Supplementary Online Content

Donker T, Cornelisz I, van Klaveren C, et al. Effectiveness of self-guided app-based virtual reality cognitive behavior therapy for acrophobia: a randomized clinical trial. *JAMA Psychiatry*. Published online March 20, 2019. doi:10.1001/jamapsychiatry.2019.0219

This supplementary material has been provided by the authors to give readers additional information about their work.

eAppendix 1. Literature search for app-based acrophobia treatment

We searched MEDLINE with the terms (mobile app OR app-based) AND (acrophobi* OR fear of height*) for articles published before November 22, 2018 in any language. We retrieved one paper, which was our protocol.

eAppendix 2. Deviations from the original trial protocol submitted to the ethic committee and the research protocol

Deviations from the original trial protocol submitted to the ethic committee

1. In the research proposal submitted to the ethic committee, we estimated a sample size of N=154 based on an expected AQ post-treatment effect size of 0.5, power of 0.80 and an anticipated drop-out of 20%, which is typically observed in online interventions and virtual reality interventions. However, these interventions were al guided by a coach or therapist. In the current study, there is no guidance from a therapist or coach. Prior to the start of trial, we expected a higher drop-out rate because this is more commonly observed in unguided interventions [1]. Therefore, we estimated drop-out to be around 40%, thereby requiring a sample of N=180.

2. Contrary to the research proposal for the ethic committee, we did not include the GAD-7 (measuring general anxiety) as secondary outcome measure, because we wanted to lower the number of questions to be filled in by participants, and we already had an general anxiety measure included.

3. We have not used ANOVA tests with missing observations carried forward as was described in the ethics proposal, because more valid analytical techniques are used that acknowledge the high drop-out that was observed due to ineligible smartphones of participants. The experiment suffered from two complications (non-compliance and missing outcomes), which is why intention-to-treat analyses (ITT) are performed and reported [cf. 2]. We have added Lee bounds analysis and Multiple Imputation techniques to account for potential concerns of bias and precision of the single regression-based imputations reported in the paper [cf. 3].

4. Exploratory analyses (clinical meaningful change, reliable change and number needed to treat) were added to the current paper to examine robustness of results.

5. The intervention period lasted three weeks instead of six weeks as was originally planned. This has been adjusted following findings from user-tests, in which most of test-participants finished the modules in three weeks. 6. The anxiety subscale was used to calculate the sample size. This was mistakenly taken over as a primary outcome in the registered protocol. In line with previous studies reporting on acrophobia questionnaires, we used the complete AQ as the primary outcome measure. We do, also in line with previous studies, report on both the anxiety subscale and avoidance subscale as well. The suicidal ideation question was only assessed at baseline to exclude suicidal participants.

Deviations from the protocol study

7) In the protocol paper [1], we mentioned that we would examine two types of per protocol analyses: 1) participants who returned post-test and/or follow-up questionnaires, and 2) participants who also completed 50% or more of the ZeroPhobia modules. In the current paper we only report the first category of per protocol (sensitivity) analyses (participants who returned post-test and/or follow-up questionnaires). For consideration of space, the per protocol analyses of participants who also completed 50% or more of the modules will be reported in another paper focusing on usage data. In addition, ecological momentary assessment (EMA) and the variable "practice in real life" will be reported in this paper as well. With studying EMA, we will be able to examine - when participants start practising with exposure, and if this is different compared to traditional face-to-face therapy.

8) In the protocol paper [1], we mentioned that usage data (EMA) would be collected to model missing data. However we have not included usage data in our analysis yet. Different techniques to analyse missing data were used for the pre-treatment attrition sample and the dropout sample. Missing outcome values for the pre-treatment attrition sample were imputed using the waitlist control sample, whereas for the dropout sample initial treatment assignment was used.

9) The 23% attrition in the intervention group was caused by ineligible smartphones. After post-test, participants in the waitlist condition who had an ineligible smartphone were asked to borrow a smartphone

from someone else or, in case they could not access an eligible phone, and for ethical reasons, were offered the intervention at the VU University. The same procedure applied for the 23% of participants who could not participate in the intervention group. Because their data would not be used in the analysis, this would not cause any harms to data integrity.

[1] Donker T, Bennett K, Bennett A, et al. Internet-Delivered Interpersonal Psychotherapy Versus Internet-Delivered Cognitive Behavioral Therapy for Adults With Depressive Symptoms: Randomized Controlled Noninferiority Trial. Eysenbach G, ed. *Journal of Medical Internet Research*. 2013;15(5):e82. doi:10.2196/jmir.2307.

[2] Gupta SK Intention-to-treat concept: a review. *Perspectives in clinical research*, 2011; *2*(3): 109.

[3] Little RJ, D'agostino R, Cohen ML, et al. The prevention and treatment of missing data in clinical trials. *New England Journal of Medicine*, 2012;367(14): 1355-1360.

eTable 1. Comparison of missing outcome participants to non-missing outcome participants after randomisation

Measure	Non-missing (n=129)	Missing overall (n=64)	P Value	Missing Tech issues (n=21)	P Value	Missing Unknown reasons (n=43)	P Value
Demographic variables							
Age (M, SD)	41.15 (13.30)	41.69 (14.41)	.80	45.91 (12.21)	.10	39.63 (15.07)	.36
Female, no. (%)	89 (69)	40 (62.5)		12 (57.14)	.32	28 (65.11)	.79
Education, no. (%)							
None or primary	2 (1.55)	0 (0)	.32	0 (0)	.62	0 (0)	.45
Secondary	11 (8.53)	7 (10.94)	.59	3 (14.29)	.41	4 (9.30)	.99
Post-secondary	116 (89.92)	57 (89.06)	.86	18 (85.71)	.54	39 (90.70)	.80
Psychotropic medication, no. (%)	4 (3.1)	0 (0)	.76	0 (0)	.88	0 (0)	.82
Primary outcome-baseline (M, SD)							
AQ tot	84.73 (18.79)	84.53 (2.18)	.94	85.24 (19.22)	.88	84.19 (16.68)	.85
AQ Anxiety	68.52 (14.74)	67.73 (13.94)	.72	67.86 (14.76)	.89	67.67 (13.70)	.76
AQ –Avoid	16.21 (5.11)	16.80 (4.81)	.44	17.38 (5.95)	.35	16.51 (4.20)	.87
Secondary outcome-baseline							
ATHQ	45.08 (9.01)	43.87 (9.51)	.39	45.99 (7.76)	.49	42.83 (10.18)	.14
BAI	33.32 (18.19)	31.98 (18.69)	.64	26.71 (20.27)	.10	34.56 (17.54)	.50
Mastery	26.52 (6.41)	25.30 (7.40)	.24	26.62 (4.91)	.72	24.65 (8.33)	.11
PHQ	2.19 (3.19)	2.41 (3.09)	.65	2.29 (2.67)	.97	2.47 (3.31)	.63

Abbreviations: AQ, Acrophobia Questionnaire; ATHQ, Attention to Height Questionnaire; BAI, Beck Anxiety Inventory; PHQ, Patient Health Questionnaire.

eTable 2. Inferential statistics of treatment outcome measures, complete cases (n=134-144)

	ZeroPhobia (n: [51-57])			**Wait list control group (n: [83-87])**		**P Value[b]**	**Effect size (Cohen`s d) (95% CI)**	
Primary outcome (M, SD)	Baseline	Post-test	Follow-up	Baseline	Post-test		Between, post-test	Within, Bas-FU [c]
AQ – total	84.88 (19.28)	39.43 (24.36)	32.20 (19.11)	84.79 (18.40)	75.02 (22.47)	<.001	1.53 (1.15-1.91)	2.68 (2.09-3.22)
AQ-Anxiety	68.25 (15.36)	32.81 (19.22)	26.76 (14.74)	68.83 (14.64)	60.99 (17.72)	<.001	1.54 (1.16-1.91)	2.74 (2.15-3.28)
AQ-Avoidance	16.21 (5.28)	6.57 (5.79)	5.64 (5.06)	16.24 (4.92)	14.16 (5.35)	<.001	1.37 (1.00-1.74)	2 (1.48-2.48)
Secondary outcomes (M,SD)								
ATHQ	44.93 (8.45)	28.67 (13.03)	25.28 (12.45)	44.47 (9.64)	45.53 (9.92)	<.001	1.50 (1.12-1.88)	1.75 (1.25-2.22)
BAI	30.45 (18.16)	31.35 (10.30)	25.45 (5.64)	34.77 (17.66)	38.13 (14.87)	.005	0.51 (0.15-0.86)	0.33 (0.11-0.76)
Mastery[a]	26.00 (6.74)	28.52 (4.42)	28.89 (4.93)	26.72 (5.97)	27.57 (5.11)	.26	-0.20 (-0.54-0.15)	-0.42 (-.84—.001)
PHQ	2.04 (3.38)	1.98 (2.48)	1.89 (2.56)	2.20 (3.01)	2.87 (3.64)	.12	0.27 (-0.07-0.62)	0.06 (-0.36-0.48)

Abbreviations: AQ, Acrophobia Questionnaire; ATHQ, Attention to Height Questionnaire; BAI, Beck Anxiety Inventory; Bas, Baseline; FU, Follow-up; PHQ, Patient Health Questionnaire.

[a]A higher score means a higher sense of mastery.

[b] T-test comparisons pre (baseline) and post-test outcomes for complete cases.

[c] Within-group effect sizes are based on subjects who filled in both baseline and follow-up.

eTable 3. Heterogeneity of treatment effects of ZeroPhobia with respect to baseline acrophobia symptoms (n=193)

Variable	**AQ total post**					
	Model 1			**Model 2**		
	b ITT (SE)	t	*P* Value[a]	b ITT (SE)	t	*P* Value[a]
Intervention	-26.7 (2.73)	-9.78	<.0001	-0.880 (13.00)	-0.07	0.95
AQ total baseline	0.664 (0.082)	8.07	<.0001	0.823 (0.113)	7.28	<.0001
Intervention x AQ total baseline				-0.306 (0.151)	-2.02	0.043
R^2	0.52			0.54		

Abbreviations: AQ, Acrophobia Questionnaire; SE, standard error,

[a]two-sided. The outcome variable in Models 1 and 2 is AQ-total post-score. Treatment effects reported in Model 1 and 2 are intention-to-treat (ITT) effects, using dual-strategy regression imputation based on the source of missing data. Pre-score and background controls are included for improved precision of the regression point estimates reported. Background control variables are gender, age, education level and a dummy for whether a participant had missing values on one or more background characteristics.

eTable 4. Estimation results when covarying for potential general anxiety symptoms (n=193)

Measure	**Post-test**			
	b ITT (SE)	95% CI	t	*P* Value[a]
Primary outcome				
AQ- total	-23.89 (2.58)	-28.98 - -18.79)	-9.24	<.0001
General anxiety symptoms	.656 (0.12)	0.42-0.89	5.55	
Adj. R	0.57			

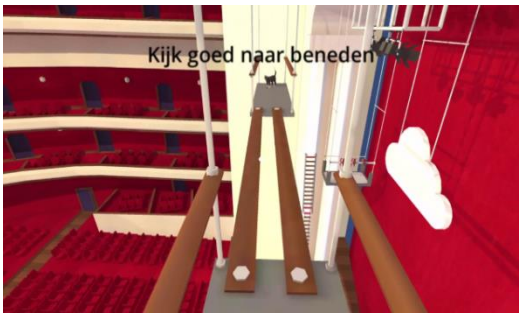
Abbreviations: AQ, Acrophobia Questionnaire; ITT: intention-to-treat; SE, standard error.
[a]two-sided. The outcome variable is AQ-total post-score. Treatment effects are intention-to-treat (ITT) effects, using dual-strategy regression imputation based on the source of missing data. Pre-score and background controls are included for improved precision of the regression point estimates reported. Background control variables are gender, age, education level, general anxiety symptoms and a dummy for whether a participant had missing values on one or more background characteristics.

eTable 5. The regression-predicted reductions in post-test for 10-point bins of pre-scores

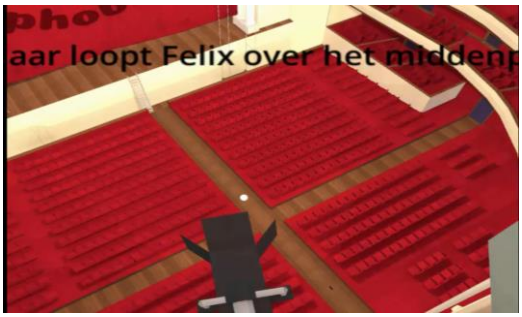
AQ Total (baseline)	Mean
50	-31.02
60	-35.53
70	-40.44
80	-45.22
90	-50.24
100	-53.89
110	-61.19
120	-66.36
130	-68.71

Abbreviations: AQ, Acrophobia Questionnaire

eFigure 1. Virtual reality screen shots



Virtual Reality Theatre



Virtual reality theatre