

Mobile-App-Guided Exposure Therapy for Panic Disorder With and Without Agoraphobia: A Randomized Controlled Trial

Matthias Guth, Annika Wiebe, Mogda Ekhlash-Schumann, Alexa Rodenjohnann, Felicia Rohlfesen, Jacqueline Schilbach, Benjamin Selaskowski, Alexandra Philipsen, Niclas Braun

Submitted to: Journal of Medical Internet Research
on: April 24, 2025

Disclaimer: © The authors. All rights reserved. This is a privileged document currently under peer-review/community review. Authors have provided JMIR Publications with an exclusive license to publish this preprint on its website for review purposes only. While the final peer-reviewed paper may be licensed under a CC BY license on publication, at this stage authors and publisher expressly prohibit redistribution of this draft paper other than for review purposes.

Table of Contents

Original Manuscript.....	5
Supplementary Files.....	39
Figures	40
Figure 1.....	41
Figure 2.....	42
Figure 3.....	43
Figure 4.....	44
Multimedia Appendixes	45
Multimedia Appendix 1.....	46
Multimedia Appendix 2.....	46
Multimedia Appendix 3.....	46
Multimedia Appendix 4.....	46
CONSORT (or other) checklists.....	47
CONSORT (or other) checklist 0.....	47
TOC/Feature image for homepages	48
TOC/Feature image for homepage 0.....	49

Mobile-App-Guided Exposure Therapy for Panic Disorder With and Without Agoraphobia: A Randomized Controlled Trial

Matthias Guth^{1*} MD; Annika Wiebe^{1*} MSc; Mogda Ekhlas-Schumann¹; Alexa Rodenjohn¹; Felicia Rohlfen¹; Jacqueline Schilbach¹; Benjamin Selaskowski¹ PhD; Alexandra Philipsen¹ MD; Niclas Braun¹ PhD

¹Department of Psychiatry and Psychotherapy University Hospital Bonn Bonn DE

*these authors contributed equally

Corresponding Author:

Matthias Guth MD

Department of Psychiatry and Psychotherapy

University Hospital Bonn

Venusberg-Campus 1

Bonn

DE

Abstract

Background: Mobile apps that implement disorder-specific psychotherapy for panic disorder with and without agoraphobia (PD/A) can be used in real-life situations that trigger symptoms and, therefore, are a promising novel therapeutic tool.

Objective: This RCT aims to expand the currently limited evidence base for the efficacy of mobile interventions for PD/A by evaluating a mobile app focusing on interoceptive and in-vivo exposure therapy.

Methods: After establishing the diagnosis of PD/A using a secure video communication platform, we randomized 111 adults to three equally sized groups: PD/A-specific exposure therapy app; mindfulness meditation app unrelated to the disorder (active control condition); and waiting list (passive control condition). Participants used the apps self-guided. Additional psychotherapy was not allowed during the study. Self-reported PD/A symptom severity was our primary outcome parameter and measured by the Panic and Agoraphobia Scale (PAS) and symptom subset scales. Secondary outcome parameters included depressive symptoms and quality of life (QoL). We conducted assessments at baseline, after the intervention (allocation + 5 weeks), and at follow-up (allocation + 13 weeks).

Results: We observed significantly lower overall PD/A symptom severity in the exposure app group compared to waiting list post-treatment ($P = .04$, Cohen's $d = 0.55$) and at follow-up ($P = .04$, $d = 0.60$). At follow-up, the exposure app group demonstrated significantly stronger improvements than the waiting list group in depressive symptoms ($P = .007$, $d = 0.75$) and psychological QoL ($P = .01$, $d = 0.63$). We observed no significant differences between the exposure and the meditation apps. In the exposure app group, 35 % of participants showed reliable improvement in overall PD/A symptom severity at post-treatment (meditation app: 6 %, waiting list 7 %). The dropout rate after five weeks was 14 % in the exposure app group (meditation app: 8 %, waiting list 16 %). No adverse outcomes were reported in the exposure app group.

Conclusions: Our findings suggest that app-guided exposure therapy can be a useful further treatment option in addition to established psychological and pharmacological strategies. However, superiority compared to the disorder-unspecific meditation app remains unclear. Clinical Trial: German Clinical Trials Register (DRKS00022204)

(JMIR Preprints 24/04/2025:76389)

DOI: <https://doi.org/10.2196/preprints.76389>

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

Please make my preprint PDF available to anyone at any time (recommended).

✓ Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.

Only make the preprint title and abstract visible.

No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain visible to the public.

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <https://www.jmir.org/preprint/76389>, the full text will be available to the public.

No. Please do not make my accepted manuscript PDF available to anyone. I understand that if I later pay to participate in <https://www.jmir.org/preprint/76389>, the full text will be available to the public.



Original Manuscript

Original Paper**Mobile-App-Guided Exposure Therapy for Panic Disorder With and Without Agoraphobia: A Randomized Controlled Trial****Authors**

Matthias Guth^{*a}, Annika Wiebe^{*a}, Mogda Ekhlas-Schumann^a, Alexa Rodenjohann^a, Felicia Rohlfen^a, Jacqueline Schilbach^a, Benjamin Selaskowski^a, Alexandra Philipsen^a, Niclas Braun^a

* both authors equally contributed to this work.

Affiliations

a) Department of Psychiatry and Psychotherapy, University Hospital Bonn, Venusberg-Campus 1, 53127 Bonn, Germany

Corresponding author:

Dr. Matthias Guth

University Hospital Bonn, Department of Psychiatry and Psychotherapy, Venusberg-Campus 1, 53127 Bonn, Germany

Tel.: + 49 228 287-15723

Fax: + 49 228 287-16097

mguth@uni-bonn.de

Abstract

Background: Mobile apps that implement disorder-specific psychotherapy for panic disorder with and without agoraphobia (PD/A) can be used in real-life situations that trigger symptoms and, therefore, are a promising novel therapeutic tool.

Objective: This RCT aims to expand the currently limited evidence base for the efficacy of mobile interventions for PD/A by evaluating a mobile app focusing on interoceptive and in-vivo exposure therapy.

Methods: After establishing the diagnosis of PD/A using a secure video communication platform, we randomized 111 adults to three equally sized groups: PD/A-specific exposure therapy app; mindfulness meditation app unrelated to the disorder (active control condition); and waiting list (passive control condition). Participants used the apps self-guided. Additional psychotherapy was not allowed during the study. Self-reported PD/A symptom severity was our primary outcome parameter and measured by the Panic and Agoraphobia Scale (PAS) and symptom subset scales. Secondary outcome parameters included depressive symptoms and quality of life (QoL). We conducted assessments at baseline, after the intervention (allocation + 5 weeks), and at follow-up (allocation + 13 weeks).

Results: We observed significantly lower overall PD/A symptom severity in the exposure app group compared to waiting list post-treatment ($P = .04$, Cohen's $d = 0.55$) and at follow-up ($P = .04$, $d = 0.60$). At follow-up, the exposure app group demonstrated significantly stronger improvements than the waiting list group in depressive symptoms ($P = .007$, $d = 0.75$) and psychological QoL ($P = .01$, $d = 0.63$). We observed no significant differences between the exposure and the meditation apps. In the exposure app group, 35 % of participants showed reliable improvement in overall PD/A symptom severity at post-treatment (meditation app: 6 %, waiting list 7 %). The dropout rate after five weeks was 14 % in the exposure app group (meditation app: 8 %, waiting list 16 %). No adverse outcomes were reported in the exposure app group.

Conclusions: Our findings suggest that app-guided exposure therapy can be a useful further treatment option in addition to established psychological and pharmacological strategies. However, superiority compared to the disorder-unspecific meditation app remains unclear.

Trial registration: German Clinical Trials Register (DRKS00022204)

Keywords:

Panic disorder; Agoraphobia; Mobile intervention; Digital health; Mobile health; Exposure therapy.

1 Introduction

A panic disorder is characterized by the repeated occurrence of unexpected panic attacks, by worries about the recurrence or consequences of another panic attack, and/or by behavioral changes related to these worries [1]. Panic disorders often occur comorbidly with agoraphobia, which is defined by a strong fear and/or avoidance of situations in which escape is difficult or help is not immediately available in the event of panic-like symptoms, such as being in public or crowded spaces or out of the house alone [1]. According to the World Mental Health Surveys, the estimated cross-national lifetime prevalence of panic disorder is approximately 1.7 % [2] and that of agoraphobia is 1.5 % [3]. Both disorders have a high risk of becoming chronic if left untreated [1].

International guidelines recommend treating panic disorders with and without agoraphobia (PD/A) with cognitive behavioral therapy (CBT) and/or pharmacotherapy [4]–[7]. Psychological models highlight the role of dysfunctional appraisal of panic symptoms as severely threatening and the presence of fear of fear in PD/A [8]. Exposure therapy aims at correcting these maladaptive psychological patterns by systematically confronting patients with their panic symptoms and associated fears. Interoceptive exposure and in vivo exposure are considered core features of CBT for PD/A, and a component network meta-analysis identified interoceptive exposure as a particularly effective component of such therapies [7], [9].

A range of internet- or mobile-based interventions (IMIs) has been developed for PD/A, and international guidelines advise to use CBT based IMIs either as a stand-alone treatment or as a companion intervention during psychotherapy [5]–[7]. The terms digital CBT (dCBT) or internet-based CBT (ICBT) are frequently used as alternative terminology for this type of IMIs [5], [7]. IMIs are generally considered a useful and important addition to the spectrum of treatments available for PD/A and other anxiety disorders. They can be used flexibly at all locations and times, offer the possibility of pseudonymous use, and might empower patients by increasing perceived control and self-efficacy. Further, they can be tailored to specific audiences (e.g., by offering multiple languages) and, in particular in the case of mobile interventions, could improve the transfer of therapeutic content into daily life, for example by facilitating regular exercises [10], [11]. The latter could be especially relevant for PD/A, as patients often experience difficulties leaving their homes and exposure exercises are a key therapeutic component.

Two recent meta-analyses [12], [13], with 16 and 13 included controlled studies respectively, investigated the effectiveness of IMIs for individuals with PD/A. These meta-analyses identified only two RCT that investigated mobile interventions, and none of these studies focused specifically on exposure therapy [14], [15]. Both meta-analyses found that compared to passive control groups, such as waiting list (WL), IMIs were highly effective in reducing primary panic and agoraphobia symptoms. Further, medium to large effect sizes in favor of IMIs were also found for secondary outcomes such as reduction in depression and anxiety symptoms and increase in quality of life (QoL). However, none of the trials comparing IMIs and passive control groups included a follow-up assessment. When comparing IMIs to other active treatment groups, mainly face-to-face individual or group CBT, no significant differences were found in panic and agoraphobia symptoms, comorbid depression, anxiety, or QoL at post-assessment or follow-up. Overall, however, the number of eligible trials included in both meta-analyses is relatively small, and existing trials show considerable differences in terms of type, duration, and amount of therapist guidance of the interventions compared. More research, especially the conduction of further high-quality, rigorously conducted RCTs with appropriate control interventions and follow-up assessments is therefore needed; in particular, this applies to studies featuring mobile interventions.

The objective of our longitudinal RCT was to evaluate the efficacy of a mobile app with a particular emphasis on interoceptive and in-vivo exposure for individuals with PD/A. We compared a group of patients using the exposure therapy app with a passive (WL) and an active control group using a non-

specific mobile app for mindfulness meditation. Changes in panic and agoraphobia symptoms were the primary outcome parameters. Secondary outcome parameters included symptoms of depression, anxiety, stress, QoL, and self-efficacy expectations. We hypothesized that using the exposure therapy app would result in a significant symptom reduction, particularly in comparison to the WL group, in which we generally did not expect symptom improvements. Furthermore, we anticipated that the exposure therapy app would lead to greater symptom reduction than the meditation app, since the latter is not an intervention specifically designed for the treatment of PD/A. Our study extends the current evidence in several important ways: First, it evaluates a purely mobile IMI, a type of intervention that is still underrepresented in recent meta analyses [12], [13]. Second, by evaluating two control conditions – a passive WL and an active, but not disorder-specific intervention – the study might help to disentangle general IMI effects from specific effects of mobile delivered exposure therapy in PD/A. And third, in contrast to previous IMI studies that typically studied broad CBT interventions featuring a wider range of therapeutic interventions from cognitive restructuring to stress management, the current RCT specifically focused on facilitating exposure, a key component of disorder-specific CBT in PD/A.

2 Methods

2.1 Design

In this RCT, participants were randomly allocated to three equally sized groups: a group that received app-based exposure therapy, a group that received app-based mindfulness meditation (active control), and a WL group (passive control). Panic and agoraphobia symptoms, depression and anxiety, QoL, anxiety sensitivity, and self-efficacy expectations were assessed before the intervention (T_0), after the five-week intervention phase (T_1), and at follow up 13 weeks after T_0 (T_2).

The study was conducted in accordance with the Helsinki Declaration, approved by the University of Bonn's medical ethics committee (protocol number: 321/20), and preregistered in the German Clinical Trials Register (DRKS00022204).

2.2 Participants and Procedure

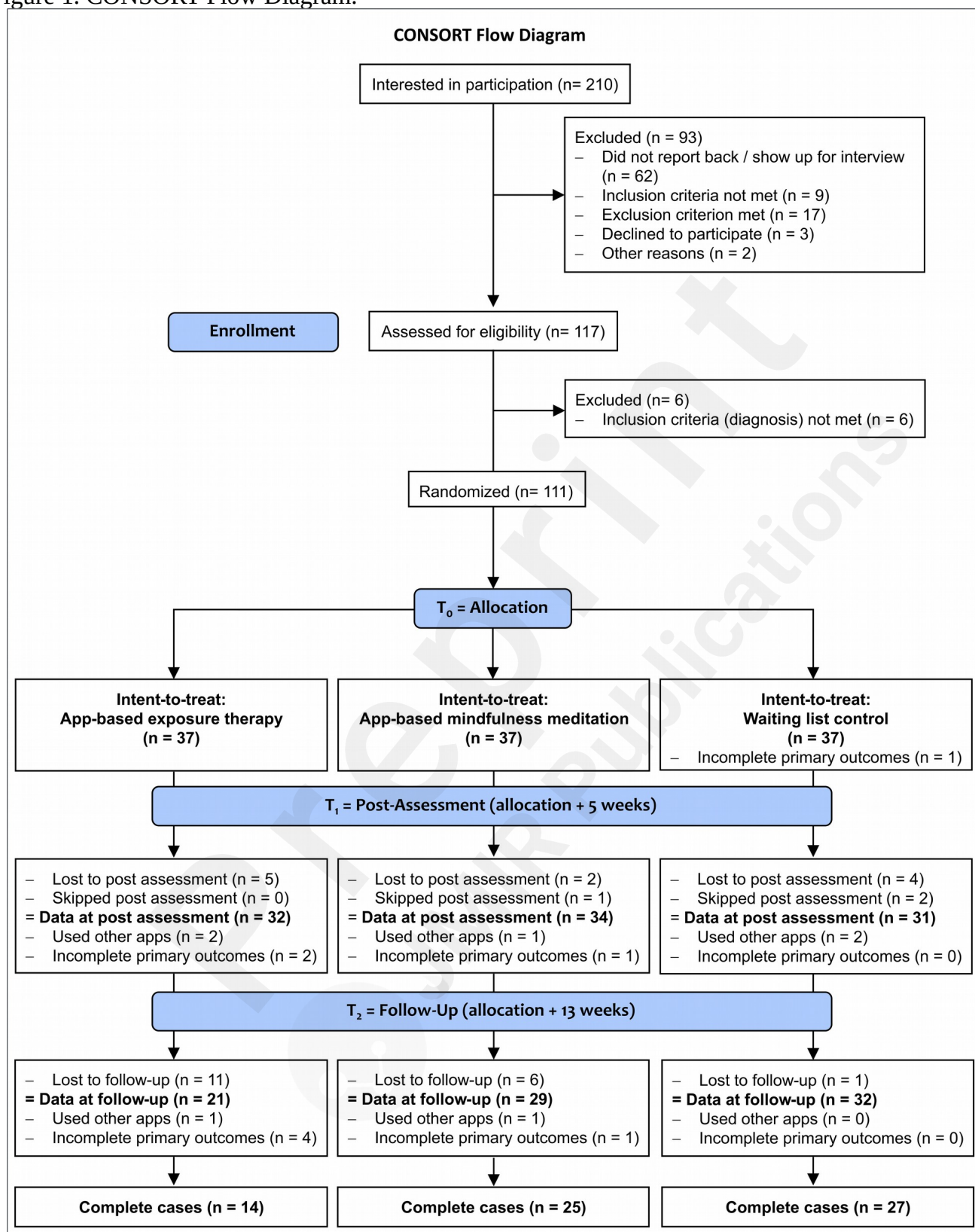
Individuals between the ages of 18 and 60 years who met the diagnostic criteria for a PD/A were eligible to participate in the study. Current psychotherapeutic treatment (face-to-face individual or group therapy) was an exclusion criterion. Other exclusion criteria were acute psychosis, pregnancy, unstable cardiovascular disease, inadequately controlled hypertension, other medical conditions that precluded the administration of exposure therapy, and insufficient knowledge of the German language, as we evaluated German versions of the studied apps. Mobile device literacy was not explicitly specified as an inclusion criterion but implicitly required for study participation. Participants were allowed to continue psychiatric medication during the study.

Recruitment and data collection for the study took place between November 2020 and January 2023 (first participant in: November 17, 2020; last participant out: January 11, 2023). Participants were recruited via the outpatient clinic of the Department of Psychiatry and Psychotherapy of the University Hospital Bonn; by advertising among psychotherapists, psychiatrists, general practitioners, self-help groups, psychological study counseling; and via social media and newspapers. Since the study was conducted entirely online, the participants were recruited nationwide in Germany as well as in Austria and Switzerland.

For the complete participant flow, see Figure 1. A total of 210 individuals expressed interest in the study and were contacted by the study team. The study team outlined the study design and presented the inclusion and exclusion criteria. After this initial step, 117 individuals progressed further and underwent the full eligibility assessment. All participants gave their written informed consent. To establish the diagnosis of PD/A and to determine comorbid disorders, participants underwent a structured clinical interview at the baseline assessment (T_0). The clinical interview was based on the Brief Diagnostic Interview for Mental Disorders Mini-DIPS, [16] and carried out remotely, using an secure video communication platform [17]. If the diagnostic criteria were not met or any exclusion criterion was fulfilled, study participation was terminated at this point. If participants met the diagnostic criteria for PD/A and did not show any exclusion criterion, we randomly assigned them to one of the three groups by drawing from an urn. The study team prepared the urn prior to the start of the study, and the allocation sequence was unpredictable. After screening, 111 individuals were included and assigned to one of the three groups; recruitment was stopped after reaching this sample size, which was pre-specified based on a power calculation (tool used: G*Power [18]; parameters used: effect size = 0.253, α = 0.05, power = 0.95, number of groups = 3, number of measurements = 2, anticipated dropout rate = 40 %). Following the baseline assessment, the five-week intervention period started for all included participants. At the end of the intervention period, all participants received a link to a post-intervention survey via email (T_1). At T_2 , all participants received a link to a follow-up survey. Once the participants had completed this questionnaire, participation in the study was completed. Participants in the WL group equally underwent assessments at T_0 , T_1 , and T_2 and could request access to one of the two apps following T_2 .



Figure 1. CONSORT Flow Diagram.



2.3 Enrollment and Assessment

Six members of the study team rated the Mini-DIPS interview and performed randomization (one physician, one post-graduate psychologist, four advanced-level psychology/medical students under supervision). Raters received training and conducted mock interviews prior to the first rating. All self-report data was obtained via an online survey tool [19].

2.3.1 Primary Study Outcomes

Primary outcomes were panic and agoraphobia symptoms based on the following self-reported questionnaires:

Panic and Agoraphobia Scale (PAS): The PAS is a 13-item scale measuring overall symptom severity in individuals diagnosed with PD/A [8]. The 13 items cover the frequency and severity of panic attacks, agoraphobic avoidance, anticipatory anxiety, functional impairment, and worry related to somatic health. The PAS is recommended by guidelines [5], [20] and has been widely used in clinical trials e.g., [21]–[23]. The total sum score ranges from 0 to 52, with scores above 8 indicating mild, above 18 intermediate, above 28 severe and above 39 very severe symptoms.

Agoraphobic Cognitions Questionnaire (ACQ) and Body Sensations Questionnaire (BSQ): The ACQ and the BSQ are companion measures focusing on the concept of fear of fear in PD/A [24]–[26]. The ACQ thereby addresses negative/catastrophic thoughts associated with fear and nervousness, including the expectation of severe somatic illness such as stroke or brain tumor, and the expectation of loss of control. The ACQ is provided as a 14-item scale and its score ranges from 1 to 5. The BSQ, in turn, assesses anxiety associated with bodily sensations typically present during panic attacks (e.g., palpitations, dizziness, or depersonalization) on a 17-item scale and its score ranges from 1 to 5.

Mobility Inventory (MI): The MI [26]–[28] outlines 27 situations often avoided by individuals with PD/A like supermarkets, busses or crowded places. Participants rate their degree of avoidance for each of these situations separately for avoidance alone and avoidance when accompanied by a trusted person. The derived scores range from 1 to 5.

Texas Safety Maneuver Scale (TSMS): The TSMS [29]–[31] measures dysfunctional strategies employed to reduce anxiety and panic in situations perceived as dangerous (e.g., checking the location of nearby hospitals or distracting oneself through music). The TSMS lists 50 common safety maneuvers and prompts participants to rate the degree to which they use these strategies. The possible overall score ranges between 0 and 250.

2.3.2 Secondary Study Outcomes

Secondary outcomes were general anxiety, depression, and stress symptoms, QoL, anxiety sensitivity, and self-efficacy expectations, and assessed by the following self-reported questionnaires:

World Health Organization Quality of Life (WHO-QOL-BREF): This self-report questionnaire developed by the WHO [32], [33] assesses the global health status of patients across four health domains: physical health, psychological QoL, social relationships and environmental QoL. The total scores for each domain range from 0 to 100.

Depression, Anxiety and Stress Scale – 21 items (DASS-21): The DASS-21 [34], [35] consists of 21 items, equally divided into three subscales – depression, anxiety and stress. The subscales each range from 0 to 42 (sum score multiplied by two to retain consistency with scoring of the long version DASS).

Anxiety Sensitivity Index 3 (ASI-3): The ASI-3 [36]–[38] measures anxiety sensitivity, a construct similar to fear of fear. The questionnaire contains 18 items, and its overall scale varies between 0 and 72.

General Self-Efficacy Scale (GSES): The GSES [39]–[41] was used to assess perceived general self-efficacy, that is, the extent to which people believe they are able to successfully cope with challenging life situations. The GSES is administered as a 10-items questionnaire and its total sum score ranges from 10 to 40.

2.4 Interventions

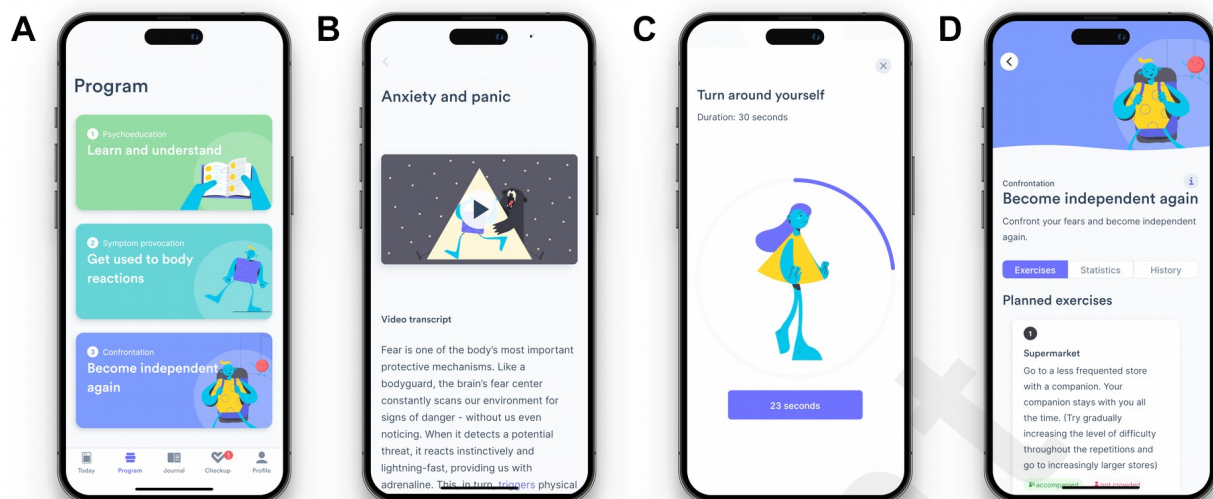
Participants assigned to the exposure app group and to the meditation app group each were instructed to use the respective mobile app for a period of five weeks. During this period, they received no

further instructions from the study team regarding the app content besides a short introduction at baseline (i.e. self-guided therapy). To minimize the risk of performance bias, none of the two apps was presented as superior to the participants during the onboarding procedure. However, we could not guarantee that participants were kept completely blind, as information documents about the study stating that the exposure therapy intervention was the intervention of interest had to be made available online (e.g. in the study registration) and had to be disclosed to obtain informed consent. Both apps were available for Android and iOS devices. After the five-week intervention period, access to the app was not restricted, but no further instructions were given, so that participants could continue using their respective app as they wished.

2.4.1 Exposure Therapy Intervention

Participants in the exposure app group received free access to Mindable: Panic & Agoraphobia [42], an e-health application otherwise available to individuals with PD/A in Germany on prescription [43]. The app is based on the manual for cognitive-behavioral exposure therapy published by Lang et al. [31] and focuses on preparing, accompanying and debriefing various exposure exercises. After working through a base module featuring psychoeducative content (e.g., psychological models of causes, maintaining factors and consequences of the disorder), patients are guided through a second module focusing on interoceptive exposure (e.g., accustoming to self-provoked body sensations such as dizziness or tachycardia) and a third module facilitating in-vivo exposure of agoraphobic situations. For the third module, patients define situations relevant for themselves (e.g., public transportation, supermarket, elevators) and are encouraged to gradually increase exposure difficulty (e.g., first accompanied, later alone). Content is delivered through videos and accompanying texts. Patients interact with the application, for example by answering self-reflection questions, logging anxiety during exposure exercises, and repetitively filling short symptom questionnaires. Patients are instructed to use the app daily and schedule at least one exposure exercise per week. Customizable mobile-operating-system-based notifications reminded participants to frequently use the app and complete scheduled tasks. We recommended participants to work through the three therapy modules chronologically, starting with the psychoeducation module in the first week. A study-specific “frozen” version of the app was used and not changed throughout the study. Figure 2 presents exemplary screenshots of the exposure therapy app. Panel A shows an app overview featuring the three main modules: Psychoeducation, interoceptive exposure and in-vivo exposure. Panel B features an exemplary video transcription from the psychoeducation module, and panel C an exemplary interoceptive exposure exercise focused on the elicitation of dizziness. An exemplary in-vivo exposure exercise guiding a visit to a supermarket together with a trusted person is displayed in panel D. The operating language was set to German for the study.

Figure 2. Visualization of the Exposure Therapy App. Visualization copyright: Mindable Health GmbH.



2.4.2 Mindfulness Meditation Intervention

Participants in the meditation app group received free access to 7Mind [44], a meditation app which is otherwise available in Germany through selected health insurances, some employers, and on a self-payment basis. The application teaches mindfulness meditation through a range of courses (e.g., “foundation”, “from stress to silence”) and themed exercises (e.g., “falling asleep”, “bus or train”). The app dynamically recommends content and provides advice. Participants can schedule meditations within the program and receive reminders. The intervention is not specifically focused on PD/A. Participants were encouraged to use the app daily and to choose courses and relaxation exercises according to their own needs and interests.

2.5 Statistical Analyses

To evaluate the results under real-world conditions, we conducted an intent-to-treat (ITT) analysis for all primary and secondary parameters, in which we included all available data from all randomized participants. In addition, to examine the results under ideal conditions, we conducted an additional completer analysis on all primary outcome parameters (see also 2.6.). In both analyses, separate linear mixed effects models (LMM) were calculated for each analyzed outcome parameter. The model specification controlled for baseline values by including the factor time (T_0 , T_1 , T_2) and the interaction of group (exposure, meditation, WL) and time [45]. Further, participant ID was added as a random effects term, and age, gender, and education served as covariates. Post-hoc tests were conducted by running Bonferroni corrected t-tests on the estimated marginal means from the model. All reported statistical tests are based on a significance level of $\alpha = .05$. Partial Eta-squared (η_p^2) was used as a measure of LMM effect sizes. For significant effects, post-hoc t-test effect sizes were reported as Cohen’s d , using the estimated marginal means and the pooled standard deviation from the observed means. We analyzed reliable changes based on the reliable change index suggested by Jacobson and Truax [46] using normative data for the PAS [8]. Confidence intervals for the number needed to treat (NNT) were estimated using the Wald method [47]. All statistical analyses were conducted in R 4.2.2 [48]. The packages lme4 [49] and emmeans [50] were used for model estimation/post-hoc testing.

2.6 Missing Data Handling and Exclusions

From a total of 111 participants enrolled (37 participants per group), 11 participants had dropped out of the study at T_1 . Furthermore, three participants did not participate at the T_1 assessment but

returned for T₂. In total, T₁ data was missing for five participants in the exposure app group (14 %), three participants in the meditation app group (8 %), and six participants in the waiting list group (16 %). At T₂, an additional 18 participants had dropped out. Additionally, at all three timepoints, individual participants only partially completed the questionnaires. Dropout rates were highest in the exposure app group. For detailed information, see participant flowchart (Figure 1).

For the ITT analysis, data of all 37 participants per group were analyzed, including data from those with missing data at one or more of the timepoints. A detailed summary of the number of complete cases that were included into the ITT analysis for each questionnaire per group and timepoint can be found in the supplementary data S1.

For the completer analysis, we considered only those participants who had no missing data in any questionnaires at all three timepoints. Seven participants (three in exposure app group, two in meditation app group, two in WL group) indicated at the T₁ and/or T₂ assessment that, in addition to the apps evaluated in the study, they had used other meditation or therapy apps to reduce their PD/A symptoms. We also excluded these individuals from the completer analysis. Thus, the completer analysis included 14 participants from the exposure app group, 25 participants from the meditation app group, and 27 participants from the WL group.

3 Results

3.1 Baseline Characteristics

Demographic information, diagnostic status, and baseline symptom scores for all three groups are summarized in Table 1. A total of 111 participants (mean age: 36.78, 73 % female) were randomized. Of these, 59 % suffered from combined PD/A, 22 %, from mere panic disorder and 20 % from mere agoraphobia. Compared to the WL group, significantly more participants ($p = .048$) in the meditation app group had concomitant other anxiety disorders (e.g., social anxiety disorder, generalized anxiety disorder). All other demographic and diagnostic status variables showed no significant group differences. Mean PAS Score across all groups was 22.13 at baseline, corresponding to intermediate symptom severity. A significant baseline group difference was detected for the PAS score, driven by significantly lower scores in the meditation app group compared to the WL group. Of note, however, we controlled for baseline differences in all statistical models (see 2.5.). There were no other significant baseline group differences in any of the other primary or secondary outcomes.

Table 1. Demographic and clinical characterization

	App-based exposure therapy	App-based mindfulness meditation	Waiting list	<i>P</i> ^a
	n = 37	n = 37	n = 37	
Demographics				
Age in years, mean (SD)	39.05 (13.28)	38.49 (12.14)	32.81 (11.63)	.06
range	20 – 60	19 – 59	19 – 57	
Gender				
male	7 (19 %)	9 (24 %)	12 (32 %)	.44
female	29 (78 %)	27 (73 %)	25 (68 %)	
not specified	1 (3 %)	1 (3 %)	0	
Education				
primary or below	0	0	0	.78
lower secondary education	13 (35 %)	7 (19 %)	6 (16 %)	
higher secondary education	12 (32 %)	19 (51 %)	16 (43 %)	
university degree	11 (30 %)	10 (27 %)	14 (38 %)	
postgraduate	1 (3 %)	1 (3 %)	1 (3 %)	
PD/A diagnostic status				
Panic disorder and agoraphobia	25 (68 %)	17 (46 %)	23 (62 %)	.32
Panic disorder only	7 (19 %)	9 (24 %)	8 (22 %)	
Agoraphobia only	5 (14 %)	11 (30 %)	6 (16 %)	
Comorbidities				
Other concomitant anxiety disorder	25 (68 %)	29 (78 %)	19 (51 %)	.048 ^d
Concomitant affective disorder	6 (16 %)	1 (3 %)	6 (16 %)	.43
Other concomitant psychiatric diagnosis	18 (49 %)	12 (32 %)	16 (43 %)	.54
PD/A symptom severity^b				
PAS, mean (SD)	23.3 (7.19)	18.7 (9.89)	24.44 (9.09)	.02 ^e
ACQ, mean (SD)	2.29 (0.60)	2.14 (0.68)	2.11 (0.53)	.41
BSQ, mean (SD)	2.89 (0.62)	2.69 (0.84)	2.94 (0.71)	.30
MI accompanied, mean (SD)	2.09 (0.80)	1.96 (0.87)	1.78 (0.66)	.22
MI alone, mean (SD)	2.63 (0.94)	2.45 (1.11)	2.29 (0.84)	.32
TSMS, mean (SD)	117.97 (42.91)	98.62 (42.47)	105.84 (41.19)	.14
Secondary outcomes^c				
WHO-QOL psychological, mean (SD)	54.84 (15.45)	58.22 (12.36)	55.41 (16.28)	.58
WHO-QOL physical, mean (SD)	66.31 (17.22)	66.80 (15.27)	65.44 (17.82)	.94
WHO-QOL social, mean (SD)	62.84 (20.93)	65.77 (20.39)	64.41 (17.96)	.82
WHO-QOL environmental, mean (SD)	73.40 (12.25)	71.03 (12.73)	74.41 (12.21)	.49
DASS depression, mean (SD)	15.08 (10.88)	11.14 (9.56)	15.30 (10.08)	.15
DASS stress, mean (SD)	21.78 (10.02)	17.62 (9.07)	20.38 (10.25)	.18
DASS anxiety, mean (SD)	20.97 (10.54)	17.57 (9.49)	21.89 (8.88)	.13
ASI, mean (SD)	40.19 (13.37)	36.68 (12.86)	35.86 (12.88)	.32
GSES, mean (SD)	25.35 (4.64)	25.95 (3.95)	26.54 (5.13)	.54

^a P-column states ANOVA (continuous variables)/chi-squared (categorical measures) test results. If $P < 0.05$, we conducted post-hoc tests (pairwise comparisons).

^b Abbreviations: PAS = Panic and Agoraphobia Scale, ACQ = Agoraphobic Cognitions Questionnaire, BSQ = Body Sensations Questionnaire, MI = Mobility Inventory, TSMS = Texas Safety Maneuver Scale

^c Abbreviations: WHO-QOL = World Health Organization Quality of Life, DASS = Depression, Anxiety and Stress Scale, ASI = Anxiety Sensitivity Index, GSES = General Self-Efficacy Scale

^d significant post-hoc test: meditation vs. waiting list ($P = .03$), other post-hoc test not significant.

^e significant post-hoc test: meditation vs. waiting list ($P = .02$), other post-hoc test not significant, statistical models controlled for baseline differences (see below).

3.2 Primary Outcomes

Changes in PD/A symptom severity during the study are depicted in Figure 3. In the exposure app group, scores decreased from T_0 to T_1 as well as from T_1 to T_2 on all six symptom severity scales. Also, the meditation app group showed a decrease in all six symptom scores from T_0 to T_1 , however,

here the decreases were descriptively smaller, and they increased again at T₂ in four of the measures (overall PD/A symptoms, agoraphobic cognitions, fear of bodily sensations and safety maneuvers). The WL group showed varying courses of outcome with no clear overall tendencies.



Figure 3. Changes in primary outcome means in each group across the three timepoints. Error bars represent the standard error of the mean.

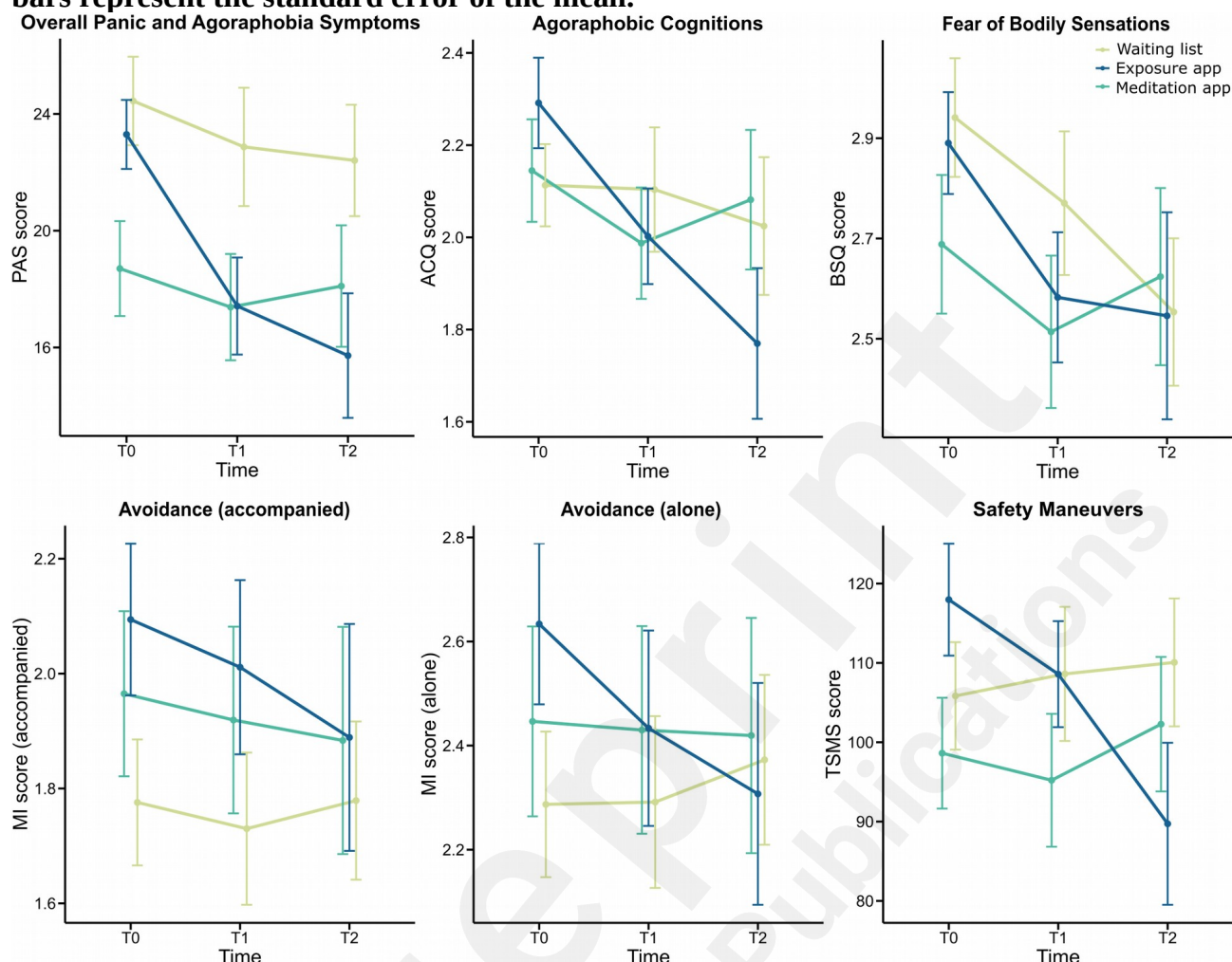


Table 2 summarizes the LMM results of our ITT analysis (for more detailed results, see supplementary data S2). For overall PD/A symptoms (PAS), we found a significant time $\times$ group interaction ($F(6,148.48) = 3.00, P = .009, \eta_p^2 = 0.11$). Post-hoc t-tests revealed significant differences between the exposure app group and the WL group at both T₁ ($t(177) = 2.48, P = .04, d = 0.55$) and T₂ ($t(211) = 2.50, P = .04, d = 0.60$), but no other significant differences in group comparisons. In addition, for the exposure app group, we observed significant reductions in overall PD/A symptoms between T₀ and T₁ ($t(174) = 4.24, P < .001, d = 0.72$) and between T₀ and T₂ ($t(181) = 4.18, P < .001, d = 0.90$), while for all other within-group comparisons we did not find any significant differences. For safety maneuvers (TSMS), we likewise found a significant time $\times$ group interaction ($F(6,142.69) = 2.88, P = .01, \eta_p^2 = 0.11$). Here our post-hoc t-tests did not reveal any significant between-group effects, but significant within-group decreases for the exposure app group between T₀ and T₁ ($t(172) = 2.74, P = .02, d = 0.32$) and T₀ and T₂ ($t(175) = 4.69, P < .001, d = 0.62$). All other within-group comparisons revealed no significant differences.

Regarding agoraphobic cognitions, fear of bodily sensations, and avoidance, we found no significant time $\times$ group interactions in the ITT analysis. However, across groups the LMM for fear of body sensations (BSQ) revealed a significant main effect of time ($F(2, 173.61) = 5.06, P = .007, \eta_p^2 = 0.06$).

In the additional completer analysis, we again found a significant time $\times$ group interaction ($F(6,91.90) = 2.69, P = .02, \eta_p^2 = 0.15$) for overall PD/A symptom severity (PAS). When following up the interaction with post-hoc t-tests, we, however, did not find significant between-group differences at any of the three timepoints. Regarding within-group comparisons, we found a

significant reduction in overall PD/A symptom severity in the exposure app group between T_0 and T_1 ($t(126) = 3.50$, $P = .002$, $d = 0.79$) and between T_0 and T_2 ($t(126) = 3.65$, $P = .001$, $d = 0.84$), but no other within-group comparison reached level of significance. Regarding safety maneuvers (TSMS), we again observed a significant time $\times$ group interaction ($F(6,91.90) = 2.94$, $P = .01$, $\eta_p^2 = 0.16$) that we followed up using post-hoc t-tests. These indicated a significant within-group decrease for the exposure app group between T_0 and T_2 ($t(126) = 3.95$, $P < .001$, $d = 0.56$), while all other post-hoc tests did not show significant effects. Detailed results of the completer analysis can be found in supplementary data S3.



Table 2. Estimated mean differences of primary outcomes – intention-to-treat analyses

Comparison ^a	T ₁ (allocation + 5 weeks)				T ₂ (allocation + 13 weeks)			
	Δ (SE) ^b	95%-CI ^b	P	d ^b	Δ (SE) ^b	95%-CI ^b	P	d ^b
Overall panic and agoraphobia symptoms								
Exposure App – Waiting list	-5.64 (2.27)	-11.13 to -0.14	.04	0.55	-6.15 (2.46)	-12.10 to -0.21	.04	0.60
Exposure app – Meditation app	-1.16 (2.21)	-6.50 to 4.19	1.00	0.12	-2.65 (2.45)	-8.56 to 3.26	.84	0.25
Waiting list – Meditation app	4.48 (2.23)	-0.91 to 9.87	.14	-0.41	3.50 (2.26)	-1.95 to 8.96	.37	-0.32
Safety maneuvers								
Exposure app – Waiting list	0.74 (9.98)	-23.44 to 24.93	1.00	-0.02	-15.90 (10.48)	-41.24 to 9.45	.39	0.35
Exposure app – Meditation app	4.79 (9.72)	-18.76 to 28.35	1.00	-0.11	-12.81 (10.36)	-37.85 to 12.23	.65	0.29
Waiting list – Meditation app	-4.05 (9.82)	-19.73 to 27.84	1.00	-0.08	3.90 (9.89)	-20.86 to 27.04	1.00	-0.07

^a All models control for baseline differences.

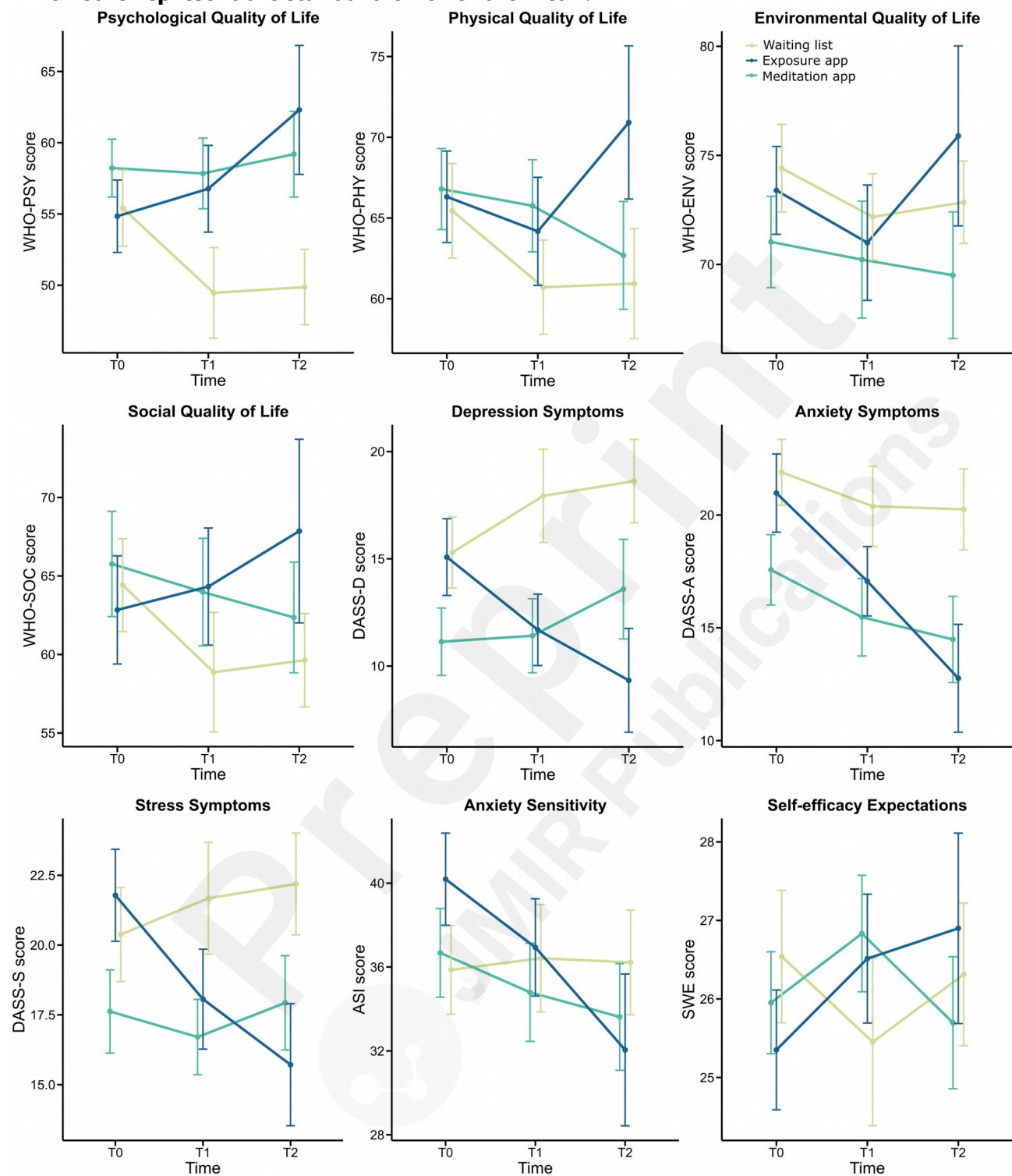
^b Δ = Group differences based on estimated marginal means, SE = standard error; CI = confidence interval; d = Cohen's d effect size.

3.1 Secondary Outcomes

Changes in secondary outcomes are depicted in Figure 4. Descriptively, the exposure app group improved in all secondary outcomes over the course of the study, while most measures stayed stable or even deteriorated in the other two groups.



Figure 4. Changes in secondary outcome means in each group across the three timepoints. Error bars represent the standard error of the mean.



In the ITT analysis, the LMM on psychological QoL (as measured by the psychological subscale of the WHO-QOL) revealed a significant main effect of time ($F(2, 175.68) = 4.45, P = .01, \eta_p^2 = 0.05$) as well as a significant time x group interaction ($F(6, 144.29) = 3.12, P = .007, \eta_p^2 = 0.11$). The post-hoc t-tests attributed this interaction to a significantly higher psychological QoL in the exposure app group than in the WL group at T_2 ($t(160) = -2.88, P = .01, d = 0.63$), since all other between-groups comparisons showed no significant differences. Moreover, for the exposure app group, we found significant within-group increases in the psychological QoL between T_0 and T_2 ($t(178) = -3.14, P = 0.006, d = 0.36$) and between T_1 and T_2 ($t(175) = -2.98, P = .01, d = 0.32$). For the WL group, in turn, we found significant decreases in psychological QoL between T_0 and T_1 ($t(175) = 2.54, P = .04, d = -0.26$) and between T_0 and T_2 ($t(175) = 2.59, P = .03, d = -0.28$), while all other within-group comparisons showed no significant differences.

Further, there was a significant time x group interaction for depressive symptoms as measured by the DASS questionnaire ($F(6, 148.84) = 3.00, P = .009, \eta_p^2 = 0.11$). Our post-hoc between-group comparison thereby revealed significantly lower depressive symptoms in the exposure app group compared to the WL group at T_2 ($t(206) = 3.09, P = .007, d = 0.75$). Moreover, within the exposure app group, we found a significant decrease in the severity of depressive symptoms from T_0 to T_2 ($t(184) = 2.87, P = .01, d = 0.48$), while all of our other comparisons did not show any significant effects.

Regarding our ITT analyses on all remaining secondary outcomes (physical, social, and environmental QoL; stress and overall anxiety; anxiety sensitivity; perceived general self-efficiency), we found no significant time effects or time x group interactions. Detailed results of our analyses for secondary outcomes can be found in supplementary data S4.

3.4 Response and Safety Analyses

Because we found a significant between-group effect for overall PD/A symptoms, we analyzed individual change for each participant based on this outcome (see Table 3 and Table 4). At T_1 , we observed reliable improvement for 35 % of participants in the exposure app group (meditation app: 6 %, WL: 7 %). 23 % of participants in the exposure app group showed reliable improvement to a score below the PAS cut-off value of 8 suggested for remission [8] compared to 3 % in the WL group and no participant in the meditation app group. Compared to WL, these differences translate into a NNT of 3.47 for reliable improvement and 5.20 for reliable improvement to remission (compared to meditation app: NNT = 3.39 and 4.43). Follow-up data showed a similar pattern with a NNT of 3.85 for reliable improvement compared to WL and a NNT of 3.16 compared to the meditation app. However, three remitted participants in the exposure app group dropped out between T_1 and T_2 and we observed lower rates of remission at T_2 compared to T_1 for the exposure group. NNT confidence intervals for this measure included $\pm \infty$, implying that no significant differences in remission rates were observed between the groups at T_2 .

To assess the safety of the meditation app, we also analyzed data for reliable individual deterioration. At T_1 , one participant (3 %) in the exposure app group showed reliable deterioration (WL: 3%, meditation app: 0 %). The observed score for this participant increased from 12 (mild symptoms) to 24 (intermediate symptom severity). At T_2 , reliable deterioration was not observed in the exposure app group (WL: 6%, meditation app: 10 %). Between-group NNT comparisons for reliable deterioration showed no significant group differences (all confidence intervals included $\pm \infty$). One participant from the WL group and one participant from the meditation app group were admitted to hospital and subsequently dropped out of the study. No adverse events were reported to the study team in the exposure app group.

Table 3. Response Analyses – Overall panic and agoraphobia symptoms

	Post-Assessment (allocation + 5 weeks)		Follow-up (allocation + 13 weeks)	
	<i>n</i>	<i>Percentage</i>	<i>n</i>	<i>Percentage</i>
Reliable improvement^a				
Exposure app	11	35 %	7	39 %
Meditation app	2	6 %	2	7 %
Waiting list	2	7 %	4	13 %
Reliable improvement to remission^a				
Exposure app	7	23 %	3	17 %
Meditation app	0	0 %	1	3 %
Waiting list	1	3 %	2	6 %
Reliable deterioration^a				
Exposure app	1	3 %	0	0 %
Meditation app	0	0 %	3	10 %
Waiting list	1	3 %	2	6 %

^a Reliable change calculated based on the reliable change index suggested by Jacobson and Truax [46] using normative data for the Panic and Agoraphobia Scale [8].

Table 4. Response Analyses – Overall panic and agoraphobia symptoms - NNT

	Post-Assessment (allocation + 5 weeks)	Follow-up (allocation + 13 weeks)
	<i>NNT</i> ^a	<i>NNT</i> ^a
Reliable improvement ^b		
Exposure app – waiting list	3.47 [2.09 – 10.25]	3.85 [2.09 – 10.25]
Exposure app – meditation app	3.39 [2.07 – 9.10]	3.16 [1.78 – 13.06]
Waiting list – meditation app	127.50 [-∞, – -8.98] ∪ [7.87, ∞]	16.65 [-∞, – -11.15] ∪ [4.77, ∞]
Reliable improvement to remission ^b		
Exposure app – waiting list	5.20 [2.83 – 31.37]	9.79 [-∞, – -11.05] ∪ [3.39, ∞]
Exposure app – meditation app	4.43 [2.68 – 12.72]	7.57 [-∞, – -19.10] ∪ [3.16, ∞]
Waiting list – meditation app	30.00 [-∞, – -32.36] ∪ [10.25, ∞]	33.30 [-∞, – -12.66] ∪ [7.19, ∞]
Reliable deterioration ^b		
Exposure app – waiting list	31.00 [-∞, – -33.40] ∪ [10.59, ∞]	-15.50 [-∞, – -6.62] ∪ [45.52, ∞]
Exposure app – meditation app	351.33 [-∞, – -12.29] ∪ [11.48, ∞]	-9.67 [-∞, – -4.67] ∪ [135.24, ∞]
Waiting list – meditation app	-34.00 [-∞, – -11.60] ∪ [36.52, ∞]	-25.69 [-∞, – -5.57] ∪ [9.84, ∞]

^a NNT = number needed to treat, 95 % confidence intervals in square brackets.

^b Reliable change calculated based on the reliable change index suggested by Jacobson and Truax [46] using normative data for the Panic and Agoraphobia Scale [8].

4 Discussion

In this longitudinal RCT we evaluated the efficacy of an exposure therapy app for panic disorder and agoraphobia (PD/A) in comparison to an active (mindfulness meditation app) and passive (WL) control group. We found that the primary outcome parameter overall PD/A symptom severity, was significantly lower in the exposure app group compared to WL at both post-treatment and follow-up. Regarding secondary outcomes, we found significant between-group effects for the exposure app versus waiting list comparisons in depressive symptoms and psychological QoL at follow-up.

4.1 Primary Outcomes

The findings demonstrate the efficacy of the exposure app as a treatment for core symptoms of PD/A. Between group effect sizes compared to WL were medium according to the classification suggested by Cohen [51]. Treatment effects were sustained throughout the entire study period of 13 weeks. PD/A symptom scores declined even further between post-assessment and follow-up, suggesting potential for further therapeutic gains in long-time users. Compared to WL, participants who used the exposure therapy app showed stronger improvement in overall severity of PD/A symptoms. This finding is in line with a recently published RCT that investigated the same exposure therapy app and also found significantly greater improvement in overall PD/A severity in the exposure app group compared to a waiting list control group [52].

Further, in the exposure app group we also found a reduction in maladaptive safety maneuvers. These strategies have been found to contribute to the persistence of symptoms by reaffirming the assessment that the associated situations are dangerous and by preventing full exposure and subsequent anxiety reduction [29], [53]. Safety maneuvers are explicitly addressed within the exposure app and our findings suggest that participants engaged in less of these maneuvers following exposure app use. Descriptively, scores for the remaining symptom clusters associated with PD/A (agoraphobic cognitions, fear of bodily sensations, avoidance behavior) also declined within the exposure app group, but the effects failed to reach statistical significance. As the exposure therapy app possibly targets all these symptom subsets through framed interoceptive and in-vivo exposure, the overall improvement in PD/A symptoms observed might have been the result of the sum of smaller improvements in all these symptom subsets. In consequence, the study might not have been sufficiently powered to detect significant group differences for the possibly smaller improvements in symptom subset scores.

In addition to comparing the app-based exposure therapy to WL, we evaluated its efficacy against an alternative active control intervention, the mindfulness meditation app. The mindfulness meditation app was not specifically designed to treat PD/A but provides a similar experience of using a high-quality digital intervention. In theory, comparing the two apps should permit to separate general effects of using digital interventions (e.g., stemming from the sense of being able to self-therapy symptoms with a novel type of intervention) from specific effects of app-guided exposure therapy. Descriptively, symptom scores declined more in the exposure app group compared to the mindfulness meditation app group. Additionally, significant within-group symptom reductions as well as significantly lower post-treatment symptoms compared to WL were only observed for the exposure therapy app. These observations seem to support the assumption that app-guided exposure therapy is specifically suitable as a digital intervention in PD/A. However, direct comparisons between the exposure app and the meditation app did not demonstrate statistically significantly greater symptom improvement in the exposure app group. One possible explanation might be that the meditation app itself had additional specific therapeutic effects exceeding those generally associated with app usage. These may have mitigated exposure therapy effects in between-group comparisons, which is supported by previous research that found positive effects of mindfulness meditation in anxiety disorders including PD/A [54], [55]. Another possible explanation is that the degree to which

exposure therapy effects exceed general effects associated with mobile intervention usage was too small to be detected with our study design and sample size. As previous literature on similar comparisons is very limited [12] it is difficult to judge what effect size is to be expected for this type of comparison and our study might have been underpowered. Another aspect to consider is that recent meta-analyses showed that digital interventions in PD/A are often provided as guided interventions [12]. Also, the exposure therapy app studied here was built based on material from a therapeutic manual intended to be used with therapist guidance [31]. The effects of using the intervention unguided might have been smaller than those of a guided intervention and this might have resulted in insignificant comparisons with the meditation app. Furthermore, despite randomization, baseline PD/A symptom severity scores were significantly lower in the meditation app group. Meditation might be particularly difficult to practice when symptoms are more severe. Therefore, meditation could have lower effects among participants with higher symptom scores, a subgroup that was underrepresented in the meditation app group of our study. In addition, our between-app comparisons suggested that advantages of the exposure app become more pronounced at follow-up compared to post-assessment. This suggests that our overall study period might have been too short to detect significant between-group effects of app-guided exposure therapy compared to the meditation app.

4.2 Secondary Outcomes

Compared to WL, psychological QoL and depressive symptoms improved significantly among exposure therapy app participants at follow-up but not at post-treatment. The observed delayed significant improvement in these secondary outcomes might suggest that app-guided exposure therapy effects on these measures might be downstream effects of initial improvements in core PD/A symptoms. The observed effects are partially consistent with the results of the two previous meta-analyses [12], [13] which also found that IMIs for PD/A improved QoL and comorbid depressive symptoms. Of note, however, in the current study, the QoL effect was specifically found for psychological QoL, while for the remaining QoL domains, no significant effect was observed. Psychological QoL might be the primary domain of improvement as it is the domain most closely related to mental health. This view is supported by our finding that, across all QoL domains, the lowest baseline values were consistently reported in the psychological domain for all groups. Also, a recent validity analysis of the WHO-QOL [56] found that, out of all QoL domains, the psychological domain is the best predictor for overall QoL.

Of note, no significant treatment effects were found for comorbid general anxiety, although in the previous meta-analyses [12], [13], similar positive effects of IMIs were found for anxiety as for comorbid depression and QoL. Also, we found no effects for anxiety sensitivity and perceived general self-efficacy. These findings are unexpected, given that CBT interventions, including interventions with interoceptive exposure/symptom provocation, seem to effectively reduce anxiety sensitivity [57] and increases in self-efficacy as well as decreases in anxiety sensitivity appear to predict subsequent symptom severity reductions in CBT for PD/A [58]. Potential explanations for these findings might, however, be insufficient statistical power to detect such changes in secondary outcomes and divergent mechanisms of change in IMIs compared to classical CBT. Furthermore, the focused nature of the therapeutic material featured in the exposure therapy app might also have reduced its ability to improve anxiety symptoms in a broader sense beyond improvements in core PD/A symptoms.

4.3 Response and Safety Analyses

Our analyses estimated that approximately one in four participants achieved substantial improvement in the exposure app group compared to WL and the meditation app, and that rates for reliable improvement and remission were highest in this group. As no adverse events were reported for the

exposure app and rates of deterioration were lower or comparable to the other groups, our results indicate that using the exposure app is safe for individuals with PD/A. In sum, these findings suggest that the exposure app can be a valuable treatment option in addition to established psychological and pharmacological strategies. However, to avoid patient demotivation, clinicians should clearly indicate that, while some patients might greatly benefit from using the exposure app, many patients might also see no relevant change in their symptom severity.

4.4 Limitations and Future Directions

One major limitation of the current study is the high dropout rate, especially the high rate of participants who were lost at follow-up in the exposure app group (~27%). This attrition resulted in complete data sets for only 14 participants in this group and clearly reduced statistical power. The particularly high number of dropouts in the exposure app group might be caused by the study design. After completing the core part of the intervention, some participants in the exposure app group presumably no longer saw any benefit in continuing their study participation. This might be particularly applicable for those participants who responded best, as three of seven participants who had reached remission according to the PAS at post-assessment subsequently dropped out of the study. Conversely, the WL group exhibited the highest completer rate, as they might have been highly motivated to complete the study to gain access to the treatment. Future studies should further investigate long-term effects and implement strategies that increase participants' incentives to complete the study, regardless of their group allocation.

Secondly, overall PD/A symptom severity at baseline was intermediate in this study according to the classification of Bandelow [8]. In addition, due to the lack of stratified randomization, group differences in baseline symptom severity were observed in that the meditation group had lower overall PD/A symptom scores than the other groups. While we controlled for symptom severity at baseline in our statistical models, it is still possible that treatment effects of both the exposure therapy app as well as the meditation app to some degree depend on the symptom severity level. For example, some types of exercises, such as longer meditations or longer in-vivo exposure sessions, might be difficult for participants with very severe symptoms. Therefore, further studies of relatively small or moderate sample sizes should use a stratified randomization procedure and explore the effects of app-guided exposure therapy systematically in samples with high and/or low symptom severity.

A further caveat of the current study is the relatively short intervention period of five weeks. Previous RCTs on IMIs for PD/A featured in the meta-analysis of Polak et al. [13] typically used treatment periods ranging from 6 to 12 weeks. This may also be reflected in the delayed treatment effects for our secondary outcomes, which only became apparent at follow-up, but not immediately at post-treatment. Therefore, the intervention period of this study might not have been sufficient to reveal all potential effects and a longer intervention period could possibly lead to stronger post-treatment effects in future studies.

An additional limitation of our study is that we could not guarantee that participants were kept completely blind, as information documents about the study stating that the exposure therapy intervention was the intervention of interest had to be made available online (e.g. in the study registration) and had to be disclosed to obtain informed consent.

Another important note about the current study is that a large part of data collection took place during Covid-19 restrictions, including temporary lockdowns. This might have seriously restricted the realization and effectiveness of in-vivo exposure exercises, especially exercises that required using public transport and being in crowded places. Hence, it is conceivable that the effects found in the exposure app group would have been stronger if there had been no Covid-19 restrictions and exposure exercises could have been carried out more easily.

More general, future research should attempt to systematically investigate where, when, and how participants used the exposure therapy app, in order to investigate what factors do influence the

efficacy of this type of IMI. In addition, the relationship between participant characteristic, such as attitudes to digital technology or previous treatment experiences, and the efficacy of app-guided exposure therapy should be analyzed. A further point to systematically investigate is the optimal level of guidance required for the use of IMIs. In our study, participants received only a brief introduction before beginning the unguided self-help interventions. Of note, however, a recent meta-analysis suggests that guided self-help interventions in PD/A, i.e. self-help interventions with regular therapist-contact, are more beneficial than unguided interventions [59]. Therefore, future research should also investigate how systematic variation of therapist-guidance might influence the efficacy of the here tested IMI.

4.2 Conclusions

This RCT found evidence for the efficacy of app-guided exposure therapy for PD/A. The evaluated app improved overall PD/A symptom severity as well as depressive symptoms and psychological QoL and reduced safety maneuvers. Greater improvement for overall PD/A symptoms in exposure app versus WL comparisons was found. However, we were unable to demonstrate any superiority of the exposure app compared to our active control condition, an app focused on mindfulness meditation. While descriptively, participants assigned to the exposure app showed greater and more sustained improvements than participants in the meditation app group, these differences were not statistically significant. The extent to which this was due to a lack of statistical power needs to be determined in future large-scale confirmatory studies, since the general feasibility, safety and efficacy of mobile exposure therapy for PD/A has been indicated here. Additional directions for future research include investigations aimed at the role of therapist-guidance, usage intensity, situational factors of app use, and patient characteristics. The observed response and remission rates indicate that, while some patients might greatly benefit from using the exposure app, many patients might also see no relevant change in their symptom severity. Overall, our findings suggest that app-guided exposure therapy can be a useful further treatment option in addition to established psychological and pharmacological strategies available for PD/A.

Acknowledgments

We thank Mindable Health GmbH for providing free access to the exposure app to study participants and for granting permission to feature app visualizations in this article (Figure 2). We thank 7 Mind GmbH for providing free access to the mindfulness meditation app to study participants.

Conflicts of Interest

The project was funded by budgets from the University of Bonn. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. During the project initiation phase, we consulted with Mindable Health GmbH regarding the initial study design; Mindable Health GmbH had no role in the final decision on the study design. Mindable Health GmbH had no role in study conduction, collection, analysis, or interpretation of the data; preparation, review, or approval of the manuscript, and decision to submit the manuscript for publication. In 2021 and 2022, Matthias Guth was involved in two projects with HelloBetter (GET.ON Institut für Online Gesundheitstrainings GmbH) as a scientific advisor for evaluation studies of digital interventions for chronic pain and sleep disorders. In 2021 und 2022, Matthias Guth delivered lectures sponsored by CompuGroup Medical. From 2020 to 2023, Niclas Braun received funding from BONFOR, by which Niclas Braun and Benjamin Selaskowski were partly funded. From 2020 to 2023, Alexandra Philipsen and Niclas Braun received funding from Medice, by which Annika Wiebe's position was funded. Alexandra Philipsen receives funding by the German Federal Ministry of Education and Research as part of the ITEA project DAIsy, by which Annika Wiebe's and Benjamin Selaskowski's positions is funded and Matthias Guth's position is partly funded. Over the past three years, Alexandra Philipsen received further funding by the German Federal Ministry of Education and Research, Horizon2020, Medice, and DFG; she reports serving on advisory boards for Takeda, Medice, and Boehringer; and delivering lectures sponsored by Medice, Takeda; and being the author of books and articles on psychotherapy. The other authors declare no competing interests.

Supplementary Material/Multimedia Appendix

Table S1: Completers per group and outcome

Table S2: Detailed LMM results of our ITT-analysis for primary outcomes

Table S3: LMM results of our completer analyses for primary outcomes with significant first stage F-statistic

Table S4: LMM results of our ITT-analyses for secondary outcomes with significant first stage F-statistic

CONSORT-EHEALTH checklist

Abbreviations

PD/A = panic disorders, with and without agoraphobia; IMIs = internet- or mobile-based interventions; CBT = cognitive behavioral therapy; ITT = intent-to-treat; LMM = linear mixed effects models, NNT = number needed to treat, WL = waiting list, d = Cohen's d effect size, PAS = Panic and Agoraphobia Scale, ACQ = Agoraphobic Cognitions Questionnaire, BSQ = Body Sensations Questionnaire, MI = Mobility Inventory, TSMS = Texas Safety Maneuver Scale, WHO-QOL = World Health Organization Quality of Life, DASS = Depression, Anxiety and Stress Scale, ASI = Anxiety Sensitivity Index, GSES = General Self-Efficacy Scale

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.



References

1. American Psychiatric Association. Diagnostic and statistical manual of mental disorders (DSM-5®). Arlington: American Psychiatric Publishing; 2013.
2. de Jonge P, Roest AM, Lim CCW, Florescu SE, Bromet E, Stein D, Harris M, Nakov V, Caldas-de-Almeida JM, Levinson D, Al-Hamzawi AO, Haro JM, Viana MC, Borges G, O'Neill S, de Girolamo G, Demyttenaere K, Gureje O, Iwata N, Lee S, Hu C, Karam A, Moskalewicz J, Kovess-Masfety V, Navarro-Mateu F, Browne MO, Piazza M, Posada-Villa J, Torres Y, ten Have ML, Kessler RC, Scott KM. Cross-national Epidemiology of Panic Disorder and Panic Attacks in the World Mental Health Surveys. *Depress Anxiety*. 2016 Dec;33(12):1155–1177. PMID: 27775828
3. Roest AM, de Vries YA, Lim CCW, Wittchen HU, Stein DJ, Adamowski T, Al-Hamzawi A, Bromet EJ, Viana MC, de Girolamo G, Demyttenaere K, Florescu S, Gureje O, Haro JM, Hu C, Karam EG, Caldas-de-Almeida JM, Kawakami N, Lépine JP, Levinson D, Medina-Mora ME, Navarro-Mateu F, O'Neill S, Piazza M, Posada-Villa JA, Slade T, Torres Y, Kessler RC, Scott KM, de Jonge P, Collaborators O behalf of the WWMHS. A comparison of DSM-5 and DSM-IV agoraphobia in the World Mental Health Surveys. *Depress Anxiety*. 2019;36(6):499–510. PMID: 30726581
4. National Institute for Health and Care Excellence (NICE). Clinical guideline. Generalised anxiety disorder and panic disorder in adults: management [Internet]. 2020. Available from: <https://www.nice.org.uk/guidance/cg113>
5. Andrews G, Bell C, Boyce P, Gale C, Lampe L, Marwat O, Rapee R, Wilkins G. Royal Australian and New Zealand College of Psychiatrists clinical practice guidelines for the treatment of panic disorder, social anxiety disorder and generalised anxiety disorder. *Aust N Z J Psychiatry*. 2018 Dec;52(12):1109–1172.
6. Bandelow B, Aden I, Alpers GW, Benecke C, Deckert J, Domschke K, Eckhardt-Henn A, Geiser F, Gerlach AL, Harfst T, Hau S, Hoffmann S, Hoyer J, Hunger-Schoppe C, Kellner M, Köllner V, Kopp IB, Langs G, Liebeck H, Matzat J, Ohly M, Rüddel HP, Rudolf S, Scheufele E, Simon R, Staats H, Ströhle A, Waldherr B, Wedekind D, Werner AM, Wiltink J, Wolters JP, Zwanzger P, Beutel ME. Deutsche S3-Leitlinie Behandlung von Angststörungen [German guideline: Treatment of anxiety disorders] [Internet]. 2021. Available from: <https://register.awmf.org/de/leitlinien/detail/051-028>
7. Katzman MA, Bleau P, Blier P, Chokka P, Kjernisted K, Van Ameringen M, the Canadian Anxiety Guidelines Initiative Group on behalf of the Anxiety Disorders Association of Canada/ Association Canadienne des troubles anxieux and McGill University. Canadian clinical practice guidelines for the management of anxiety, posttraumatic stress and obsessive-compulsive disorders. *BMC Psychiatry*. 2014;14(Suppl 1):S1. PMID: 25081580
8. Bandelow B. PAS. Panik- und Agoraphobie-Skala (Manual). 2nd ed. Göttingen: Hogrefe; 2016.
9. Pompoli A, Furukawa TA, Efthimiou O, Imai H, Tajika A, Salanti G. Dismantling cognitive-behaviour therapy for panic disorder: a systematic review and component network meta-analysis. *Psychol Med*. 2018 Sep;48(12):1945–1953. PMID: 29368665
10. Andersson G, Titov N. Advantages and limitations of Internet-based interventions for common

mental disorders. *World Psychiatry*. 2014 Feb;13(1):4–11. PMID: 24497236

11. Ebert DD, Van Daele T, Nordgreen T, Karekla M, Compare A, Zarbo C, Brugnera A, Øverland S, Trebbi G, Jensen KL, Kaehlke F, Baumeister H. Internet- and Mobile-Based Psychological Interventions: Applications, Efficacy, and Potential for Improving Mental Health: A Report of the EFPA E-Health Taskforce. *Eur Psychol*. 2018 May;23(2):167–187.
12. Domhardt M, Letsch J, Kybelka J, Koenigbauer J, Doebler P, Baumeister H. Are Internet- and mobile-based interventions effective in adults with diagnosed panic disorder and/or agoraphobia? A systematic review and meta-analysis. *J Affect Disord*. 2020 Nov 1;276:169–182. PMID: 32697696
13. Polak M, Tanzer NK, Bauernhofer K, Andersson G. Disorder-specific internet-based cognitive-behavioral therapy in treating panic disorder, comorbid symptoms and improving quality of life: A meta-analytic evaluation of randomized controlled trials. *Internet Interv*. 2021 Apr 1;24:100364. PMID: 33643852
14. Ivanova E, Lindner P, Ly KH, Dahlin M, Vernmark K, Andersson G, Carlbring P. 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. *J Anxiety Disord*. 2016 Dec;44:27–35. PMID: 27721123
15. Christoforou M, Sáez Fonseca JA, Tsakanikos E. Two Novel Cognitive Behavioral Therapy–Based Mobile Apps for Agoraphobia: Randomized Controlled Trial. *J Med Internet Res*. 2017 Nov 24;19(11):e398. PMID: 29175809
16. Margraf J, Cwik JC. Mini-DIPS Open Access: Diagnostisches Kurzinterview bei psychischen Störungen [Internet]. OMP Ruhr-Universität Bochum. 2017. Available from: <https://omp.ub.rub.de/index.php/RUB/catalog/book/102>
17. RED Medical Systems GmbH. RED connect [Internet]. Munich, Germany; 2020. Available from: <https://www.redmedical.de/>
18. Faul F, Erdfelder E, Buchner A, Lang AG. Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behav Res Methods*. 2009 Nov;41(4):1149–1160. PMID: 19897823
19. SosciSurvey. SosciSurvey.de. SosciSurvey GmbH; 2024.
20. European Medicines Agency. Guideline on clinical investigation of medicinal products indicated for the treatment of panic disorder [Internet]. 2005 [cited 2024 Jan 9]. Available from: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-clinical-investigation-medicinal-products-indicated-treatment-panic-disorder_en.pdf
21. Broocks A, Bandelow B, Pekrun G, George A, Meyer T, Bartmann U, Hillmer-Vogel U, Rüther E. Comparison of Aerobic Exercise, Clomipramine, and Placebo in the Treatment of Panic Disorder. *Am J Psychiatry*. 1998 May;155(5):603–609.
22. Casey LM, Oei TPS, Newcombe PA, Kenardy J. The role of catastrophic misinterpretation of bodily sensations and panic self-efficacy in predicting panic severity. *J Anxiety Disord*. 2004 Jan;18(3):325–340.

23. Ishiguro S, Watanabe T, Ueda M, Saeki Y, Hayashi Y, Akiyama K, Saito A, Kato K, Inoue Y, Shimoda K. Determinants of pharmacodynamic trajectory of the therapeutic response to paroxetine in Japanese patients with panic disorder. *Eur J Clin Pharmacol*. 2011 Dec;67(12):1213–1221.
24. Chambless DL, Caputo GC, Bright P, Gallagher R. Assessment of fear of fear in agoraphobics: The Body Sensations Questionnaire and the Agoraphobic Cognitions Questionnaire. *J Consult Clin Psychol*. 1984 Dec;52(6):1090–1097. PMID: 6520279
25. Chambless DL, Gracely EJ. Fear of fear and the anxiety disorders. *Cogn Ther Res*. 1989 Feb;13(1):9–20.
26. Ehlers A, Margraf J. AKV. Fragebogen zu körperbezogenen Ängsten, Kognitionen und Vermeidung. Manual. 2nd ed. Göttingen: Beltz Test GmbH; 2001.
27. Chambless DL, Caputo GC, Jasin SE, Gracely EJ, Williams C. The Mobility Inventory for Agoraphobia. *Behav Res Ther*. 1985;23(1):35–44. PMID: 3985915
28. Chambless DL, Sharpless BA, Rodriguez D, McCarthy KS, Milrod BL, Khalsa SR, Barber JP. Psychometric Properties of the Mobility Inventory for Agoraphobia: Convergent, Discriminant, and Criterion-Related Validity. *Behav Ther*. 2011 Dec;42(4):689–699. PMID: 22035997
29. Helbig-Lang S, Richter J, Lang T, Gerlach AL, Fehm L, Alpers GW, Ströhle A, Kircher T, Deckert J, Gloster AT, Wittchen HU. The role of safety behaviors in exposure-based treatment for panic disorder and agoraphobia: Associations to symptom severity, treatment course, and outcome. *J Anxiety Disord*. 2014 Dec;28(8):836–844. PMID: 25445073
30. Kamphuis JH, Telch MJ. Assessment of strategies to manage or avoid perceived threats among panic disorder patients: the Texas Safety Maneuver Scale (TSMS). *Clin Psychol Psychother*. 1998 Sep;5(3).
31. Lang T, Helbig-Lang S, Westphal D, Gloster AT, Wittchen HU. Expositionsbasierte Therapie der Panikstörung mit Agoraphobie: ein Behandlungsmanual. 2., aktualisierte Auflage. Göttingen: Hogrefe; 2018.
32. World Health Organization. WHOQOL User Manual [Internet]. 1998. Available from: https://iris.who.int/bitstream/handle/10665/77932/WHO_HIS_HSI_Rev.2012.03_eng.pdf
33. Angermeyer MC, Kilian R, Matschinger H. WHOQOL-BREF. Deutsche Version (German Version) [Internet]. 2000. Available from: <https://cdn.who.int/media/docs/default-source/publishing-policies/whoqol-bref/german-whoqol-bref.zip>
34. Lovibond SH, Lovibond P. Manual for the Depression Anxiety Stress Scales. 2nd ed. Sydney: Psychology Foundation; 1995.
35. Nilges P, Essau C. Die Depressions-Angst-Stress-Skalen: Der DASS – ein Screeningverfahren nicht nur für Schmerzpatienten. *Schmerz*. 2015 Dec;29(6):649–657. PMID: 26205682
36. Taylor S, Zvolensky MJ, Cox BJ, Deacon B, Heimberg RG, Ledley DR, Abramowitz JS, Holaway RM, Sandin B, Stewart SH, Coles M, Eng W, Daly ES, Arrindell WA, Bouvard M,

- Cardenas SJ. Robust dimensions of anxiety sensitivity: Development and initial validation of the Anxiety Sensitivity Index-3. *Psychol Assess*. 2007 Jun;19(2):176–188. PMID: 17563199
37. Kemper CJ, Ziegler M, Taylor S. Überprüfung der psychometrischen Qualität der deutschen Version des Angstsensitivitätsindex-3. *Diagnostica*. 2009 Oct;55(4):223–233.
38. Kemper CJ, Finnern M. Die deutsche Version des Angstsensitivitätsindex-3 (ASI-3) [Internet]. 2011. Available from: https://christoph-kemper.net/fileadmin/user_upload/Publications/Kurzbeschreibung_ASI-3_Angstsensitivitaetsindex-3_2011.pdf
39. Schwarzer R, Jerusalem M. Generalized Self-Efficacy scale. *Meas Health Psychol User's Portf Causal Control Beliefs*. Windsor, England: NFER-NELSON; 1995. p. 35–37.
40. Jerusalem M, Schwarzer R. General Self-Efficacy Scale (GSE) [Internet]. 2012. Available from: <http://userpage.fu-berlin.de/~health/selfscal.htm>
41. Schwarzer R, Jerusalem M, editors. Skalen zur Erfassung von Lehrer- und Schülermerkmalen: Dokumentation der psychometrischen Verfahren im Rahmen der Wissenschaftlichen Begleitung des Modellversuchs Selbstwirksame Schulen. Berlin: Freie Universität Berlin; 1999.
42. Mindable Health GmbH. Mindable: Panik & Agoraphobie [Internet]. Berlin, Germany; 2024. Available from: <https://www.mindable.health/>
43. BfArM. Bundesinstitut für Arzneimittel und Medizinprodukte (Federal Institute for Drugs and Medical Devices). DiGA-Verzeichnis. Mindable: Panikstörung und Agoraphobie [Internet]. 2024. Available from: <https://diga.bfarm.de/de/verzeichnis/00329>
44. 7Mind GmbH. 7Mind [Internet]. Berlin, Germany; 2024. Available from: <https://www.7mind.de/>
45. Twisk J, Bosman L, Hoekstra T, Rijnhart J, Welten M, Heymans M. Different ways to estimate treatment effects in randomised controlled trials. *Contemp Clin Trials Commun*. 2018 Jun 1;10:80–85. PMID: 29696162
46. Jacobson NS, Truax P. Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *J Consult Clin Psychol*. 1991;59(1):12–19. PMID: 2002127
47. Bender R. Calculating Confidence Intervals for the Number Needed to Treat. *Control Clin Trials*. 2001 Feb;22(2):102–110. PMID: 11306148
48. R Foundation for Statistical Computing. R: A language and environment for statistical computing [Internet]. Vienna, Austria; 2022. Available from: <https://www.R-project.org/>
49. Bates D, Mächler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models Using lme4. *J Stat Softw*. 2015;67(1).
50. Russell VL. emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.8.4-1 [Internet]. 2023. Available from: <https://CRAN.R-project.org/package=emmeans>
51. Cohen J. Statistical Power Analysis for the Behavioral Sciences [Internet]. 2nd ed. Routledge;

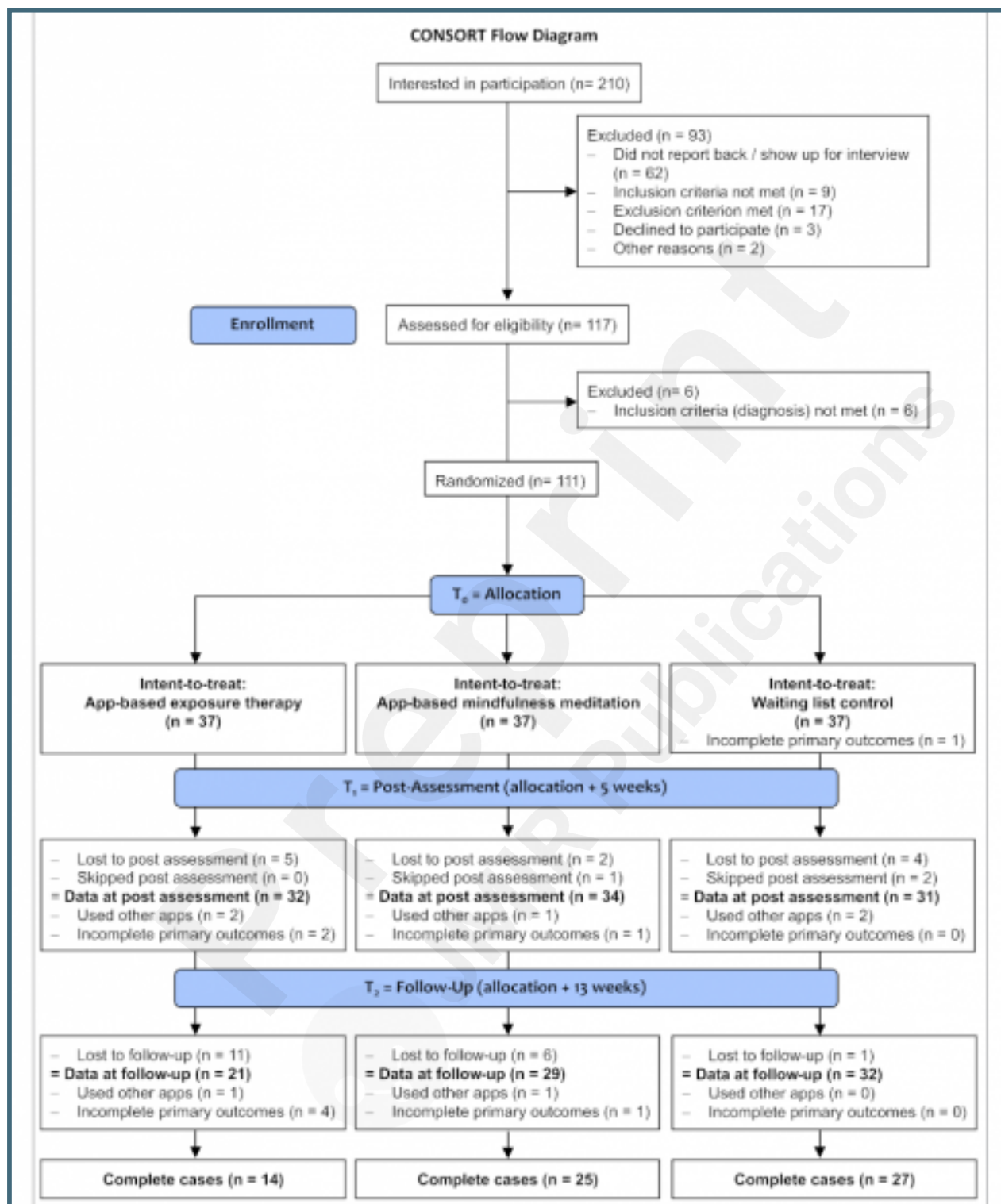
1988. Available from: <https://www.taylorfrancis.com/books/9781134742707>

52. Spies J, Lang T, Helbig-Lang S. A Smart Solution? Efficacy of a Self-Management Smartphone App for Panic Disorder and Agoraphobia. *Cogn Ther Res*. 2024 Jul 20;
53. Salkovskis PM, Clark DM, Hackmann A, Wells A, Gelder MG. An experimental investigation of the role of safety-seeking behaviours in the maintenance of panic disorder with agoraphobia. *Behav Res Ther*. 1999 Jun 1;37(6):559–574. PMID: 10372469
54. Kabat-Zinn J, Massion AO, Kristeller J, Peterson LG, Fletcher K, Pbert L, Lenderking WR, Santorelli SF. Effectiveness of a meditation-based stress reduction program in the treatment of anxiety disorders. *Am J Psychiatry*. 1992 Jul;149(7):936–943. PMID: 1609875
55. Saeed SA, Cunningham K, Bloch RM. Depression and Anxiety Disorders: Benefits of Exercise, Yoga, and Meditation. *Am Fam Physician*. 2019 May 15;99(10):620–627. PMID: 31083878
56. Oliveira SEH, Carvalho H, Esteves F. Toward an understanding of the quality of life construct: Validity and reliability of the WHOQOL-Bref in a psychiatric sample. *Psychiatry Res*. 2016 Oct 30;244:37–44. PMID: 27455149
57. Smits JAJ, Berry AC, Tart CD, Powers MB. The efficacy of cognitive-behavioral interventions for reducing anxiety sensitivity: A meta-analytic review. *Behav Res Ther*. 2008 Sep 1;46(9):1047–1054. PMID: 18687421
58. Gallagher MW, Payne LA, White KS, Shear KM, Woods SW, Gorman JM, Barlow DH. Mechanisms of change in cognitive behavioral therapy for panic disorder: The unique effects of self-efficacy and anxiety sensitivity. *Behav Res Ther*. 2013 Nov;51(11):767–777. PMID: 24095901
59. Papola D, Ostuzzi G, Tedeschi F, Gastaldon C, Purgato M, Giovane CD, Pompoli A, Pauley D, Karyotaki E, Sijbrandij M, Furukawa TA, Cuijpers P, Barbui C. CBT treatment delivery formats for panic disorder: a systematic review and network meta-analysis of randomised controlled trials. *Psychol Med*. Cambridge University Press; 2023 Feb;53(3):614–624. PMID: 37132646

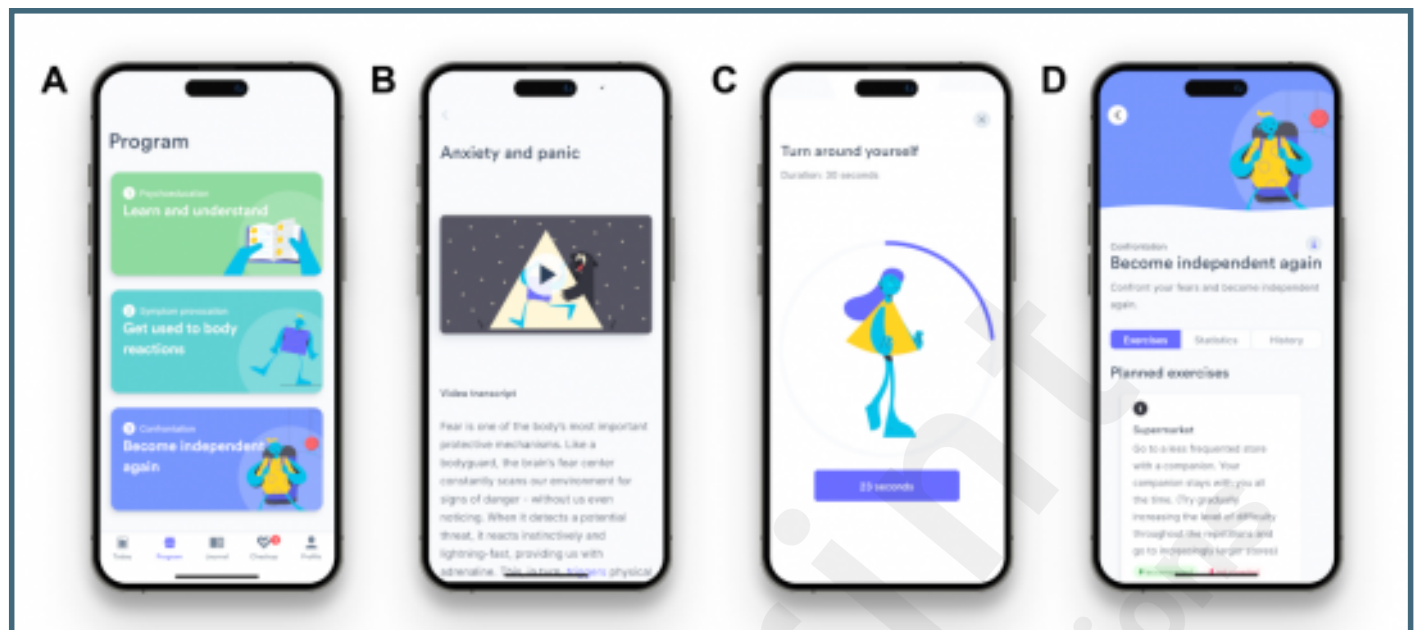
Supplementary Files

Figures

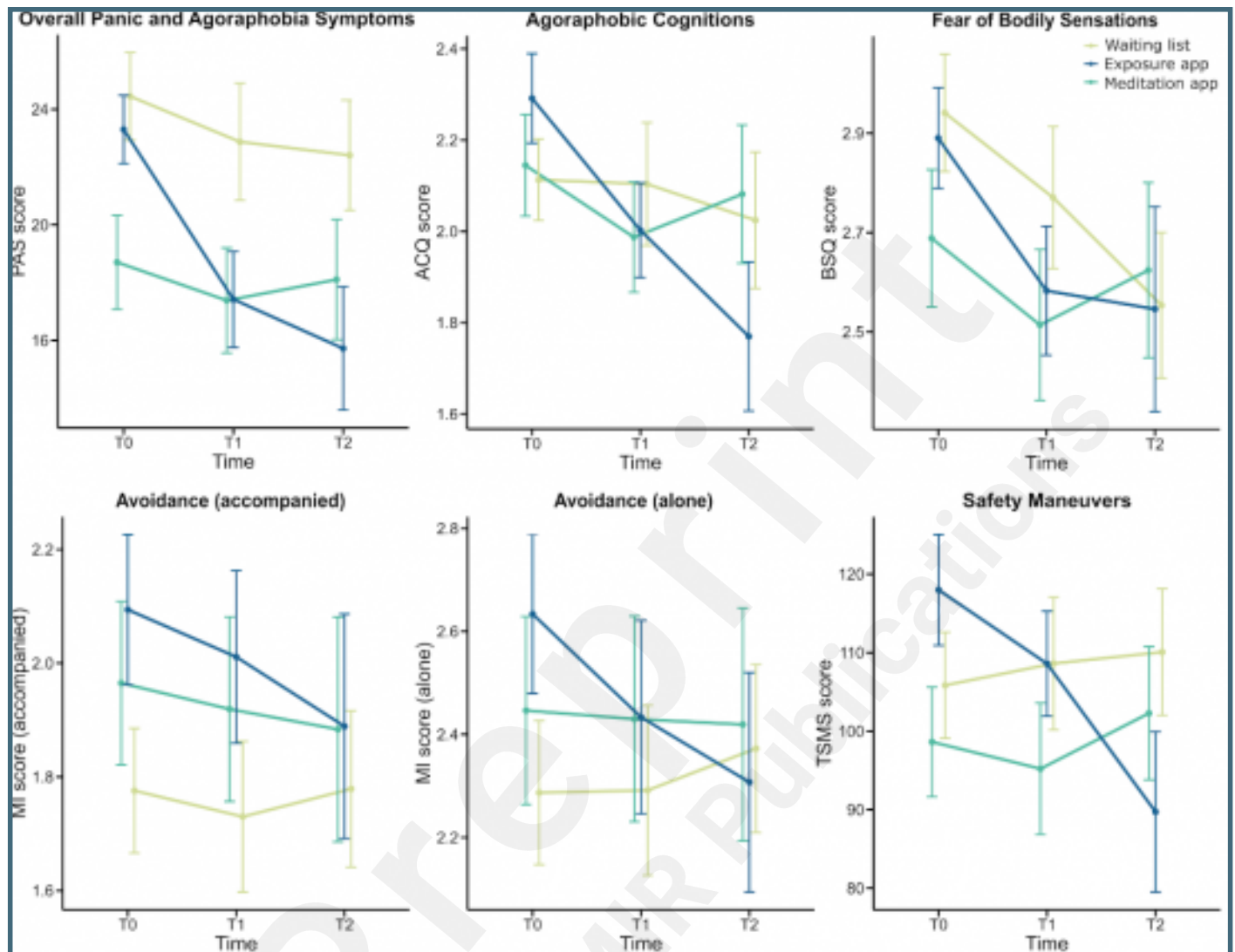
CONSORT flow diagram.



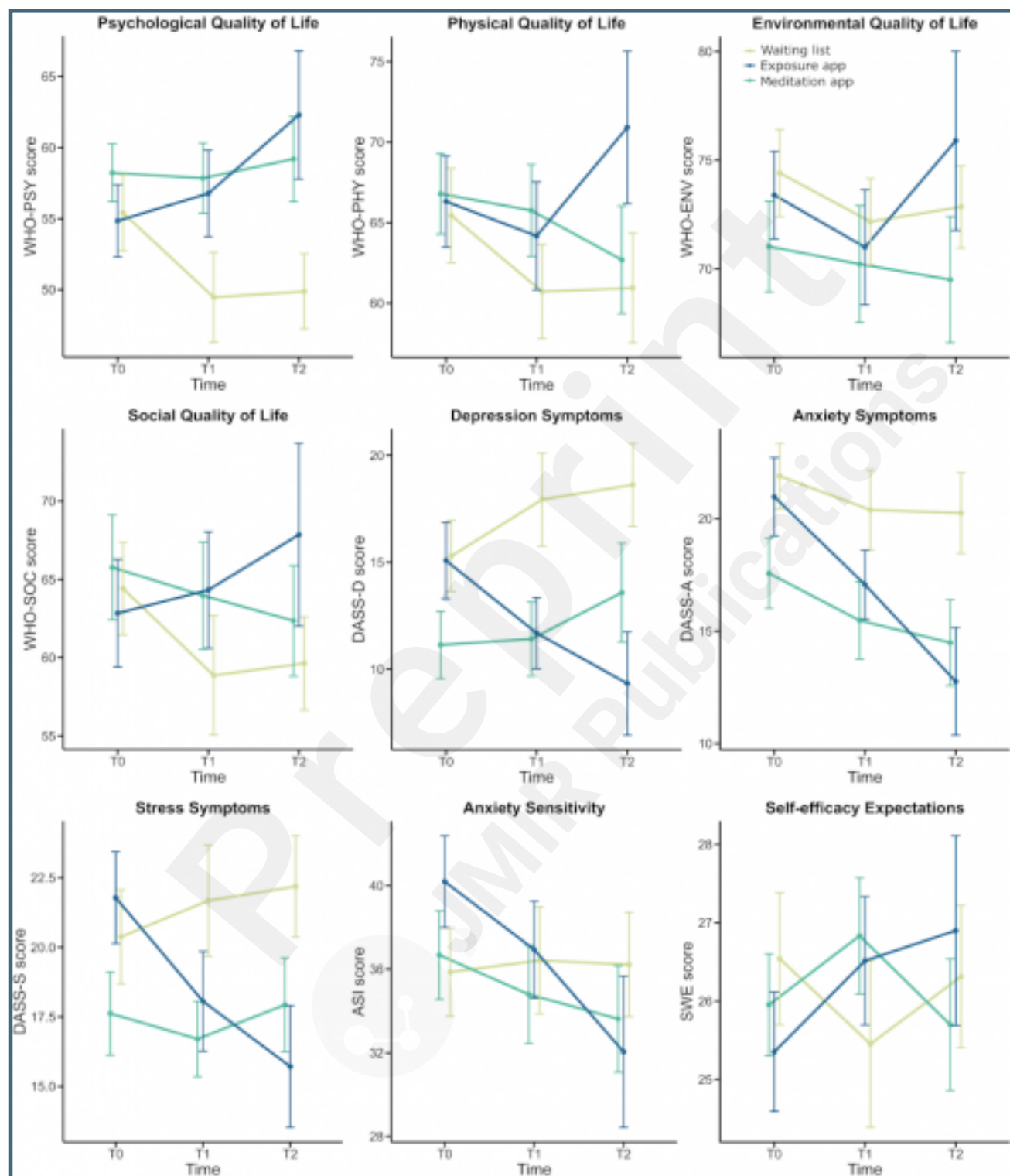
Visualization of the exposure therapy app. Visualization copyright: Mindable Health GmbH.



Changes in primary outcome means in each group across the three timepoints. Error bars represent the standard error of the mean.



Changes in secondary outcome means in each group across the three timepoints. Error bars represent the standard error of the mean.



Multimedia Appendixes

Table S1: Completers per group and outcome.

URL: <http://asset.jmir.pub/assets/8fec99625dd6108303f4bfcd2e8e3398.pdf>

Table S2: Detailed LMM results of our ITT-analysis for primary outcomes.

URL: <http://asset.jmir.pub/assets/e531a62f83630f34cdc1015e5a7019ac.pdf>

Table S3: LMM results of our completer analyses for primary outcomes with significant first stage F-statistic.

URL: <http://asset.jmir.pub/assets/4f72b0c8adf287abfd3e0e0ff81b580c.pdf>

Table S4: LMM results of our ITT-analyses for secondary outcomes with significant first stage F-statistic.

URL: <http://asset.jmir.pub/assets/8185b1eb1058862baca027f4446ee54c.pdf>



CONSORT (or other) checklists

CONSORT-EHEALTH checklist.

URL: <http://asset.jmir.pub/assets/5c0612b5e4bb0f5b25005cf859c46b35.pdf>

TOC/Feature image for homepages

AI-generated image, in response to the request "a photo of a person's arm and hand seen from the side holding a mobile phone with a crowded public space in the background" (Generator: Stable Diffusion Text to Image; Date: April 24, 2025; Requestor: Matthias Guth).

