseline variables were tested for significant biases concerning the factors “group”, “study center” and “dropout”.

To account for possible selection effects, outcome analyses were conducted in two data sets: (a) full analysis set (FAS) or intent-to-treat-sample, and (b) per protocol set (PPS) or completer sample, which consisted of all participants who took part in the baseline- and post-assessment. For exploratory analyses, a (c) quality set (QS) was analyzed, which consisted of participants from the PPS set which had an adherence score of at least 75% in the use of the app “Mindable”.

Missing Values

As the questionnaires were presented electronically, there were no missing values, as participants were forced to make a choice if they wanted to continue with the questionnaire. In the case of the analyses in the FAS set, it was made sure that missing values due to dropout were missing completely at random. After that, missing values were estimated by means of multiple imputation.

Data Analysis

App efficacy was analyzed by means of linear mixed models (LMM). The dependent variables were scores of primary and secondary outcome measures. The variable “participant” was utilized as a random effect in the model in order to control for individual differences in participants. Furthermore, the variables “group” (app vs. waitlist control), “time” (baseline vs. post), the group x time-interaction, as well as meeting the criteria of an agoraphobia and either being recruited “online” vs. “face-to-face” were utilized as fixed effects. All analyses were two-tailed ($\alpha=0.05$). To account for multiple testing, secondary outcome analyses were additionally analyzed considering a Bonferroni-corrected α of 0.017.

Analyses for all four hypotheses were performed using a “full-model” including the interaction that was compared to a “null-model” that differed only by excluding the interaction. This was done in order to be able to evaluate whether including the interaction significantly improved the model x data-fit. In reference to Nakagawa and Schielzeth (2013), the difference in variance explained by each of the two models (R^2) was calculated in order to specify the variance explained by the interaction between time and group.

Response was evaluated by calculating the Reliable Change Index (RCI, Christensen & Mendoza, 1986; Jacobson et al., 1984). Remission was defined by achieving reliable change and a PAS score below the cut-off of 8 points in the post-assessment.

Data were prepared and analyzed using IBM SPSS 29.0 statistical software and the R packages *lme4* for calculating the LMM (Bates et al., 2015), *r2glmm* for calculating the R^2 (Nakagawa & Schielzeth, 2013) and *mitml* for the multiple imputation (Grund et al., 2023). In order to impute missing data, 20 estimations per missing value were made, using all variables relevant to the LMM as predictors. These 20 estimations were then pooled in accordance with Rubin’s rules (Rubin, 1987). The multiple imputation’s convergence was examined and found to be satisfactory. All Rhat-scores were <1.05 and all Trace-plots varied about a constant equilibrium.

Results

Participants and Enrollment

After screening, a total of 122 participants took part in the baseline-assessment. 107 individuals (78.7%) met the inclusion criteria, and were included in the study (see Fig. 1 for a study flow chart). Seventeen participants (15.9%) dropped out during the observation period or could not be contacted anymore. Therefore, 90 participants completed the study. Dropout-rates did not significantly differ between the two study groups ($\chi^2(1)=1.06, p=.30$).

Baseline Characteristics

Table 1 shows the participants' baseline characteristics. The majority of the participants were female, younger than 40 years of age and highly educated. Most participants met the criteria of both panic disorder and agoraphobia ($n=87$), and a third of the participants had at least one comorbid disorder (mostly depressive disorders).

Bias Control

It was analyzed whether participants having dropped out of the study differed from participants completing the study in baseline variables. No significant differences were found in the demographic variables nor the primary and secondary outcomes. Hence, the risk of selection biases was estimated to be low.

Additionally, it was analyzed whether baseline variables differed across study centers. Because of the change in recruitment from face-to-face to online, a new variable "recruitment" was created with the labels "Bremen" ($n=20$), "Hamburg" ($n=17$), "Münster" ($n=3$) and "online" ($n=67$). Significant differences between these recruitment "centers" were found concerning age ($F(3)=2.92; p=.038$), the percentage of people suffering from an agoraphobia and a panic disorder ($\chi^2(1)=9.25, p=.026$) and baseline values on the PAS ($F(3)=5.27; p=.02$). Post-hoc analyses showed that the "online" group had significantly higher levels of self-rated anxiety in comparison to the other recruitment groups. Because of the low number of participants in Münster, this group could not be considered in the post-hoc analyses. For further analyses, the dichotomous variable "online" vs. "face-to-face" was created and used as a fixed effect in the main outcome analyses.

Adverse Events

Two participants reported symptom deterioration during the observation period, both due to other reasons than study

participation. One of them dropped out of the study. No other adverse events were reported.

App Usage

Two participants in the app group did not consent to the processing of their app data. Hence, app data was available for 55 participants. App usage widely differed with the highest mean adherence score for the psychoeducation module ($M=85.86, SD=33.16$), where more than 85% of participants completed all of the psychoeducation contents. On average, participants completed 7.73 ($SD=2.98$) out of nine psychoeducation lessons. The modules "symptom provocation" and "exposure" were used less frequently. Only half of the participants (56.4%) underwent five or more exercises. 25% of the participants did not document any exercise. The mean sum of interoceptive and in vivo exposure exercises was $M=9.15 (SD=15.73)$. On average, participants completed 63% ($SD=42.23$) of the recommended exercises in the exercise modules. Regarding the self-monitoring modules "daily diary" and "weekly checkup", 70% of participants made at least eight entries, with an average of $M=15.20 (SD=14.93)$ entries. On average, participants completed 78% ($SD=30.43$) of the recommended checkup entries. The mean overall adherence score was 75.6% ($SD=27.72$) and 34 out of 55 (61.8%) participants met the overall adherence score cut-off of 75.

User Satisfaction

Satisfaction with the app was moderate to high. Mean agreement was highest for the items "I would recommend this training to a friend, if he or she were in need of similar help" ($M=3.39, SD=0.71$) and "The training I attended was of high quality" ($M=3.30, SD=0.70$). The item "The training has met my needs" received the lowest agreement ($M=2.67, SD=0.94$) in the study sample. The user satisfaction mean scores and standard deviations were also calculated for the participants, who had an adherence score of $>75\%$ (QS, $n=29$). High app users did not differ in satisfaction compared to the whole app group.

Severity of Panic and Agoraphobic Symptoms

Table 2 shows the means and standard deviations at the different assessment time points in the FAS- and PPS-data sets of all outcomes. The mean PAS scores were higher than 8 and lower than 28 points indicating moderate PD/A symptom severity. In order to analyze, whether using the app would result in higher reductions in PD/A symptoms, the change in the total PAS scores over time was compared between the two groups by means of LMM. The results for

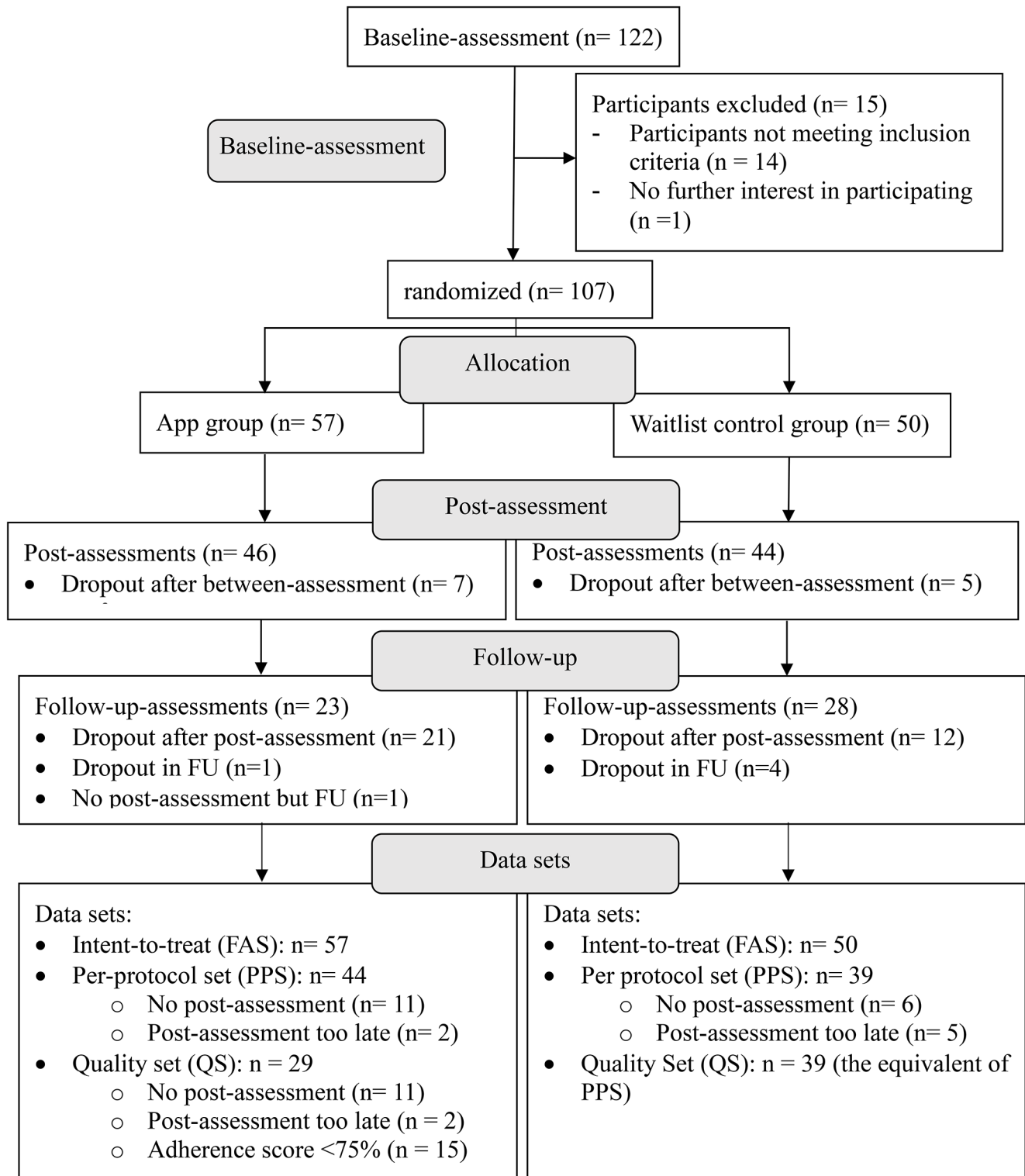


Fig. 1 Flow of participants

Table 1 Demographic sample characteristics at baseline

	Total (<i>N</i> =107) ¹	AG (<i>n</i> =57)	CG (<i>n</i> =50)	Test for differences ²
Gender*: % (n)				
female	76.6 (82)	70.2 (40)	84.0 (42)	$\chi^2(1) = 2.84, p = .092$
male	23.4 (25)	29.8 (17)	16.0 (8)	
Age: M (SD)	36.10 (14.23)	36.35 (14.40)	35.82 (14.17)	$t(105) = -0.19, p = .848$
Education: % (n)				
None (yet)	5.7 (6)	5.4 (3)	6.0 (3)	$\chi^2(3) = 0.82, p =$
student	21.7 (23)	21.4 (12)	22.0 (11)	0.844
Vocational training	48.1 (51)	51.8 (29)	44.0 (22)	
University degree	24.5 (26)	21.4 (12)	28.0 (14)	
Employment status: % (n)				
employed	65.4 (70)	63.2 (36)	68.0 (34)	$\chi^2(1) = 0.28, p =$
Not employed	34.6 (37)	36.8 (21)	32.0 (16)	0.374
Diagnosis: % (n)				
Panic disorder	94.4 (101)	94.7 (54)	94.0 (47)	$\chi^2(1) = 0.03, p = 1.00$
Agoraphobia	86.9 (93)	93.0 (53)	80.0 (40)	$\chi^2(1) = 3.95, p = .082$
Comorbidity: % (n)				
At least one comorbid disorder	64.5 (69)	61.4 (35)	68.0 (34)	$\chi^2(1) = 0.51, p = .546$
Comorbid depressive disorder	52.3 (56)	45.6 (26)	60.0 (30)	$\chi^2(1) = 2.21, p = .175$
Medication % (n)	24.3 (26)	22.8 (13)	26.0 (13)	$t(1) = 0.15, p = .822$

Note: AG = App group, CG = Waitlist control group, * = All participants either identified as male or female, no one identified as diverse

¹ Variations in the total *N* caused by missing values

² Test for differences dependent on the variable being categorical (chi square test) or metric (t-test)

Table 2 Descriptive data of the outcome variables at baseline- and post-assessment in the two groups and the FAS- and PPS-data sets

Outcome	Data set Time Point	FAS			PPS		
		<i>N</i>	AG	CG	<i>N</i>	AG	CG
PAS	Baseline	107	27.51 (8.49)	25.82 (8.49)	83	27.55 (9.25)	26.28 (8.64)
	Post	90	19.96 (9.34)	23.82 (9.27)	83	19.98 (9.45)	23.67 (9.00)
QoL	Baseline	107	55.12 (17.81)	54.33 (16.79)	83	56.91 (17.99)	53.21 (17.06)
	Post	90	60.78 (18.83)	54.17 (18.70)	83	61.00 (19.24)	53.00 (17.78)
ACQ	Baseline	107	74.39 (17.32)	71.28 (14.67)	83	76.66 (16.51)	69.77 (14.73)
	Post	90	85.41 (18.99)	72.64 (16.52)**	83	86.25 (18.67)	72.10 (16.77)**
SDS	Baseline	107	4.21 (2.21)	4.13 (2.41)	83	4.27 (2.24)	4.21 (2.52)
	Post	90	3.43 (2.30)	3.65 (2.64)	83	3.39 (2.33)	3.65 (2.57)

Note. AG = App group; CG = Waitlist control group; FAS = Full-analysis set; PPS = Per protocol-set; α -level = 0.05, ** = Bonferroni corrected α -level = 0.017; PAS = Panic and Agoraphobia Scale; QoL = WHOQOL-BREF: World Health Organization Quality of Life Scale, Domain psychological health; ACQ = Anxiety Control Questionnaire; SDS = Sheehan Disability Scale

the intent-to-treat- (FAS) and the completer- (PPS) data sets are depicted in Table 3. Exploratory analyses for all outcome measures in the quality (QS)-data set can be found in the Appendix A and B. Descriptive data on PAS subscales can be found in Appendix C.

In all three data sets the time as well as the group x time interaction effect became significant, whereas the overall group effect did not. Hence, the app group and the waitlist control group changed differently over time. In all three data sets it made a significant difference whether the interaction was included in the model or not, but the variance in the data explained by the “full-model” was low in all three cases (FAS: $F(1, 760.431) = 11.538, p < .001, R^2 = 0.20$;

PPS: $\chi^2(1) = 9.205, p = .002, R^2 = 0.19$; QS: $\chi^2(1) = 9.067, p = .003, R^2 = 0.20$). Adding the interaction to the model had a small effect in all three data sets ($\text{Diff}(R^{2full} - R^{2null})$: FAS = 0.019; PPS = 0.016; QS = 0.022).

Furthermore, the factor “site” gained significance in all three data sets indicating that people differed dependent on the way they were assessed (online vs. face-to-face). The factor “agoraphobia” was significant in the FAS-data set but not in the PPS- and QS- data sets. This suggests that suffering from agoraphobia influenced the response to the study.

Table 3 Results of the LMM in all outcome measures and the FAS- and PPS-data sets

Outcome	Data set								
	Factor	FAS				PPS			
		Estimate	Std. Error	t-Value	<i>p</i>	Estimate	Std. Error	t-Value	<i>p</i>
PAS	Intercept	22.02	2.14	10.30	<0.001	25.74	2.84	9.08	<0.001
	Group	-0.77	0.75	-1.03	0.304	-0.88	1.91	-0.46	0.645
	Time	2.43	0.38	6.48	<0.001	-7.57	1.10	-6.87	<0.001
	Agora	4.84	2.23	2.17	0.030	4.20	2.66	1.58	0.119
	Site	-5.28	1.53	-3.46	0.001	-5.53	1.78	-3.10	0.003
	Group*Time	1.27	0.38	3.40	0.001	4.95	1.61	3.08	0.003
QoL**	Intercept	57.36	4.88	11.75	<0.001	54.06	6.19	8.74	<0.001
	Group	1.60	1.70	0.94	0.349	-3.75	3.99	-0.94	0.351
	Time	-1.28	0.57	-2.24	0.025	3.98	1.57	2.53	0.013
	Agora	-3.25	5.08	-0.64	0.522	1.10	5.87	0.19	0.852
	Site	3.91	3.46	1.13	0.258	5.09	3.94	1.29	0.199
	Group*Time	-0.97	0.54	-1.79	0.073	-4.19	2.29	-1.83	0.072
ACQ**	Intercept	80.17	4.45	18.00	<0.001	75.06	5.61	13.39	<0.001
	Group	3.95	1.55	2.55	0.011	-7.04	3.72	-1.89	0.061
	Time	-3.06	0.66	-4.64	<0.001	9.59	1.90	5.06	<0.001
	Agora	-7.58	4.63	-1.64	0.101	-0.49	5.29	-0.09	0.927
	Site	6.09	3.16	1.93	0.054	5.62	3.54	1.59	0.117
	Group*Time	-1.87	0.63	0.63	0.003	-7.26	2.29	-2.62	0.010
SDS**	Intercept	3.74	0.62	6.05	0.000	4.36	0.80	5.46	<0.001
	Group	-0.10	0.22	-0.44	0.661	-0.02	0.53	-0.04	0.970
	Time	0.35	0.09	3.90	<0.001	-0.88	0.26	-3.39	0.001
	Agora	0.49	0.65	0.75	0.451	0.29	0.75	0.39	0.699
	Site	-0.91	0.44	-2.06	0.040	-0.99	0.51	-1.97	0.053
	Group*Time	0.10	0.09	1.15	0.250	0.31	0.38	0.83	0.407

Note. FAS=Full-analysis set; PPS=Per protocol-set; α -level=0.05, ** = Bonferroni corrected α -level=0.017; PAS=Panic and Agoraphobia Scale; QoL=WHOQOL-BREF: World Health Organization Quality of Life Scale, Domain psychological health; ACQ=Anxiety Control Questionnaire; SDS=Sheehan Disability Scale; Agora= Fulfilling the criteria of an agoraphobia yes/no; Site: Recruited online vs. face-to-face

Response and Remission

The Reliable Change Index as an indicator of response in participants was calculated by means of the PAS' norm sample ($r_{tt} = 0.78$). This resulted in a reliable change being defined by a difference of > 13.4 points from baseline- to post-assessment. Based on this definition, no participant displayed a reliable deterioration. 13% of the app group ($n=6$) and 9.1% ($n=4$) of the control-group reported reliable improvements ($\chi^2(1)=0.356$ $p=.551$). 4.3% of the app group ($n=2$) and none of the waitlist control group participants met the remission criteria ($\chi^2(1)=1.957$ $p=.162$) of having an RCI > 13.4 and falling below the PAS cut-off score of 8 points. At post-assessment, 86.67% (AG=84.78%, CG=88.64%) fulfilled the DSM-5 diagnostic criteria of a panic disorder and 82.22% (AG=80.43%, CG=84.09%) those of an agoraphobia. Response and remission rates did not significantly differ between groups.

Secondary Outcomes

Quality of Life

Quality of life was evaluated to be moderate. It varied around 55 and 54 points with the WHOQOL-BREF (domain psychological health) ranging from 0 to 100 points (see Table 2). The group x time interaction did not gain significance in any of the data sets indicating no significant differences over time in quality of life in the two groups. Table 3 depicts the results of the LMM in the intent-to-treat- and the completer- data sets (see Appendix A and B for results in the quality-data set). Therefore, no further analyses concerning the importance of the interaction effect for perceived quality of life were calculated. The effect of time was significant in all data sets. In the FAS- and QS-data sets the time effect did not gain significance when being analyzed by means of the Bonferroni corrected α -level of 0.017.

Anxiety-Related Control Beliefs

Both groups had slightly lower mean scores on the ACQ at baseline than the cut-off of 75.5 indicating relevant impairments in anxiety-related control beliefs (see Table 2). Next, it was examined whether the app group and the waitlist control group differed significantly over time in their change in anxiety-related control beliefs. The results of the LMM in the intent-to-treat- and the completer-data sets are shown in Table 3 (see Appendix A and B for the results in the quality-data set). The group effect was significant in the FAS analysis but not in the PPS or QS analysis whereas the time effect was significant in all three data sets. The group x time interaction was significant in all three data sets, too. In all three data sets it made a significant difference whether the interaction effect was included in the model or if not, but the variance explained by the “full-model” was low in all three data sets (FAS: $F(1, 2508.540) = 8.701, p = .003, R^2 = 0.125$); PPS: $\chi^2(1) = 6.770, p = .009, R^2 = 0.145$); QS: $\chi^2(1) = 5.908, p = .015, R^2 = 0.15$). Adding the interaction effect to the model had a small effect in all three data sets ($\text{Diff}(R^{2\text{full}} - R^{2\text{null}}$): FAS = 0.011; PPS = 0.009; QS = 0.009).

Disability

The SDS scores varied around 4 points in both groups which indicated that participants were “moderately” impaired in functioning. To examine whether the app group and the waitlist control group differed significantly in the change in functional impairment over time LMM were calculated in the three data sets FAS, PPS and QS. The results of the FAS- and PPS-data sets are shown in Table 3 (see Appendix A and B for results of the QS-data set). The time factor gained significance in all three data sets, indicating that all participants’ functional impairment improved significantly over time. The factor “site” gained significance in the FAS-data set. The latter effect was not significant anymore when considering the Bonferroni-adjusted α -level. As the time x group interaction was not significant no further analysis were conducted concerning the importance of the interaction for the model.

Discussion

Main Findings

Many people suffering from panic disorder and/or agoraphobia remain untreated due to limited access and/or availability of therapeutic treatment. Given the high chronicity of these conditions, improving mental health care for those people seems mandatory. The current study, thus, aimed to

evaluate a self-management approach using a smartphone-based app that can easily be implemented.

Our results showed that the app “Mindable” led to a significant reduction of PD/A symptoms over time compared to a waitlist control group. This effect was found across the different data analysis sets, indicating stability of the findings. Reductions in panic and agoraphobic symptoms were smaller than in face-to-face CBT (see Gloster et al., 2011), however, as the extent of PD/A symptoms at baseline was comparable between the therapy study (PAS range of 27.1–28.4) and our study (mean baseline PAS of 26.7), and given that these reductions were achieved without any therapist assistance, we believe this to be a valuable contribution to support persons with serious anxiety problems. The significant reduction of PD/A symptoms found in the current study is also in line with recent meta-analyses and systematic reviews on DHA for PD/A (Andrews et al., 2018; Domhardt et al., 2020; Stech et al., 2020; Weisel et al., 2019). Concerning smartphone-based DHA, our findings are somewhat in contrast to findings from Ivanova and colleagues (2016) who could not show significant improvements for panic symptoms. This might be attributed to the different treatment approaches used: Ivanova et al. (2016) relied on an ACT protocol, while the app “Mindable” used an exposure rationale, as the evidence-based treatment for PD/A. In line with this assumption, Ebenfeld and colleagues (2021) who also used exposure exercises within their app, also showed significant improvements. It has to be noted that the reductions found in our study were achieved by pure self-help without further therapeutic contact, thus providing further evidence that non-guided DHA might be as effective as DHA incorporating further therapeutic contact (Pauley et al., 2023).

The app group showed a significant increase of anxiety-related control beliefs in comparison to the waitlist control group over time. This is a promising finding as anxiety control is not only associated with avoidance behavior (Telch et al., 1989; White et al., 2006), it was also found to mediate the relationship between panic attacks and subsequent increases in anticipatory anxiety (Helbig-Lang et al., 2012a). Studies in therapeutic settings have already shown that an increase in anxiety-related control beliefs was accompanied by symptom reductions in patients with panic disorder and social anxiety disorder (Craske et al., 2014). It might be assumed the increases in perceived anxiety control have also mediated anxiety reductions in this study, however, this assumption needs further evaluation.

Participants in both groups showed a significant increase in quality of life and functional impairment over time. This is in line with the findings that passive psychoeducation strategies, such as receiving information or materials on a mental disorder or receiving feedback on one’s diagnosis

after a structured clinical interview have been shown to lead to symptom reduction in depression and anxiety (Donker et al., 2009). Anyhow, the group x time effects on quality of life and functional impairment were not significant in the current study. The time period may have been too short to be able to find significant changes in the perceived quality of life by means of the app. Ebenfeld and colleagues (2021) for instance, found no significant changes in quality of life in people suffering from PD/A at the post-assessment or a 3-month follow-up after using an app combined with online coaching, but at 6-month follow-up participants showed significant increases in quality of life. An explanation for the lack of change in functional impairment due to the app may be that the questionnaires used to assess functional impairment were too unspecific and therefore, external confounding influences may have had a greater impact on the answers given in these questionnaires compared to symptom specific questionnaires.

All in all, the results on the effect of smartphone-based DHA on the change in perceived quality of life and functional impairment in people suffering from PD/A are still inconsistent and further research is needed.

It can be highlighted that the study was conducted under naturalistic conditions (few exclusion criteria, participants were free in the way they used the app). As the app did not require further contact with a therapist, this treatment approach meets the current challenges faced by practitioners and patients due to the limited access to psychotherapy. The switch to online baseline-assessments allowed people from all over Germany to take part in the study and further strengthened external validity. Furthermore, the sample size was satisfactory in comparison to many other studies in the field. The drop-out rate of 15.9% was lower than the mean drop-out rate calculated in a recent meta-analysis on DHA (24%, Linardon & Fuller-Tyszkiewicz, 2020), maybe reflecting the rather high user satisfaction reported by participants.

Limitations

Findings have to be considered with caution due to some study limitations. First, the change from face-to-face to online assessments may have led to a distortion of the study results. Because of the COVID-19 pandemic, face-to-face assessments were not feasible for most of the participants. As a result, more severe cases of PD/A may have been able to participate because people did not have to leave their home to take part. Differences in the baseline-assessment between participants of the “face-to-face” vs. the “online” group in anxiety symptoms underline this assumption. On the one hand, variance in the baseline-assessments and, therefore, the general representativeness of the sample

may have been increased; on the other hand, maybe more severely anxious participants were included in the study for whom the app “Mindable” may not be indicated. Anyhow, exploratory analyses showed no significant differences between “face-to-face” and “online” participants in the way PD/A symptoms changed over time.

Second, suffering from an agoraphobia had a significant impact on how participants responded to the study. It is possible that this effect was overestimated because only very few participants did not meet the diagnostic criteria of an agoraphobia ($n = 14$) and, thus, variation in this fixed factor was low. Anyhow, other studies have found that agoraphobic avoidance at baseline leads to less improvement (e.g. Porter & Chambless, 2015). Future studies should further examine whether the self-management apps are equally suitable for individuals with panic disorder, agoraphobia, and both disorders. In this regard, future studies may also profit from a clinical evaluation of panic and agoraphobia symptoms next to the self-report by participants.

Third, it needs to be noted that even though the study was conducted under naturalistic conditions, the findings may be limited in their real-life applicability. Participants in our study had contact to professionals at baseline- and post-assessment. Research on DHA where participants could take part in studies without any contact with a researcher led to higher attrition rates compared to participants who at least had a telephone or an in-person interview with a researcher (Linardon et al., 2020). As the use of Mindable has a medical prescription as a prerequisite, this at least partially reflects naturalistic health care. Additionally, the app “Mindable” offers reminder messages because this has been shown to lead to lower attrition rates (Linardon et al., 2020).

A last, but serious limitation of the study is the lack of conclusive follow-up data. As participants in the waitlist group gained access to the app during the follow-up period, group comparisons at follow-up were not meaningful to analyze long-term effects. An exploratory analysis of follow-up data suggested that improvements in PD/A symptoms were not stable, as the PAS total score deteriorated from post to follow-up (Appendix E). As these analyses were based on a limited sample of $n = 23$ participants of the app group only, we cannot conclude whether this reflects an actual deterioration effect or an unspecific effect of time. These effects were also limited to the PAS; improvements in anxiety-related control beliefs remained rather stable, suggesting that participants had achieved confidence in coping with anxiety-related symptoms. However, there is a need for further studies specifically examining long-term effects of DHA, especially when using these strategies as part of a stepped care approach.

All in all, future studies should replicate the current study’s results and expand on them. For instance, follow-up

analyses will be important to further be able to make assumptions on the long-term efficacy of using a self-management app for treating PD/A symptoms. Further, interindividual variables should be examined to be able to make more solid statements on the mechanisms underlying the change in PD/A symptoms over time. It is of great importance to further examine how exactly which app module impacts which symptom and which therapeutic mechanism lies beyond the symptom improvements. Comparing the app effects to an active control group and examining whether the app might further enhance effectiveness of face-to-face therapy will also foster health care for anxiety disorders in the future.

Conclusions

In conclusion, the self-management smartphone app “Mindable” has shown to significantly reduce anxiety symptoms and to improve anxiety-related control beliefs in patients suffering from panic disorder and/ or agoraphobia. It is well accepted and easy to use. As no symptom remission could be achieved and the exploratory follow-up analysis indicated low stability in PD/A symptom change, this approach may be helpful for a stepped care approach for anxiety disorders taking limited treatment capacities into account.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10608-024-10513-w>.

Acknowledgements We thank all study participants and the Christoph-Dornier-Foundation of Clinical Psychology for supporting the study. We also thank Susanne Christ for coordinating the course of the study, Dirk Schulte am Hülse (Constructor University Bremen) for the technical support, Prof. Dr. Klaus Boehnke (Constructor University Bremen) and Prof. Dr. Simon Grund (University of Hamburg) for the statistical support and all people involved in the diagnostic processes: Hannah Berger-Johannsen, Anne Löttert, Amira Blal, Nike Kirst, Annika Stauffer, Tania Vorotniak and Simone Voss.

Author Contributions Justine Spies: Data curation, formal analysis, investigation, visualization, writing – original draft, review and editing; Sylvia Helbig-Lang: Conceptualization, formal analysis, funding acquisition, methodology, project administration, resources, supervision, validation, visualization, writing – original draft, review and editing; Thomas Lang: Conceptualization, formal analysis, funding acquisition, project administration, resources, supervision.

Funding This research was funded by the Christoph Dornier Stiftung and by the Mindable Health GmbH.

Data Availability Data will be made available on request.

Declarations

Ethical Approval Ethical approval for the study was granted by the ethical committee of the German Society for Psychology (reference

number: LangThomas2020-12-14VA).

Informed Consent Informed consent was obtained from all participants included in the study.

Conflict of Interest The study was funded by Mindable GmbH and the Christoph Dornier Stiftung and conducted within an accreditation of the Federal Institute for Drugs and Medical Devices (BfArM).

References

- Andersson, G., Cuijpers, P., Carlbring, P., Riper, H., & Hedman, E. (2014). Guided internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: A systematic review and meta-analysis. *World Psychiatry, 13*(3), 288–295. <https://doi.org/10.1002/wps.20151>
- Andrews, G., Cuijpers, P., Craske, M. G., McEvoy, P., & Titov, N. (2010). Computer therapy for the anxiety and depressive disorders is effective, acceptable and practical health care: A meta-analysis. *Plos One, 5*(10). <https://doi.org/10.1371/journal.pone.0013196>
- Andrews, G., Basu, A., Cuijpers, P., Craske, M. G., McEvoy, P., English, C. L., & Newby, J. M. (2018). Computer therapy for the anxiety and depression disorders is effective, acceptable and practical health care: An updated meta-analysis. In *Journal of Anxiety Disorders* (Vol. 55, pp. 70–78). Elsevier Ltd. <https://doi.org/10.1016/j.janxdis.2018.01.001>
- Angermeyer, C., Kilian, R., & Matschinger, H. (2000). *Deutschsprachige Version Der WHO-Instrumente Zur Erfassung Von Lebensqualität WHOQOL-100 und WHOQOL-BREFM*. Hogrefe.
- Bandelow, B. (2016). *Panik- Und Agoraphobie-Skala* (2nd ed.). Hogrefe.
- Bandelow, B., Brunner, E., Broocks, A., Beinroth, D., Hajak, G., Pralle, L., & Rüther, E. (1998). The use of the panic and agoraphobia scale in a clinical trial. *Psychiatry Research, 77*(1), 43–49. [https://doi.org/10.1016/S0165-1781\(97\)00118-2](https://doi.org/10.1016/S0165-1781(97)00118-2)
- Bandelow, B., Aden, I., Alpers, G. W., Benecke, A., Benecke, C., Deckert, J., Domschke, K., Eckhardt-Henn, A., Geiser, F., Gerlach, A. L., Harfst, T., Hau, S., Hoffmann, S., Hoyer, J., Hunger-Schoppe, C., Kellner, M., Köllner, V., Kopp, I. B., Langs, G., & Beutel, M. E. (2021). *S3-Leitlinie Behandlung von Angststörungen: Version 2*
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1). <https://doi.org/10.18637/jss.v067.i01>
- Boß, L., Lehr, D., Reis, D., Vis, C., Riper, H., Berking, M., & Ebert, D. D. (2016). Reliability and validity of assessing user satisfaction with web-based health interventions. *Journal of Medical Internet Research, 18*(8), e234. <https://doi.org/10.2196/jmir.5952>
- Brown, T. A., White, K. S., Forsyth, J. P., & Barlow, D. H. (2004). The structure of perceived emotional control: Psychometric properties of a revised anxiety control questionnaire. *Behavior Therapy, 35*(1), 75–99.
- Christensen, L., & Mendoza, J. L. (1986). A method of assessing change in a single subject: An alteration of the RC index. *Behavior Therapy, 17*(3), 305–308. [https://doi.org/10.1016/S0005-7894\(86\)80060-0](https://doi.org/10.1016/S0005-7894(86)80060-0)
- Craske, M. G., Treanor, M., Conway, C. C., Zbozinek, T., & Vervliet, B. (2014). Maximizing exposure therapy: An inhibitory learning approach. *Behaviour Research and Therapy, 58*, 10–23. <https://doi.org/10.1016/j.brat.2014.04.006>
- de Vries, L. P., Baselmans, B. M. L., & Bartels, M. (2021). Smartphone-based ecological momentary assessment of well-being:

- A systematic review and recommendations for future studies. *Journal of Happiness Studies*, 22(5), 2361–2408. <https://doi.org/10.1007/s10902-020-00324-7>
- Domhardt, M., Geflehn, H., von Rezori, R. E., & Baumeister, H. (2019). Internet- and mobile-based interventions for anxiety disorders: A meta-analytic review of intervention components. *Depression and Anxiety*, 36(3), 213–224. <https://doi.org/10.1002/da.22860>
- Domhardt, M., Letsch, J., Kybelka, J., Koenigbauer, J., Doebler, P., & Baumeister, H. (2020). Are internet- and mobile-based interventions effective in adults with diagnosed panic disorder and/or agoraphobia? A systematic review and meta-analysis. *Journal of Affective Disorders*, 276, 169–182. <https://doi.org/10.1016/j.jad.2020.06.059>
- Donker, T., Griffiths, K. M., Cuijpers, P., & Christensen, H. (2009). Psychoeducation for depression, anxiety and psychological distress: A meta-analysis. *BMC Medicine*, 7(1), 79. <https://doi.org/10.1186/1741-7015-7-79>
- Ebenfeld, L., Lehr, D., Ebert, D. D., Stegemann, S. K., Riper, H., Funk, B., & Berking, M. (2021). Evaluating a hybrid web-based training program for panic disorder and agoraphobia: Randomized controlled trial. *Journal of Medical Internet Research*, 23(3). <https://doi.org/10.2196/20829>
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. <https://doi.org/10.3758/BF03193146>
- Firth, J., Torous, J., Nicholas, J., Carney, R., Rosenbaum, S., & Sarris, J. (2017). Can smartphone mental health interventions reduce symptoms of anxiety? A meta-analysis of randomized controlled trials. *Journal of Affective Disorders*, 218, 15–22. <https://doi.org/10.1016/j.jad.2017.04.046>
- Gloster, A. T., Wittchen, H. U., Einsle, F., Lang, T., Helbig-Lang, S., Fydrich, T., Fehm, L., Hamm, A. O., Richter, J., Alpers, G. W., Gerlach, A. L., Ströhle, A., Kircher, T., Deckert, J., Zwanzger, P., Höfler, M., & Arolt, V. (2011). Psychological treatment for panic disorder with agoraphobia: A randomized controlled trial to examine the role of therapist-guided exposure in situ in CBT. *Journal of Consulting and Clinical Psychology*, 79(3), 406–420. <https://doi.org/10.1037/a0023584>
- Grund, S., Robitzsch, A., & Luedtke, O. (2023). *mitml: Tools for multiple imputation in multilevel modeling. R package version 0.4-5*.
- Hähnel, A., Weigel, B., Hoyer, J., & Helbig, S. (2004). Wartezeit für Psychotherapiepatienten – Und Wie Sie zu Nutzen Ist. *Verhaltenstherapie*, 14(4), 294–302. <https://doi.org/10.1159/000082839>
- Helbig, S., & Hoyer, J. (2007). Hilft Wenig Viel? Eine Minimalintervention für Patienten während Der Wartezeit auf ambulante Verhaltenstherapie. *Verhaltenstherapie*, 17(2), 109–115.
- Helbig-Lang, S., Lang, T., Petermann, F., & Hoyer, J. (2012a). Anticipatory anxiety as a function of panic attacks and panic-related self-efficacy: An ambulatory assessment study in panic disorder. *Behavioural and Cognitive Psychotherapy*, 40(5), 590–604. <https://doi.org/10.1017/S1352465812000057>
- Helbig-Lang, S., Schneider, A., & Petermann, F. (2012b). Angstbezogenes Kontrollierenleben. *Zeitschrift Für Klinische Psychologie Und Psychotherapie*, 41(4), 258–270. <https://doi.org/10.1026/1616-3443/a000171>
- Heron, K. E., & Smyth, J. M. (2010). Ecological momentary interventions: Incorporating mobile technology into psychosocial and health behaviour treatments. *British Journal of Health Psychology*, 15(1), 1–39. <https://doi.org/10.1348/135910709X466063>
- Hodgins, D. C. (2013). Reliability and validity of the Sheehan disability scale modified for pathological gambling. *BMC Psychiatry*, 13. <https://doi.org/10.1186/1471-244X-13-177>
- Ivanova, E., Lindner, P., Ly, K. H., Dahlin, M., Vernmark, K., Andersson, G., & Carlbring, P. (2016). Guided and unguided acceptance and commitment therapy for social anxiety disorder and/or panic disorder provided via the internet and a smartphone application: A randomized controlled trial. *Journal of Anxiety Disorders*, 44, 27–35. <https://doi.org/10.1016/j.janxdis.2016.09.012>
- Jacobi, F., Höfler, M., Siegert, J., Mack, S., Gerschler, A., Scholl, L., Busch, M. A., Hapke, U., Maske, U., Seiffert, L., Gabel, W., Maier, W., Wagner, M., Zielasek, J., & Wittchen, H. U. (2014). Twelve-month prevalence, comorbidity and correlates of mental disorders in Germany: The mental health module of the German health interview and examination survey for adults (DEGS1-MH). *International Journal of Methods in Psychiatric Research*, 23(3), 304–319. <https://doi.org/10.1002/mpr.1439>
- Jacobson, N. S., Follette, W. C., & Revenstorf, D. (1984). Psychotherapy outcome research: Methods for reporting variability and evaluating clinical significance. *Behavior Therapy*, 15(4), 336–352. [https://doi.org/10.1016/S0005-7894\(84\)80002-7](https://doi.org/10.1016/S0005-7894(84)80002-7)
- Lang, T., Helbig-Lang, S., & Petermann, F. (2009). Was wirkt in Der Kognitiven Verhaltenstherapie Der Panikstörung Mit Agoraphobie? *Zeitschrift für Psychiatrie Psychologie und Psychotherapie*, 57(3), 161–175. <https://doi.org/10.1024/1661-4747.57.3.161>
- Lang, T., Helbig-Lang, S., Westphal, D., Gloster, A., & Wittchen, H. U. (2012). *Expositionsbasierte Therapie Der Panikstörung Mit Agoraphobie. Ein Behandlungsmanual*. Hogrefe.
- Leon, A. C., Shear, M. K., Portera, L., & Klerman, G. L. (1992). Social psychiatry and psychiatric epidemiology assessing impairment in patients with panic disorder: The Sheehan disability scale. *Social Psychiatry and Psychiatric Epidemiology*, 27, 78–82.
- Linardon, J., & Fuller-Tyszkiewicz, M. (2020). Attrition and adherence in smartphone-delivered interventions for mental health problems: A systematic and meta-analytic review. *Journal of Consulting and Clinical Psychology*, 88(1), 1–13. <https://doi.org/10.1037/ccp0000459>
- Linardon, J., Cuijpers, P., Carlbring, P., Messer, M., & Fuller-Tyszkiewicz, M. (2019). The efficacy of app-supported smartphone interventions for mental health problems: A meta-analysis of randomized controlled trials. *World Psychiatry*, 18(3), 325–336. <https://doi.org/10.1002/wps.20673>
- Maaß, L., Freye, M., Pan, C. C., Dassow, H. H., Niess, J., & Jahnel, T. (2022). The definitions of health apps and medical apps from the perspective of public health and law: Qualitative analysis of an interdisciplinary literature overview. *JMIR MHealth and UHealth*, 10(10), e37980. <https://doi.org/10.2196/37980>
- Margraf, J., Cwik, J. C., Suppiger, A., & Schneider, S. (2017). DIPS open access: Diagnostic interview for mental disorders. *Bochum: Mental Health Research and Treatment Center, Ruhr-Universität Bochum*.
- Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, 4(2), 133–142. <https://doi.org/10.1111/j.2041-210x.2012.00261.x>
- National Institute for Health and Care Excellence (NICE) Generalised anxiety disorder and panic disorder in adults: Management. (2020). Guidance and guidelines. <https://www.nice.org.uk/guidance/cg113/resources>
- Neutens, T. (2015). Accessibility, equity and health care: Review and research directions for transport geographers. *Journal of Transport Geography*, 43, 14–27. <https://doi.org/10.1016/j.jtrangeo.2014.12.006>
- Pauley, D., Cuijpers, P., Papola, D., Miguel, C., & Karyotaki, E. (2023). Two decades of digital interventions for anxiety disorders: A systematic review and meta-analysis of treatment effectiveness. *Psychological Medicine*, 1–13. <https://doi.org/10.1017/S0033291721001999>
- Porter, E., & Chambless, D. L. (2015). A systematic review of predictors and moderators of improvement in cognitive-behavioral

- therapy for panic disorder and agoraphobia. *Clinical Psychology Review*, 42, 179–192. <https://doi.org/10.1016/j.cpr.2015.09.004>
- RED Medical (2014). RED Medical Systems GmbH.
- Rubin, D. B. (1987). Introduction and summary of repeated-imputation inferences. In N. Hoboken (Ed.), *Multiple imputation for nonresponse in surveys* (pp. 1–26). John Wiley & Sons, Inc.
- Skapinakis, P., Lewis, G., Davies, S., Brugha, T., Prince, M., & Singleton, N. (2011). Panic disorder and subthreshold panic in the UK general population: Epidemiology, comorbidity and functional limitation. *European Psychiatry*, 26(6), 354–362. <https://doi.org/10.1016/j.eurpsy.2010.06.004>
- Skevington, S. M., Loffy, M., & O’Connell, K. A. (2004). The World Health Organization’s WHOQOL-BREF quality of life assessment: Psychometric properties and results of the international field trial. A report from the WHOQOL group. *Quality of Life Research*, 13(2), 299–310. <https://doi.org/10.1023/B:QURE.0000018486.91360.00>
- Stech, E. P., Lim, J., Upton, E. L., & Newby, J. M. (2020). Internet-delivered cognitive behavioral therapy for panic disorder with or without agoraphobia: A systematic review and meta-analysis. *Cognitive Behaviour Therapy*, 49(4), 270–293. <https://doi.org/10.1080/16506073.2019.1628808>
- Telch, M. J., Brouillard, M., Telch, C. F., Agras, W. S., & Taylor, C. B. (1989). Role of cognitive appraisal in panic-related avoidance. *Behaviour Research and Therapy*, 27(4), 373–383. [https://doi.org/10.1016/0005-7967\(89\)90007-7](https://doi.org/10.1016/0005-7967(89)90007-7)
- Villatoro, A. P., DuPont-Reyes, M. J., Phelan, J. C., & Link, B. G. (2022). Me versus them: How mental illness stigma influences adolescent help-seeking behaviors for oneself and recommendations for peers. *Stigma and Health*, 7(3), 300–310. <https://doi.org/10.1037/sah0000392>
- Weisel, K. K., Fuhrmann, L. M., Berking, M., Baumeister, H., Cuijpers, P., & Ebert, D. D. (2019). Standalone smartphone apps for mental health—a systematic review and meta-analysis. *Npj Digital Medicine*, 2(1), 118. <https://doi.org/10.1038/s41746-019-0188-8>
- White, K. S., Brown, T. A., Somers, T. J., & Barlow, D. H. (2006). Avoidance behavior in panic disorder: The moderating influence of perceived control. *Behaviour Research and Therapy*, 44(1), 147–157. <https://doi.org/10.1016/j.brat.2005.07.009>

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.