A Mobile App for Social Anxiety Disorder: A Three-Arm Randomized Controlled Trial Comparing Mobile and PC-Based Guided Self-Help Interventions

Article *in* Journal of Consulting and Clinical Psychology · June 2018

DOI: 10.1037/ccp0000301

CITATIONS
78

READS
4,253

8 authors, including:

Timo Stolz
Universität Bern
9 PUBLICATIONS **279** CITATIONS
SEE PROFILE

Ava Schulz
University of Zurich
23 PUBLICATIONS **577** CITATIONS
SEE PROFILE

Tobias Krieger
Universität Bern
115 PUBLICATIONS **2,849** CITATIONS
SEE PROFILE

Alessia Vincent
Universitätsspital Basel
26 PUBLICATIONS **600** CITATIONS
SEE PROFILE

American Psychological Association

Journal of Consulting and Clinical Psychology

0022-006X/18/$12.00

2018, Vol. 86, No. 6, 493–504
http://dx.doi.org/10.1037/ccp0000301

A Mobile App for Social Anxiety Disorder: A Three-Arm Randomized Controlled Trial Comparing Mobile and PC-Based Guided Self-Help Interventions

Timo Stolz, Ava Schulz, Tobias Krieger, Alessia Vincent, Antoine Urech, Christian Moser, Stefan Westermann, and Thomas Berger
University of Bern

Objective: Internet-based cognitive–behavioral treatments (ICBT) have shown promise for various mental disorders, including social anxiety disorder (SAD). Most of these treatments have been delivered on desktop computers. However, the use of smartphones is becoming ubiquitous and could extend the reach of ICBT into users' everyday life. Only a few studies have empirically examined the efficacy of ICBT delivered through a smartphone app and there is no published study on mobile app delivered ICBT for SAD. This three-arm randomized-controlled trial (RCT) is the first to compare the efficacy of guided ICBT for smartphones (app) and conventional computers (PC) with a wait list control group (WL). ***Method:*** A total of 150 individuals meeting the diagnostic criteria for SAD were randomly assigned to one of the three conditions. Primary endpoints were self-report measures and diagnostic status of SAD. ***Results:*** After 12 weeks of treatment, both active conditions showed superior outcome on the composite of all SAD measures (PC vs. WL: $d = 0.74$; App vs. WL: $d = 0.89$) and promising diagnostic response rates ($NNT_{PC} = 3.33$; $NNT_{App} = 6.00$) compared to the WL. No significant between-groups effects were found between the two active conditions on the composite score (Cohen's $d = 0.07$). Treatment gains were maintained at 3-month follow-up. Program use was more evenly spread throughout the day in the mobile condition, indicating an integration of the program into daily routines. ***Conclusions:*** ICBT can be delivered effectively using smartphones.

What is the public health significance of this article?
This study demonstrates that internet-based cognitive–behavioral treatments can be implemented by mobile devices, extending the reach into daily routines and fostering dissemination beyond the prevalence of desktop computers, for example in developing countries.

Keywords: social anxiety disorder, cognitive–behavioral treatment, internet intervention, smartphone app, mobile treatment

Supplemental materials: http://dx.doi.org/10.1037/ccp0000301.supp

Internet-based cognitive–behavioral treatments (ICBT) have shown clinical efficacy and effectiveness for a variety of mental disorders in a number of randomized controlled trials (RCTs; Andersson, Cuijpers, Carlbring, Riper, & Hedman, 2014; Hedman, Ljótsson, & Lindefors, 2012). Most of the promising evidence comes from studies evaluating guided ICBT, a treatment format in which patients are supported by therapists or coaches while they work their way through a structured self-help program (Andersson, 2016).

One of the most researched disorders in ICBT is social anxiety disorder (SAD; Hedman, Botella, & Berger, 2016). A recent review identified 21 studies investigating Internet-based treatments for SAD comprising a total of 1,801 individuals suffering from SAD (Boettcher, Carlbring, Renneberg, & Berger, 2013). Overall, the vast majority of the RCTs investigating ICBT for SAD reported substantial reductions of social anxiety symptoms with large within-group effect sizes (Cohen's $d > 0.80$; Boettcher et al., 2013). Trials directly comparing ICBT for SAD with face-to-face therapy revealed that both conditions led to large and similar clinical improvements (Andrews, Davies, & Titov, 2011; Botella et al., 2010; Hedman, Andersson, et al., 2011). Furthermore, treatment gains were stable for up to 5 years after treatment termination and the available data also shows that ICBT can be a cost-effective alternative to traditional CBT for SAD (Hedman,

Timo Stolz, Ava Schulz, Tobias Krieger, Alessia Vincent, Antoine Urech, Christian Moser, Stefan Westermann, and Thomas Berger, Department of Clinical Psychology and Psychotherapy, University of Bern.

This research was supported by the Swiss National Science Foundation Grant PP00P1_144824/1 awarded to Thomas Berger.

Correspondence concerning this article should be addressed to Timo Stolz, Department of Clinical Psychology and Psychotherapy, University of Bern, Fabrikstrasse 8, 3012 Bern, Switzerland. E-mail: timo.stolz@psy.unibe.ch

Andersson, et al., 2011; Hedman, Furmark, et al., 2011). One study also examined negative effects of ICBT for SAD (Boettcher, Rozental, Andersson, & Carlbring, 2014). In this study, 14% of the participants described unwanted negative events that they related to the treatment. The emergence of new symptoms was the most commonly experienced negative effect, followed by the aggravation of social anxiety symptoms and negative well-being.

SAD is characterized by a marked and persistent fear of negative evaluation in social situations (American Psychiatric Association, 2013) and is one of the most common mental disorders with 12-month prevalence estimates of 7–8% in large-scale community-based studies (Kessler et al., 1994; Magee, Eaton, Wittchen, McGonagle, & Kessler, 1996). ICBT may be especially suitable for this patient group because the absence of face-to-face contact might reduce fear and thereby improve learning of psychoeducational contents of the treatment (Hedman et al., 2016). Furthermore, the majority of individuals suffering from SAD do not seek treatment or if they do, only after a long delay of up to several years (Wang et al., 2005). Olfson et al. (2000) found that a considerable number of individuals with SAD avoid seeking treatment because of their fear of what others (including the therapist) might think of them. ICBT may lower the threshold to seek help because socially anxious individuals experience more control and less threat of negative evaluation in online interactions (Lee & Stapinski, 2012).

So far, most evaluated Internet-based treatments have been delivered via desktop computers. However, the increased usage of smartphones presents promising opportunities to extend the reach of ICBT into users' everyday life, as these devices tend to be always on and within an arm's reach (Boschen & Casey, 2008; Ly et al., 2015). Mobile treatments might also have further advantages over more traditional ICBT because they allow users to perform and record exercises in their natural environment (Bang et al., 2007). Furthermore, the exercise instructions and supportive materials can be reviewed on demand in everyday situations (Watts et al., 2013). Moreover, assessment can occur right before and after critical events (Heron & Smyth, 2010) and thus may be less biased by, for example, postevent processing (i.e., the tendency for individuals with SAD to engage in a detailed review of events following a social interaction). This might lead to more valid diary entries and thereby to more specific feedback and a more appropriate sense of increasing progress later on (Mellings & Alden, 2000). More generally, the use of ICBT on a smartphone might lead to shorter but more frequent sessions and therefore facilitate learning (Bjork, Dunlosky, & Kornell, 2013). Lastly, a smartphone can be used more privately than a desktop computer that might be used by multiple people (Yuen, Goetter, Herbert, & Forman, 2012).

On the other hand, mobile treatments may also have specific drawbacks. For example, the small screen might cause problems with longer texts and the lack of a keyboard might reduce comfort. Consequently, contents might be skipped over and diary entries or exercise records might be shorter than on a computer with larger screens and keyboards, all lowering the depth of processing and thereby probably lowering outcome.

Up to now, only few studies have evaluated smartphone-delivered treatments. In their systematic review, Donker et al. (2013) could only identify a handful of smartphone applications (apps) that have been empirically and systematically evaluated (with a total $N = 227$). This low number is surprising given the fact that about 18% of all 13,600 apps that had been published in 2013 targeted health behavior (Donker et al., 2013). To date, studies on mobile interventions target various health-related areas, such as depression (Burns et al., 2011; Ly et al., 2015; Ly, Trüschel, et al., 2014; Watts et al., 2013), SAD (Dagöö et al., 2014), stress reduction (Ly, Asplund, & Andersson, 2014), fitness (Fanning, Mullen, & McAuley, 2012), weight loss (Carter, Burley, Nykjaer, & Cade, 2013), smoking cessation (Whittaker, McRobbie, Bullen, Rodgers, & Gu, 2016), and borderline personality disorder with substance abuse (Rizvi, Dimeff, Skutch, Carroll, & Linehan, 2011). In an earlier study targeting SAD (Dagöö et al., 2014), interventions could be accessed through smartphones, tablet computers, and standard computers. Thus, and in contrast to the current study, Dagöö et al. (2014) investigated an intervention that was optimized for the use on various devices and did not evaluate an app that only could be installed and used on smartphones. In their rather small trial with a total of 52 participants, Dagöö et al. (2014) reported that participants used the smartphone 42.8%, the computer 50.1%, and the tablet computer 7.1% of the time. This study gave preliminary evidence that an intervention for SAD patients could be provided via smartphones. However, this study compared Internet-based cognitive behavior therapy with Internet-based interpersonal psychotherapy and did not include a comparison of smartphone versus computer-delivered treatment, therefore, more research is needed regarding smartphone-delivered treatments for SAD.

Overall, results from the abovementioned studies indicate that smartphone-delivered treatments might be feasible, acceptable, efficacious, and suited for CBT-based treatments. However, the available studies are rather small or uncontrolled, and larger controlled studies are needed. Moreover, some of the studies evaluated blended treatments in which smartphone apps and face-to-face therapies or regular ICBT delivered on desktop computers were combined (Dagöö et al., 2014; Ly et al., 2015; Proudfoot et al., 2013). Only two pilot studies compared mobile versus PC-based treatment delivery (Carter et al., 2013; Watts et al., 2013) and found no significant differences in efficacy. However, in these cases, this could be explained by limitations of sample size and power.

Objective and Study Aim

To evaluate whether a previously validated Internet-based self-help program for SAD remains efficacious when delivered via smartphone app, we conducted a RCT in which we compared guided ICBT disseminated by means of a personal computer (PC) or an app with a wait list control group (WL). We hypothesized that both active treatment conditions would be superior to the WL. With regard to the active conditions, we did not have a clear hypothesis. There is neither evidence nor theory that could give conclusive hints to how assets and drawbacks of a mobile treatment might ultimately balance out. We also analyzed negative effects, usage patterns and predictors of outcome in the two active treatment conditions. To our knowledge, the new delivery option has never been systematically compared to the already well-established PC-based distribution in a large RCT.

Method

Recruitment and Participant Characteristics

Participants were recruited from the general population through newspaper articles, mental health related magazines, online forums, and interviews on radio and TV. In addition, several participants found our recruitment page via different search engines or links from other sites. The study website provided general information on SAD and its treatment, an outline of the study, a 24-hr emergency phone number and a registration form. Further study information and an informed consent form were provided via e-mail after registration.

Criteria for inclusion were (a) a minimum age of 18 years, (b) access to a computer and a smartphone both with Internet connection, (b) sufficient command of the German language, (c) exceeding the cut-off score on the Social Interaction Anxiety Scale (SIAS) or Social Phobia Scale (SPS), and (d) a primary diagnosis of SAD according to *DSM–IV*. Criteria for exclusion were (a) a history of psychotic or bipolar disorders, (b) other ongoing psychological treatment, (c) prescribed medication for anxiety or depression if the dosage had been changed during the last month prior to participation, or (d) active suicidal plans. Candidates with active suicidal plans were referred to a local psychiatrist or psychotherapist. Excluded persons were given access to the materials outside of the study if they were in a stable condition and were referred to other treatment options when needed.

Enrollment and Sampling Procedures

After returning a signed copy of the informed consent form, participants were asked to fill in the web-based baseline questionnaires, including SAD measures, questions concerning demographic variables, current medication, and prior or ongoing psychological treatment. When exceeding predefined cut-offs on one of the two social anxiety measures (SPS $>$ 22 or SIAS $>$ 33; Mattick & Clarke, 1998; German version: Stangier, Heidenreich, Berardi, Golbs, & Hoyer, 1999), participants were interviewed by phone using the Structured Clinical Interview for *DSM–IV*—Axis I disorders (SCID; First, Spitzer, Gibbon, & Williams, 2002; German version by Wittchen, Zaudig, & Fydrich, 1997). The interviews were conducted by the authors and 10 advanced master students. All of them were previously trained in a workshop and further educated and tested in personal trial interviews with detailed feedback. In addition, interviewers were supervised by Timo Stolz and Thomas Berger. Interviewers could not be kept blind regarding group assignment at postassessment because some participants disclosed aspects of the group assignment during the interview.

A total of 853 individuals signed up on the study website. Two hundred fifteen individuals signed the informed consent and completed the questionnaires. Out of those, 65 did not pass inclusion and exclusion criteria, as depicted in Figure 1. The remaining 150 participants were then randomly assigned to one of the three conditions: app, PC, or WL. The digital allocation tickets were concealed from the investigators and were assigned using a computerized random number generator. After randomization, the participants received an e-mail regarding their allocation.

Study Design

This RCT compared two active groups with either mobile or PC-based self-help to a WL. The WL received access to the mobile version after the treatment groups had completed the program. In the PC-based condition, access to the intervention was denied for mobile devices; participants in the mobile condition could only use the app on smartphones (iPhone or Android) to prevent shifts between the active conditions. Both treatments were not publicly available, as their use required participation in the study. All participants were guided by individual coaches during their active period. All groups were followed up until 3 months after treatment completion, as depicted in the flowchart (see Figure 1).

The trial was registered with controlled-trials.com (ISRCTN 10627379) and was approved by the Ethics Committee of the Canton of Bern, Switzerland (13/05/2014, ref: 063/14).

Clinical Intervention

Participants of both active groups received access to the same self-help materials. The intervention is based on the cognitive model of Clark and Wells (1995) and informed by the treatment manual of Stangier, Heidenreich, and Peitz (2009). The same materials have already been tested and proven efficacious in earlier RCTs (Berger, Boettcher, & Caspar, 2014; Berger et al., 2011; Berger, Hohl, & Caspar, 2009, 2010; Boettcher, Berger, & Renneberg, 2012; Schulz et al., 2016). The layout was adjusted for the smaller screen sizes of smartphones. The treatment targets both behavioral and cognitive maintaining factors, such as self-focused attention, negative automatic thoughts, experiential avoidance, processing of self as an object under social evaluation, ruminative postevent processing of social situations, safety behaviors and biased perception of somatic and cognitive anxiety symptoms.

Contents. The self-help program and smartphone app were structured into eight modules:

1. Motivational enhancement: Participants are engaged in thinking of reasons to initiate change; they define goals and start to record difficult situations.
2. Psychoeducation: This module delivers information on SAD with a focus on maintaining processes.
3. Cognitive restructuring: This module includes a thought diary to track negative beliefs in daily routine, alongside with exercises to formulate helpful thoughts.
4. Self-focused attention: Various exercises to intentionally direct attention away from oneself.
5. Behavioral experiments: A training module for planning and tracking in vivo exposures. This is further detailed in the supplementary online materials, Supplemental Figure S5 and S6.
6. Summary and repetition: The treatment's key elements in brief with an emphasis on the importance of repeated practice.
7. Healthy lifestyle and problem solving: General information on health improving behaviors.

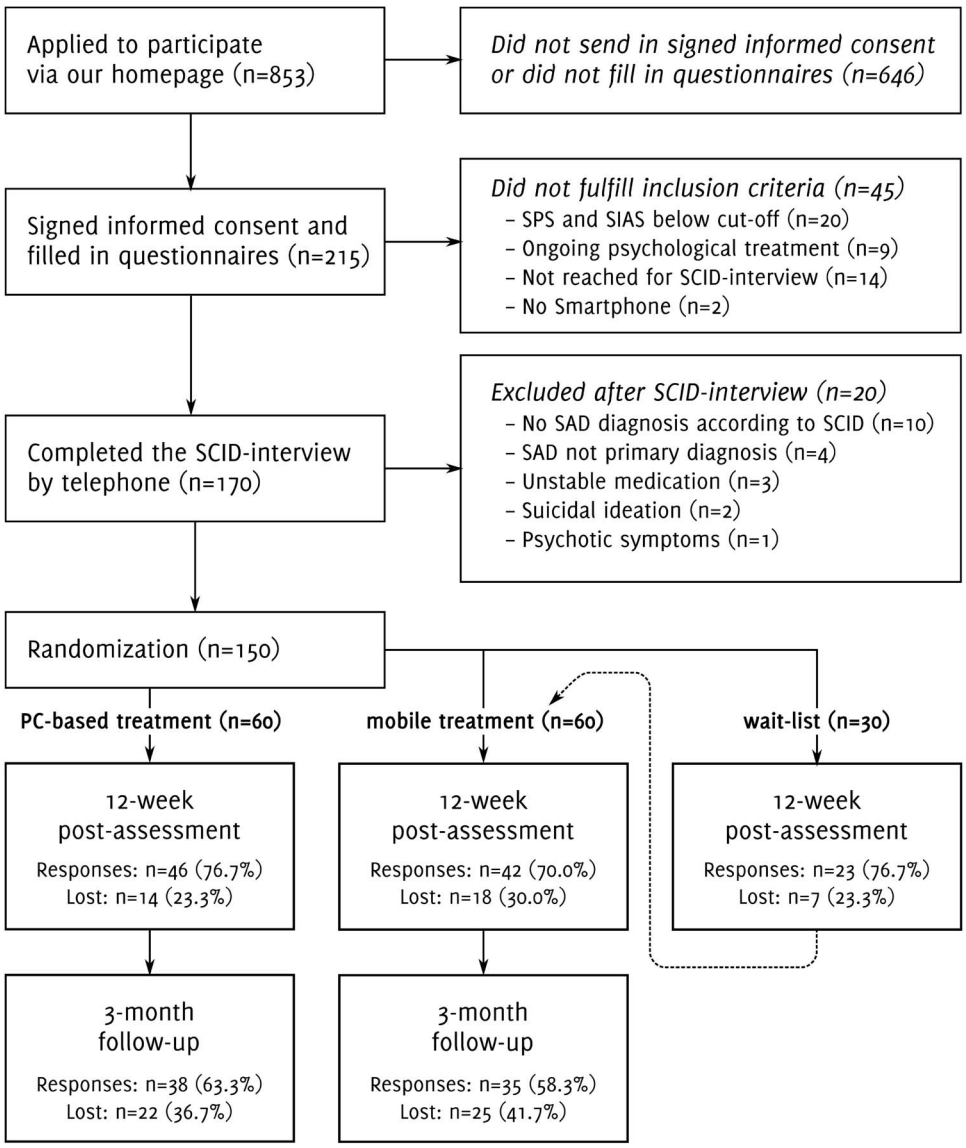

Figure 1. Selection, randomization and flow of participants throughout the trial. SCID = Structured Clinical Interview for *Diagnostic and Statistical Manual of Mental Disorders*, fourth edition (*DSM–IV*) Axis I disorders; SPS = Social Phobia Scale; SIAS = Social Interaction Anxiety Scale.

8. Relapse prevention: Strategies for maintaining the skills learned and preparation for possible relapses.

The participants were advised to work through one module per week and to start a new session after receiving a weekly feedback on their written diary entries and recent progress. At the end of the last module, participants were asked to repeat the exercises (e.g., exposure exercises) and to review the material as needed till postassessment at the end of Week 12. Completing one module unlocks the next, therefore, users are required to work through the contents sequentially.

Guidance. Timo Stolz and 10 advanced master's students in clinical psychology provided guidance during treatment. All coaches received a 1-day workshop in using the self-help program and on how to write the weekly feedback, based on case material from earlier trials. In addition, the coaches were continuously supervised by Timo Stolz and Thomas Berger. Each coach cared for seven to 20 participants, monitored their progress in the program, and contacted them via a secured text-based messaging system once a week to provide feedback and encourage further engagement. Besides, participants could send messages with specific questions and expect a response within three working days. The coaches' main role was to reinforce independent program use and keep up the participants' motivation. In case there was no activity by a participant during the week, coaches offered their help and assistance and asked if he or she was facing any problem with the program or with the tasks.

Outcome Measures

Participants completed all self-report measures using web-based versions at pretreatment, posttreatment (12 weeks) and follow-up (3 months after post).

Primary outcome measures. Primary outcome measures assessed self-reported symptoms of SAD using German versions of the SPS and SIAS and the Liebowitz Social Anxiety Scale—Self Report (LSAS-SR, Baker, Heinrichs, Kim, & Hofmann, 2002; Stangier & Heidenreich, 2004). These measures are commonly used in studies on SAD. The SPS and SIAS consist of each 20 items on a 5-point Likert scale, ranging from 0 to 4, resulting in a total score between 0 and 80. The LSAS-SR has 24 items, each of which is rated on a scale from 0 to 3. These three measures have shown to be valid and reliable with very good internal consistencies (Cronbach's alpha, SPS = .77–.94, SIAS = .84–.89, LSAS = .95), good retest-reliability (r_{tt}, SPS = .92, SIAS = .96, LSAS = .83) in English (Mattick & Clarke, 1998) and German samples (Stangier et al., 1999) and demonstrate similar properties when administered online (Cronbach's alpha, SPS = .89, SIAS = .86, Hedman et al., 2010). In the current sample, Cronbach's alpha for the SPS was .86 at pretreatment and .91 at posttreatment, the SIAS showed αs of .88 at pretreatment and .94 at posttreatment, and the LSAS had αs of .92 at pretreatment and .95 at posttreatment.

Besides these self-report measures, a second diagnostic interview was conducted at posttreatment to assess the number of participants no longer fulfilling the diagnostic criteria of SAD.

Secondary outcome measures. Secondary outcome measures consisted of self-reported assessments of depressive symptoms, interpersonal problems, quality of life, and overall psychiatric symptoms using the German versions of the Beck Depression Inventory-II (Hautzinger, Keller, & Kühner, 2006), the Inventory of Interpersonal Problems (IIP, Horowitz, Strauß, & Kordy, 2000), the psychological subscale of the Short Form-12 Health Survey (SF-12, Gandek et al., 1998), and the Brief Symptom Inventory (BSI, Derogatis, 1993; German version by Franke, 2000).

The Beck Depression Inventory-II is widely used to assess depressive symptoms both in research and in clinical practice. It consists of 21 items with Likert-ratings from 0 to 3 and robust psychometric properties for SAD clients in online assessments (Hedman et al., 2010). Cronbach's alpha in the current sample was .90 at pretreatment and .94 at posttreatment. The IIP was used to assess interpersonal problems. It has shown adequate psychometric properties (Horowitz, Rosenberg, Baer, Ureño, & Villaseñor, 1988) and consists of 64 items to be answered on a Likert scale from 0 to 4. We report mean scores over all items as a measure of interpersonal distress (McFarquhar, Luyten, & Fonagy, 2018; Tracey, Rounds, & Gurtman, 1996). Cronbach's alpha in the current sample was .91 at pretreatment and .95 at posttreatment.

The SF-12 is a briefer variant of the Short Form Health Survey (SF-36) and measures both physical and mental health. We only report the mental health subscale because physical health is not addressed by the treatment. Participants were asked to report the presence and severity of mental problems over the course of the last four weeks. The 12 items cover pain, psychological problems, as well as impairments in everyday functioning. With its good retest reliability and brevity, the SF-12 is widely used as an estimate of the general quality of life (Gandek et al., 1998). The BSI assesses psychological distress on nine dimensions such as anxiety, insecurity in social situations, depressiveness and compulsivity. In the current sample, Cronbach's alpha for the BSI was .94 at pretreatment and .96 at posttreatment. The participants are asked to rate the occurrence of symptoms within the last week on a 5-point Likert scale from 0 to 4. In this study, we report the Global Severity Index (GSI), which is the mean score over all items. As an economic instrument with robust psychometric properties, the BSI is commonly administered to detect prepost changes (Franke, 2000).

Power Calculation

The power analyses were conducted with G*Power 3 (Faul, Erdfelder, Lang, & Buchner, 2007). We aimed at detecting small to medium effect sizes of $f = 0.175$ and larger with regard to Time × Group interactions for the active conditions at an alpha error level of .05. Smaller effect sizes were considered to be irrelevant from a clinical point of view. A power analysis revealed that 37 participants were needed per active treatment group to show such an effect with a power ($1 - \beta$) of .80, assuming correlations of $r = .45$ between pre- and postmeasures as previously found in similar trials on ICBT for SAD (Berger et al., 2011, 2009). Sample size was further estimated based on a drop-out rate of 25%. Based on this assumption, we would have needed 47 participants in each treatment group. We finally decided to randomize 60 participants to each of the active conditions. A sensitivity power analysis revealed that given this sample size, a significance level of $p = .05$, and minimum power of .80, the test of interaction effect was sufficiently sensitive to detect effect sizes of $f = 0.130$ which corresponds to a Cohen's d of 0.26. Furthermore, 30 participants were estimated to be sufficient for the WL because effect sizes between the WL and treatment groups were assumed to be large based on the aforementioned earlier trials.

Statistical Analyses

All statistical analyses were performed with R (R Core Team, 2016) and the packages lme4 (Bates, Mächler, Bolker, & Walker, 2015) and mice (van Buuren & Groothuis-Oudshoorn, 2011). Analyses of variance (ANOVAs) and χ^2 tests were used to detect differences in demographic data, pretreatment measures, clinically significant improvement, and diagnostic status at posttreatment. Patterns of dropout were analyzed using a multivariate analysis of variance (MANOVA) and Little's (missing completely at random) MCAR test (Little, 1988). For the primary analysis of our results, we computed a composite score from the three social anxiety measures (SIAS, SPS, LSAS) to avoid multiple testing and potential inflation of alpha-error. Z-scores were standardized based on the pretreatment means and standard deviations. The composite score and, as secondary analyses, all primary and secondary outcome measures were analyzed with mixed-effect models using unstructured covariance matrices and restricted maximum likelihood estimation (Bates et al., 2015) with time-points nested within subjects. This approach uses all available data of each subject without substituting missing values and allows the inclusion of all participants in the analyses, following the intention-to-treat principle. For robustness, these analyses were repeated using multiple imputation and repeated-measures ANOVA with time (prepost) as a within-group factor and treatment condition as a between-groups

factor. The models were further examined using contrast analyses. Within- and between-groups effect sizes (Cohen's *d*) were calculated based on estimated means and the pooled standard deviation from the observed means. In addition, differences between the two active conditions were analyzed using a JZS Bayes factor *t* test. This test was computed using JASP (JASP Team, 2017) and JASPs' default priors together with a robustness analysis for various prior widths. Within-group changes in outcome scores from posttreatment to follow-up were analyzed for the active conditions only, as the WL was assigned to treatment after 12 weeks. This was done with another mixed-effects model per outcome measure that involved all time points to retain as much information as possible in case of missing values. Changes from post to follow-up were then tested with Tukey's honest significant difference (HSD) post hoc tests. We also report the number of participants who clinically significantly improved and who reliably deteriorated. Cases were counted as clinically significantly improved if they changed reliably according to the reliable change index (RCI; Jacobson, Follette, & Revenstorf, 1984) and belonged to a functional population at posttreatment according to criterion 'c' of Jacobson et al. (1984), and under consideration of the normative data provided by Stangier et al. (1999). As normative data had been unavailable for the LSAS, cases were counted as clinically significantly improved when the posttreatment score was at least 2 standard deviations below the baseline assessment, which is in accordance with criterion 'a' of Jacobson et al. (1984). The following thresholds were computed for the primary outcome measures: for the SPS, 15.34, RCI = 7.39; for the SIAS, 24.05, RCI = 10.33; and for the LSAS, 42.24, RCI = 23.89. Reliable deterioration was defined as a prepost increase on the SPS, the SIAS or the LSAS of at least the RCI defined above. With regard to clinically significant improvement we also report numbers needed to treat (NNT). NNT is the number of patients who need to be treated to achieve one additional positive outcome—in comparison to another treatment or a placebo. When p_{WL} is the probability of remissions in the WL, and p_{App} is the probability of being cured by the app treatment, the *NNT* equals to $1/(p_{App} - p_{WL})$.

Results

Pretreatment Evaluation

The conditions did not differ in symptom severity or demographic variables with two exceptions: Participants in the mobile condition were less educated, $\chi^2(6, N = 150) = 14.29$, $p = .03$, $V = 0.22$, and had a higher GSI prescore, $t(117.85) = -2.46$, $p = .02$, $d = 0.45$, than participants in the PC-based condition. Table 1 provides further details.

Dropout Analysis

In total, 39 participants (26%) did not complete the posttreatment assessment, although they had been invited three times in weekly intervals via e-mail (PC, $n = 14$; app, $n = 18$; and WL, $n = 7$). Noncompletion rates did not differ with respect to experimental group, $\chi^2(2, N = 150) = 0.83$, $p = .66$, $V = 0.07$, nor demographic data, $\chi^2(1–16, N = 150) = 0.07–19.22$, $ps = .12–.79$, $V = 0.02–0.26$, except for education: A Mann–Whitney test indicated that completers were significantly more educated than noncompleters, $W = 2654$, $Z = 2.24$, $p = .02$, $r = .18$.

Table 1
Baseline Demographics and Sample Characteristics for the Treatment and Control Groups

Variable	PC ($n = 60$)		App ($n = 60$)		Wait-list ($n = 30$)		Statistic of difference			
	n	%	*n*	%	*n*	%	*df*	χ^2/F	*p*	R^2/V
Age, M | *SD* (years)	34.6	12.0	34.7	9.9	35.2	12.1	2; 147	.04	.97	.004
Gender							2	4.82	.09	.18
Male	25	41.7	25	41.7	6	20.0				
Female	35	58.3	35	58.3	24	80.0				
Marital status							8	3.79	.88	.11
Divorced	0	.0	1	1.7	0	.0				
Single	31	51.7	27	45.0	16	53.3				
Living together	14	23.3	17	28.3	7	23.3				
Married	15	25.0	14	23.3	7	23.3				
Widowed	0	.0	1	1.7	0	.0				
Highest education							6	14.29	.03	.21
Compulsory school	1	1.7	5	8.3	2	6.7				
Apprenticeship	8	13.3	22	36.7	7	23.3				
College	20	33.3	10	16.7	6	20.0				
University	31	51.7	23	38.3	15	50.0				
Employment							10	6.66	.76	.14
Unemployed	3	5.0	5	8.3	4	13.3				
Student	19	31.7	13	21.7	8	26.7				
Retired	2	3.3	5	8.3	1	3.3				
At-home parent	1	1.7	3	5.0	2	6.7				
Part-time paid work	11	18.3	12	20.0	6	20.0				
Full-time paid work	24	40.0	22	36.7	9	30.0				
Prior psychotherapy	34	56.7	43	71.7	16	53.3	2	4.06	.13	.16
Current medication	9	15.0	15	25.0	2	6.7	2	5.07	.08	.18

Note. App = App-based treatment condition; PC = personal computer-based treatment condition.

Furthermore, dropout was unrelated to symptom severity at baseline according to Little's MCAR test, $\chi^2(6) = 9.34$, $p = .16$, and confirmed with a MANOVA across all primary and secondary measures with missingness as grouping factor, $F(6, 143) = 1.58$, $p = .16$, partial $\eta^2 = 0.06$.

Treatment Outcomes

Observed and estimated means for all self-report measures are presented in Table 2.

Composite score. A linear mixed model for composite symptoms showed a significant Time × Group interaction, $F(2, 120.09) = 14.18$, $p < .01$. A consecutive contrast analysis showed that both active treatments outperformed the WL at post, $t(119.46) = 5.08$, $p < .01$, $d = 1.07$, with no significant difference between the active conditions but a trend in favor of the app condition, $t(120.75) = 1.71$, $p = .09$, $d = 0.30$.

Separate measures. Further linear mixed models with group as a fixed factor and time as a repeated factor (pre—post) were fitted separately for each of the dependent measures. Significant Time × Group interaction effects have been found for all primary and secondary outcomes, $F(2, 117.46–121.83) = 4.08–11.71$, $ps = .01–.02$. Consecutive contrast analyses showed that both active treatments were significantly superior to the WL on all involved measures, $t(113.84–121.67) = -2.78– 4.60$, $ps \leq .01$, $ds = 0.53–1.01$. However, there was no significant difference among the active conditions—neither on primary outcomes, $t(117.90–121.49) = 1.28–1.61$, $ps = .11–.20$, $ds = 0.20–0.30$, nor on secondary outcomes, $t(114.85–123.14) = -0.71– 1.06$, $ps = .29–.71$, $ds = -0.11–0.20$, except for the GSI in which the mobile group fared significantly better, $t(120.16) = 2.06$, $p = .04$, $d = 0.38$. When adjusting the p values for multiple testing, both active treatments still significantly outperformed the WL on all involved measures, all adjusted $ps \leq .01$.

Because neither of the active treatments was found to be superior over the other, neither on the primary outcomes nor on the secondary measures, standardized differences of estimated pre-to-post changes are reported together with their confidence intervals in the supplementary online materials. On the SPS, the SIAS, and

Table 2
Observed and Estimated Means for Primary and Secondary Outcome Measures and Within- and Between-Group Effect Sizes

	Pretreatment		Posttreatment (observed)		Posttreatment (estimated)	Follow-up (estimated)	Pre–post within-group effect sizes (estimated means)		Between-group effect sizes at posttreatment (estimated means)		
Measure	*M* (*SD*)	*n*	*M* (*SD*)	*n*	*M* (*SE*)	*M* (*SE*)	*d*	95% CI	Groups	*d*	95% CI
Composite									$F(2, 120.08) = 14.21$, $p < .01$		
PC	−.08 (.9)	60	−1.26 (1.0)	46	−1.28 (.1)	−1.36 (.1)	1.25	[0.82, 1.66]	App vs PC	.07	[−.34, .49]
App	.14 (.7)	60	−1.37 (.9)	42	−1.35 (.1)	−1.41 (.1)	1.84	[1.37, 2,31]	App vs WL	.89	[.35, 1.41]
WL	−.13 (.9)	30	−.45 (1.0)	23	−.53 (.2)		.44	[−0.11, 0.99]	PC vs WL	.74	[.22, 1.25]
SPS									$F(2, 117.24) = 9.77$, $p < .01$		
PC	37.1 (13.2)	60	22.0 (12.9)	46	21.7 (1.8)	19.6 (1.9)	1.17	[0.76, 1.59]	App vs PC	−.08	[−.49, .34]
App	41.1 (13.8)	60	22.0 (11.7)	42	22.6 (1.9)	20.6 (2.0)	1.42	[0.98, 1.86]	App vs WL	.92	[.38, 1.45]
WL	38.9 (12.4)	30	34.7 (11.9)	23	33.5 (2.6)		.45	[−0.10, 1.00]	PC vs WL	.93	[.40, 1.45]
SIAS									$F(2, 120.88) = 10.38$, $p < .01$		
PC	51.3 (15.0)	60	36.2 (16.9)	46	36.0 (2.0)	34.1 (2.1)	.97	[0.56, 1.37]	App vs PC	.19	[−.23, .61]
App	52.3 (10.4)	60	33.2 (14.3)	42	33.0 (2.1)	31.3 (2.2)	1.59	[1.13, 2,03]	App vs WL	.60	[.08, 1.12]
WL	46.6 (13.9)	30	43.3 (14.6)	23	41.6 (2.8)		.35	[−0.20, 0.90]	PC vs WL	.35	[−.16, .85]
LSAS									$F(2, 120.62) = 11.74$, $p < .01$		
PC	81.5 (22.1)	60	54.8 (12.3)	42	54.6 (3.2)	51.3 (3.5)	1.18	[0.76, 1.59]	App vs PC	.05	[−.37, .47]
App	87.3 (19.6)	30	52.7 (13.9)	23	53.5 (3.3)	51.1 (3.6)	1.58	[1.13, 2,03]	App vs WL	.82	[.29, 1.35]
WL	82.6 (20.7)	60	74.4 (.5)	46	74.0 (4.5)		.36	[−0.19, 0.91]	PC vs WL	.78	[.26, 1.29]
BDI									$F(2, 117.59) = 4.45$, $p = .01$		
PC	17.2 (9.2)	60	10.3 (8.1)	46	10.2 (1.5)	9.9 (1.6)	.79	[0.39, 1.19]	App vs PC	−.30	[−.72, .12]
App	19.6 (11.6)	60	12.7 (12.3)	42	13.3 (1.6)	13.7 (1.6)	.52	[0.12, 0.92]	App vs WL	.29	[−.22, .80]
WL	17.4 (12.3)	30	18.2 (13.9)	23	17.0 (2.2)		.03	[−0.52, 0.57]	PC vs WL	.65	[.14, 1.16]
GSI									$F(2, 119.57) = 5.9$, $p < .01$		
PC	1.2 (.5)	60	.8 (.5)	46	.7 (.1)	.7 (.1)	.88	[0.47, 1.28]	App vs PC	−.08	[−.49, .34]
App	1.5 (.6)	60	.8 (.6)	42	.8 (.1)	.8 (.1)	1.14	[0.72, 1.57]	App vs WL	.35	[−.16, .86]
WL	1.3 (.6)	30	1.1 (.6)	23	1.0 (.1)		.41	[−0.14, 0.96]	PC vs WL	.48	[−.03, .98]
IIP									$F(2, 114.33) = 6.94$, $p < .01$		
PC	1.8 (.5)	60	1.5 (.5)	46	1.4 (.1)	1.3 (.1)	.70	[0.31, 1.10]	App vs PC	.06	[−.36, .48]
App	1.8 (.4)	60	1.4 (.5)	42	1.4 (.1)	1.3 (.1)	.96	[0.54, 1.37]	App vs WL	.60	[.07, 1.11]
WL	1.8 (.5)	30	1.8 (.6)	23	1.7 (.1)		.12	[−0.42, 0.66]	PC vs WL	.54	[.03, 1.05]
SF-12									$F(2, 122.39) = 4.1$, $p = .02$		
PC	34.0 (10.2)	60	41.8 (11.3)	46	41.4 (1.4)	41.8 (1.6)	−.69	[−1.09, −0.30]	App vs PC	.10	[−.31, .52]
App	31.4 (8.7)	60	40.0 (10.4)	42	40.3 (1.5)	41.8 (1.7)	−.94	[−1.35, −0.52]	App vs WL	−.34	[−.85, .18]
WL	34.9 (10.4)	30	35.9 (10.1)	23	36.8 (2.0)		−.19	[−0.73, 0.36]	PC vs WL	−.42	[−.93, .09]

Note. App = App-based treatment condition; PC = PC-based treatment condition; WL = wait list control condition; Composite = Composite score across SPS, SIAS, and LSAS; SPS = Social Phobia Scale; SIAS = Social Interaction Anxiety Scale; LSAS = Liebowitz Social Anxiety Scale; BDI-II = Beck Depression Inventory-II; GSI = Global Severity Index; IIP = Inventory of Interpersonal Problems; SF-12 = Short Form-12 health survey, mental subscale.

the LSAS, the 95% confidence intervals of these differences are bounded to be lower than 0.12 *SD*s in favor of the PC-based version. Therefore, even in the strongest case there would only be a small effect in favor of the PC-based delivery option. Furthermore, a JZS Bayes factor *t* test suggested that the data was 9.89 times more likely to be observed when the mobile treatment was as efficacious as the PC-based version. This gives moderate evidence that the mobile treatment is at least not less efficacious than the PC-based version. Further details of the test are reported in the supplementary online materials, including a robustness analysis for a broad range of priors, Supplemental Figure S7.

Sensitivity Analysis

Regarding the composite score and single primary outcomes, the same pattern of results was found (a) on completers' data, (b) using multiple imputation, and (c) when controlling for education, current medication or prior psychotherapy. The same pattern was also found for secondary outcomes, except for the GSI: The difference between both active groups was found to be significant with procedure (c) but not with procedures (a) or (b). Overall, the analyses of primary and secondary outcomes were robust. Nevertheless, results regarding the GSI have to be interpreted with caution. The significant results found in the corresponding linear mixed effects model could have been biased by dropout of more severely affected participants.

Effect Sizes

Effect sizes (Cohen's *d*) are presented in Table 2. The between-groups effect sizes at posttreatment on the composed social anxiety measures were $d = 0.07$ for app versus PC (in favor of the app), $d = 0.89$ for app vs. WL, and $d = 0.74$ for PC v. WL. Mean between-groups effect sizes on secondary outcome measures were $d = 0.10$ for PC versus app (in favor of the PC-based version), $d = 0.39$ for app vs. WL, and $d = 0.52$ for PC v. WL. Within-group comparisons on the composite outcome and based on estimated means revealed large effect sizes in the PC-based group ($d = 1.25$) as well as in the mobile group ($d = 1.84$; see Table 2 for detailed information).

Clinically Significant Improvement

The WL had no cases of clinically significant improvement (CSI) on the SPS and the SIAS, but two such cases (6.7%) on the LSAS. In contrast, the PC-based condition achieved higher amounts of such cases on the SPS ($n = 16$ [26.67%], NNT = 3.75, $>\chi^2(1, N = 90) = 7.99$, $p < .01$, $V = 0.30$), the SIAS ($n = 12$ [20.00%], NNT = 5.00, $\chi^2(1, N = 90) = 5.30$, $p = .02$, $V = 0.24$), and the LSAS ($n = 13$ [21.67%], NNT = 6.67, $\chi^2(1, N = 90) = 2.25$, $p = .13$, *ns*, $V = 0.16$). The same applies to the mobile condition, again on the SPS ($n = 12$ [20.00%], NNT = 5.00, $\chi^2(1, N = 90) = 5.30$, $p = .02$, $V = 0.24$), the SIAS ($n = 10$ [16.67%], NNT = 6.00, $\chi^2(1, N = 90) = 4.06$, $p = .04$, $V = 0.21$), and the LSAS ($n = 12$ [20.00%], NNT = 7.50, $\chi^2(1, N = 90) = 1.79$, $p = .18$, *ns*, $V = 0.14$). No significant differences were found between the two active conditions.

Diagnostic Status at Post-Treatment

In total, 92 participants could be reached for a second clinical interview after the treatment (app, $n = 34$ [57%]; PC, $n = 37$ [62%]; and WL, $n = 21$ [70%]; with $\chi^2(2) = 1.50$, $p = .47$, for differences in response rate). To deal with a total of 39% of missing interviews, we administered several procedures to examine the fraction of cured participants. First, we analyzed completers only (PC: 14/37 [37.8%], app: 9/34 [26.5%]). Second, we considered all missings as still suffering from SAD (PC: 14/60 [23.3%], app: 9/60 [15.0%]). Third, we considered all missings as still suffering from SAD unless they showed a clinically significant improvement on the self-reported primary outcomes (PC: 18/60 [30.0%], app: 10/60 [16.7%]).

Regardless of the procedure, all participants in the WL were still diagnosed with SAD. The PC-based group fared significantly better than the WL, regardless of the adopted procedure, $\chi^2(1, N = 58–90) = 6.61–9.45$, $ps \leq .01$, $V = 0.27–0.38$. The mobile group showed significant differences for the first and the third procedure, $\chi^2(1, N = 55–90) = 4.06–4.85$, $ps = .03–.04$, $V = 0.21–0.30$, but not for the second procedure, $\chi^2(1, N = 90) = 3.47$, $p = .06$, $V = 0.20$. Differences between the active groups were not significant regardless of the approach.

Treatment Effects at Three-Month Follow-Up

All analyses in this section only include the two active conditions, as the WL had already received access to the mobile version of the treatment. Mixed-models analyses including pre, post, and follow-up scores showed significant time effects for all scales, $F(2, 166.86–176.07) = 31.99–150.09$, $ps < .01$. Contrast analyses indicate that follow-up scores improved from baseline, $t(170.41–180.46) = -14.96–8.07$, $ps < .01$, $ds = 0.75–1.52$, and post hoc tests using Tukey's HSD indicate stability from posttreatment to follow-up, as no significant changes were detected, $ps = .05–.98$. All Time × Group interactions were nonsignificant, $F(2, 166.86–176.07) = 0.23–2.04$, $ps = .13–.79$, thus through all time points, neither of the delivery options were significantly superior.

Reliable Deterioration

Two participants (1.3%) achieved reliable deterioration on the SPS (app, $n = 1$, WL, $n = 1$) and one participant (0.7%) deteriorated reliably in the PC-based group on the SIAS.

Program Usage

Details on program usage are provided as supplementary online material. In summary, the mobile treatment was used less extensively in terms of (a) the number of performed exercises, (b) recorded diary entries, (c) number and length of messages written by patients, (d) single actions like clicks or taps, (e) time spent using the program, and (f) completed modules. In addition, residualized posttreatment outcome was predicted by active program usage. However, significant correlations were found in the mobile condition only. The most influential predictor was the number of tracked exposure exercises ($\rho = .47$, $p < .01$). For these correlative analyses, residual gain scores of the composite outcome and computed Spearman's rank correlation coefficients were used.

Single actions like clicks or taps were distributed differently over the course of the day in the two conditions (Figure 3 and 4). In the mobile condition, the usage was spread more evenly across the day, whereas members of the PC-based condition performed a higher fraction of their work during the night (see Figure 3). A Kolmogorov–Smirnov test suggested nonequal distributions, $D = 0.14$, $p < .01$. Furthermore, there was a significant Hour X Group interaction for the count of actions, $F(23, 1553.87) = 1.93$, $p = .01$, indicating a different usage pattern over time in the two conditions. Finally, usage decreased significantly after postassessment, however, without significant differences among the active conditions.

Discussion

This RCT compares the efficacy of a mobile and a PC-based SAD treatment program to a WL. To the best of our knowledge, this is the first large RCT comparing mobile and PC-based dissemination of the same guided self-help treatment. In addition, it is the largest trial using smartphones to target a specific disorder. We found two prior pilot studies that have been either rather small (Watts et al., 2013; $N_{PC\text{-}vs.\text{-}App} = 35$) or were not targeting a specific psychological disorder (weight loss; Carter et al., 2013; $N_{PC\text{-}vs.\text{-}App} = 85$). In the present study, both treatment formats were effective in reducing symptoms of SAD and increasing psychological well-being after 12 weeks of treatment. The within-group effect sizes of $d = 1.84$ (mobile) and $d = 1.25$ (PC) are in line with recent meta-analyses on ICBT for SAD (Andrews, Cuijpers, Craske, McEvoy, & Titov, 2010; Spek et al., 2007) and earlier trials using the same materials (Berger et al., 2009, 2010, 2011; Boettcher et al., 2012; Schulz et al., 2016). Of all treated participants, 28.3% achieved clinically significant improvement on at least one of the primary outcome scales (mobile: 23.3%, PC: 33.3%), compared to none in the WL. Furthermore, 19.2% of the participants did not fulfill the diagnostic criteria of SAD at posttreatment when missings were regarded as failures to provide conservative estimates (mobile: 15.0%, PC: 23.3%) compared to none in the WL. Improvements were maintained six months after randomization. Given that 62.0% of participants had been treated with psychotherapy in the past and that participants showed high comorbidity rates, these results can be interpreted as encouraging.

The effect sizes between the active treatment groups and the WL were at $d = 0.89$ (mobile) and $d = 0.74$ (PC) for the composite social anxiety score. These between-groups effect sizes can be compared with the effects on traditional therapist-delivered CBT reported in a meta-analysis by Acarturk, Cuijpers, van Straten, and de Graaf (2009), in which a subgroup analysis of 35 studies with wait list comparisons revealed a mean effect size of $d = 0.86$ on social anxiety measures for therapist-delivered CBT, cognitive therapy, social skills training, relaxation, and/or exposure. However, the between-groups effect sizes in the current study are slightly below the effect sizes reported in a more recent meta-analysis which found standardized mean differences of Hedges' $g = 1.19$ for individual CBT versus wait lists (Mayo-Wilson et al., 2014). Also, the percentage of participants who clinically significantly improved was below the range of that reported in studies on evidence-based face-to-face CBT, though comparisons are difficult because definitions of clinically significant improvement vary across studies. For instance, a trial on face-to-face CBT based on the same approach as our study (Clark & Wells, 1995) reported high clinically significant improvement rates of 74% on the SPS and of 63% on the SIAS for individual CBT (Stangier, Heidenreich, Peitz, Lauterbach, & Clark, 2003). However, the mean scores of the two studies at posttreatment are quite similar (e.g., SPS posttreatment Stangier et al.: 21.5 vs. SPS posttreatment present study: 22.0). Thus, the low CSI rates in the current study may be partly due to a conservative estimate. Another reason for the low CSI rates is that the present sample was highly impaired. Pretreatment scores on social anxiety measures in the current study were about half a standard deviation higher than pretreatment scores reported in studies on face-to-face CBT (e.g., Stangier, Heidenreich, Peitz, et al., 2003). The high scores at pretreatment could have reduced the chance of falling below the cut-off. Furthermore, the high pretreatment scores also indicate that our sample may be representative for a severely impaired population of SAD patients that typically seek information about the disorder on the Internet (Erwin, Turk, Heimberg, Fresco, & Hantula, 2004).

Regarding the comparison between the two active conditions, neither of the two treatment formats was significantly superior to the other, except for the GSI in which the mobile group showed greater improvements. However, the difference in gains is small and might in part be explained by regression toward the mean, as the mobile group had a significantly higher GSI prescore. With regard to the primary self-report measures, the mobile treatment can be considered as efficacious as the PC-based version, although the mobile version was less extensively used in terms of (a) performed exercises, (b) recorded diary entries, (c) number and length of written messages, (d) single actions like clicks or taps, (e) time spent using the program, and (f) completed modules. Taken together, the mobile delivery might require a smaller dose to achieve the same effect. This might be due to different mechanisms of change, which are indicated by the differing correlations of usage and outcome and, even more, by the different usage patterns during the course of a typical day. However, only little is known about the association of adherence, usage patterns, and outcome in Internet-based treatments (Donkin et al., 2011). Thus, further research should address differences in mechanisms of change for mobile and PC-based delivery. It could be assumed that exposure exercises have a stronger impact on outcome when they are prepared on a mobile device and then conducted spontaneously, whenever the opportunity arises. It might be of further advantage that the exercise can then be immediately recorded. This immediacy prevents postevent processing from diminishing the sense of progress. However, and importantly, according to the current findings an evidence-based treatment can be presented both on mobile devices and on the PC without altering the efficacy.

Limitations

Several limitations of this study need to be considered. First, participants had to work either with a PC or with a smartphone and could therefore not be blinded to treatment allocation. The same is true for their guiding coaches and the diagnostic interviewers, as most of the participants unwittingly revealed their mode of treatment within the messages and/or during the interview and no measures were taken to prevent participants from breaking the blind. In addition, the interrater reliability of SAD diagnoses was not formally established. Second, the findings provide evidence for

this particular treatment and cannot be easily generalized to other programs or disorders. However, similar results were found in small pilot studies on depression (Watts et al., 2013) and weight loss (Carter et al., 2013), which also compared mobile and PC-based treatments, and one trial with a mobile treatment for SAD is announced (Miloff, Marklund, & Carlbring, 2015). Nevertheless, findings from this trial should be replicated in other disorders. Third, the limited power of this study did not allow us to detect small effects between the two active groups at posttreatment. At the level of an individual treatment, such small effects might be less important than other therapy-related factors, but from a public health perspective, small effects with large populations can have great societal impact. Therefore, our findings should be replicated in larger samples and in an unguided format, because unguided treatments would be especially suitable for public health interventions. To deal with power limitations in this trial, interaction effects were analyzed using contrast analyses to provide a more sensitive estimation of differences in efficacy between the active conditions. Contrast analyses have more power as opposed to comparisons at one point in time, as they can account for repeated measures and thus are able to assign variance within the subjects. Furthermore, a JZS Bayes factor *t* test was added to assess the likelihood of similar effects for both active conditions. This test suggested that the data was 9.89 times more likely to be observed when the mobile treatment was at least as efficacious as the PC-based version, instead of the PC-based version being truly superior. Fifth, this study tested a self-selected sample with individuals who expressed their interest in an Internet-based delivery. In addition, treatment expectancy was not assessed. Therefore, we could not establish whether the results might depend on particularly motivated patients. However, other studies indicate that Internet-based interventions work just as well in routine practice as in research (e.g., Andersson & Hedman, 2013).

Conclusion

Despite these limitations, the present study extends the existing knowledge on ICBT by systematically testing a novel delivery option. Besides easier integration in daily life and routine, mobile CBT could have promise for low and middle-income countries due to a large unmet need for mental health services (Saraceno et al., 2007), in which the coverage of smartphones with an Internet subscription is much higher than those of online desktop computers might ever be (Aranda-Jan, Mohutsiwa-Dibe, & Loukanova, 2014).

In conclusion, this study provides evidence that SAD can be effectively treated with mobile self-help applications, with treatment gains being maintained 6 months after randomization.

References

Acarturk, C., Cuijpers, P., van Straten, A., & de Graaf, R. (2009). Psychological treatment of social anxiety disorder: A meta-analysis. *Psychological Medicine, 39,* 241–254. http://dx.doi.org/10.1017/S0033291708003590

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders (DSM–5).* Washington, DC: American Psychiatric Association Publishing.

Andersson, G. (2016). Internet-delivered psychological treatments. *Annual Review of Clinical Psychology, 12,* 157–179. http://dx.doi.org/10.1146/annurev-clinpsy-021815-093006

Andersson, G., Cuijpers, P., Carlbring, P., Riper, H., & Hedman, E. (2014). Guided Internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: A systematic review and meta-analysis. *World Psychiatry, 13,* 288–295. http://dx.doi.org/10.1002/wps.20151

Andersson, G., & Hedman, E. (2013). Effectiveness of guided internet-based cognitive behavior therapy in regular clinical settings. *Verhaltenstherapie, 23,* 140–148. http://dx.doi.org/10.1159/000354779

Andrews, G., Cuijpers, P., Craske, M. G., McEvoy, P., & Titov, N. (2010). Computer therapy for the anxiety and depressive disorders is effective, acceptable and practical health care: A meta-analysis. *PLoS ONE, 5*(10), e13196. http://dx.doi.org/10.1371/journal.pone.0013196

Andrews, G., Davies, M., & Titov, N. (2011). Effectiveness randomized controlled trial of face to face versus Internet cognitive behaviour therapy for social phobia. *Australian and New Zealand Journal of Psychiatry, 45,* 337–340. http://dx.doi.org/10.3109/00048674.2010.538840

Aranda-Jan, C. B., Mohutsiwa-Dibe, N., & Loukanova, S. (2014). Systematic review on what works, what does not work and why of implementation of mobile health (mHealth) projects in Africa. *BMC Public Health, 14,* 188. http://dx.doi.org/10.1186/1471-2458-14-188

Baker, S. L., Heinrichs, N., Kim, H.-J., & Hofmann, S. G. (2002). The liebowitz social anxiety scale as a self-report instrument: A preliminary psychometric analysis. *Behaviour Research and Therapy, 40,* 701–715. http://dx.doi.org/10.1016/S0005-7967(01)00060-2

Bang, M., Timpka, T., Eriksson, H., Holm, E., & Nordin, C. (2007). Mobile phone computing for in-situ cognitive behavioral therapy. In *Medinfo 2007: Proceedings of the 12th World Congress on Health (Medical) Informatics; building sustainable health systems* (pp. 1078–1082). Clifton, VA: IOS Press.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67,* 1–48. http://dx.doi.org/10.18637/jss.v067.i01

Berger, T., Boettcher, J., & Caspar, F. (2014). Internet-based guided self-help for several anxiety disorders: A randomized controlled trial comparing a tailored with a standardized disorder-specific approach. *Psychotherapy: Theory, Research, & Practice, 51,* 207–219. http://dx.doi.org/10.1037/a0032527

Berger, T., Caspar, F., Richardson, R., Kneubühler, B., Sutter, D., & Andersson, G. (2011). Internet-based treatment of social phobia: A randomized controlled trial comparing unguided with two types of guided self-help. *Behaviour Research and Therapy, 49,* 158–169. http://dx.doi.org/10.1016/j.brat.2010.12.007

Berger, T., Hohl, E., & Caspar, F. (2009). Internet-based treatment for social phobia: A randomized controlled trial. *Journal of Clinical Psychology, 65,* 1021–1035. http://dx.doi.org/10.1002/jclp.20603

Berger, T., Hohl, E., & Caspar, F. (2010). Internetbasierte Therapie der Sozialen Phobie: Ergebnisse einer 6-monate-katamnese [Internet-based treatment for social phobia: A 6-month follow-up]. *Zeitschrift für Klinische Psychologie und Psychotherapie, 39,* 217–221. http://dx.doi.org/10.1026/1616-3443/a000050

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology, 64,* 417–444. http://dx.doi.org/10.1146/annurev-psych-113011-143823

Boettcher, J., Berger, T., & Renneberg, B. (2012). Does a pre-treatment diagnostic interview affect the outcome of internet-based self-help for social anxiety disorder? a randomized controlled trial. *Behavioural and Cognitive Psychotherapy, 40,* 513–528. http://dx.doi.org/10.1017/S1352465812000501

Boettcher, J., Carlbring, P., Renneberg, B., & Berger, T. (2013). Internet-based interventions for social anxiety disorder: An overview. *Verhaltenstherapie, 23,* 160–168. http://dx.doi.org/10.1159/000354747

Boettcher, J., Rozental, A., Andersson, G., & Carlbring, P. (2014). Side effects in internet-based interventions for social anxiety disorder. *Inter-*

net Interventions, 1, 3–11. http://dx.doi.org/10.1016/j.invent.2014.02.002

Boschen, M. J., & Casey, L. M. (2008). The use of mobile telephones as adjuncts to cognitive behavioral psychotherapy. *Professional Psychology: Research and Practice, 39,* 546–552. http://dx.doi.org/10.1037/0735-7028.39.5.546

Botella, C., Gallego, M. J., Garcia-Palacios, A., Guillen, V., Baños, R. M., Quero, S., & Alcañiz, M. (2010). An Internet-based self-help treatment for fear of public speaking: A controlled trial. *Cyberpsychology, Behavior and Social Networking, 13,* 407–421. http://dx.doi.org/10.1089/cyber.2009.0224

Burns, M. N., Begale, M., Duffecy, J., Gergle, D., Karr, C. J., Giangrande, E., & Mohr, D. C. (2011). Harnessing context sensing to develop a mobile intervention for depression. *Journal of Medical Internet Research, 13*(3), e55. http://dx.doi.org/10.2196/jmir.1838

Carter, M. C., Burley, V. J., Nykjaer, C., & Cade, J. E. (2013). Adherence to a smartphone application for weight loss compared to website and paper diary: Pilot randomized controlled trial. *Journal of Medical Internet Research, 15*(4), e32. http://dx.doi.org/10.2196/jmir.2283

Clark, D. M., & Wells, A. (1995). Social phobia: Diagnosis, assessment, and treatment. In R. G. Heimberg, M. R. Liebowitz, D. A. Hope, & F. R. Schneier (Eds.), *Social phobia: Diagnosis, assessment and treatment* (pp. 69–93). New York, NY: Guilford Press.

Dagöö, J., Asplund, R. P., Bsenko, H. A., Hjerling, S., Holmberg, A., Westh, S., . . . Andersson, G. (2014). Cognitive behavior therapy versus interpersonal psychotherapy for social anxiety disorder delivered via smartphone and computer: A randomized controlled trial. *Journal of Anxiety Disorders, 28,* 410–417. http://dx.doi.org/10.1016/j.janxdis.2014.02.003

Derogatis, L. R. (1993). *Brief Symptom Inventory: Administration, scoring, and procedures manual*. Minneapolis, MN: National Computer Systems.

Donker, T., Petrie, K., Proudfoot, J., Clarke, J., Birch, M.-R., & Christensen, H. (2013). Smartphones for smarter delivery of mental health programs: A systematic review. *Journal of Medical Internet Research, 15*(11), e247. http://dx.doi.org/10.2196/jmir.2791

Donkin, L., Christensen, H., Naismith, S. L., Neal, B., Hickie, I. B., & Glozier, N. (2011). A systematic review of the impact of adherence on the effectiveness of e-therapies. *Journal of Medical Internet Research, 13*(3), e52. http://dx.doi.org/10.2196/jmir.1772

Erwin, B. A., Turk, C. L., Heimberg, R. G., Fresco, D. M., & Hantula, D. A. (2004). The Internet: Home to a severe population of individuals with social anxiety disorder? *Journal of Anxiety Disorders, 18,* 629–646. http://dx.doi.org/10.1016/j.janxdis.2003.08.002

Fanning, J., Mullen, S. P., & McAuley, E. (2012). Increasing physical activity with mobile devices: A meta-analysis. *Journal of Medical Internet Research, 14*(6), e161. http://dx.doi.org/10.2196/jmir.2171

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39,* 175–191. http://dx.doi.org/10.3758/BF03193146

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (1997). *Structured Clinical Interview for DSM–IV–TR Axis I Disorders, Research Version, Patient Edition with Psychotic Screen (SCID-I/P W/PSY SCREEN)*. New York, NY: Biometrics Research, New York State Psychiatric Institute.

Franke, G. H. (2000). *Brief Symptom Inventory von L. R. Derogatis (Kurzform der SCL90-R) - deutsche Version*. Göttingen, Germany.

Gandek, B., Ware, J. E., Aaronson, N. K., Apolone, G., Bjorner, J. B., Brazier, J. E., . . . Sullivan, M. (1998). Cross-validation of item selection and scoring for the SF-12 Health Survey in nine countries: Results from the IQOLA Project. International Quality of Life Assessment. *Journal of Clinical Epidemiology, 51,* 1171–1178. http://dx.doi.org/10.1016/S0895-4356(98)00109-7

Hautzinger, M., Keller, F., & Kühner, C. (2006). *Beck Depressions-Inventar (BDI-II)*. Frankfurt, Germany: Harcourt Test Services.

Hedman, E., Andersson, E., Ljótsson, B., Andersson, G., Rück, C., & Lindefors, N. (2011). Cost-effectiveness of Internet-based cognitive behavior therapy vs. cognitive behavioral group therapy for social anxiety disorder: Results from a randomized controlled trial. *Behaviour Research and Therapy, 49,* 729–736. http://dx.doi.org/10.1016/j.brat.2011.07.009

Hedman, E., Botella, C., & Berger, T. (2016). Internet-based cognitive behavior therapy for social anxiety disorder. In N. Lindefors & G. Andersson (Eds.), *Guided internet-based treatments in psychiatry* (pp. 53–78). Heidelberg, Germany: Springer Nature. http://dx.doi.org/10.1007/978-3-319-06083-5_4

Hedman, E., Furmark, T., Carlbring, P., Ljótsson, B., Rück, C., Lindefors, N., & Andersson, G. (2011). A 5-year follow-up of internet-based cognitive behavior therapy for social anxiety disorder. *Journal of Medical Internet Research, 13*(2), e39. http://dx.doi.org/10.2196/jmir.1776

Hedman, E., Ljótsson, B., & Lindefors, N. (2012). Cognitive behavior therapy via the Internet: A systematic review of applications, clinical efficacy and cost-effectiveness. *Expert Review of Pharmacoeconomics & Outcomes Research, 12,* 745–764. http://dx.doi.org/10.1586/erp.12.67

Hedman, E., Ljótsson, B., Rück, C., Furmark, T., Carlbring, P., Lindefors, N., & Andersson, G. (2010). Internet administration of self-report measures commonly used in research on social anxiety disorder: A psychometric evaluation. *Computers in Human Behavior, 26,* 736–740. http://dx.doi.org/10.1016/j.chb.2010.01.010

Heron, K. E., & Smyth, J. M. (2010). Ecological momentary interventions: Incorporating mobile technology into psychosocial and health behaviour treatments. *British Journal of Health Psychology, 15,* 1–39. http://dx.doi.org/10.1348/135910709X466063

Horowitz, L. M., Rosenberg, S. E., Baer, B. A., Ureño, G., & Villaseñor, V. S. (1988). Inventory of Interpersonal Problems: Psychometric properties and clinical applications. *Journal of Consulting and Clinical Psychology, 56,* 885–892. http://dx.doi.org/10.1037/0022-006X.56.6.885

Horowitz, L. M., Strauß, B., & Kordy, H. (2000). *Inventar zur Erfassung interpersonaler Probleme: Deutsche Version; IIP-D*. Weinheim, Germany: Beltz.

Jacobson, N. S., Follette, W. C., & Revenstorf, D. (1984). Psychotherapy outcome research: Methods for reporting variability and evaluating clinical significance. *Behavior Therapy, 15,* 336–352. http://dx.doi.org/10.1016/S0005-7894(84)80002-7

JASP Team. (2017). JASP (Version 0.8.2) [Computer software]. Retrieved from https://jasp-stats.org/

Kessler, R. C., McGonagle, K. A., Zhao, S., Nelson, C. B., Hughes, M., Eshleman, S., . . . Kendler, K. S. (1994). Lifetime and 12-month prevalence of *DSM-III–R* psychiatric disorders in the United States. Results from the National Comorbidity Survey. *Archives of General Psychiatry, 51,* 8–19. http://dx.doi.org/10.1001/archpsyc.1994.03950010008002

Lee, B. W., & Stapinski, L. A. (2012). Seeking safety on the internet: Relationship between social anxiety and problematic internet use. *Journal of Anxiety Disorders, 26,* 197–205. http://dx.doi.org/10.1016/j.janxdis.2011.11.001

Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association, 83,* 1198–1202. http://dx.doi.org/10.1080/01621459.1988.10478722

Ly, K. H., Asplund, K., & Andersson, G. (2014). Stress management for middle managers via an acceptance and commitment-based smartphone application: A randomized controlled trial. *Internet Interventions, 1,* 95–101. http://dx.doi.org/10.1016/j.invent.2014.06.003

Ly, K. H., Topooco, N., Cederlund, H., Wallin, A., Bergström, J., Molander, O., . . . Andersson, G. (2015). Smartphone-supported versus full

behavioural activation for depression: A randomised controlled trial. *PLoS ONE, 10*(5), e0126559. http://dx.doi.org/10.1371/journal.pone.0126559

Ly, K. H., Trüschel, A., Jarl, L., Magnusson, S., Windahl, T., Johansson, R., . . . Andersson, G. (2014). Behavioural activation versus mindfulness-based guided self-help treatment administered through a smartphone application: A randomised controlled trial. *British Medical Journal Open, 4*(1), e003440. http://dx.doi.org/10.1136/bmjopen-2013-003440

Magee, W. J., Eaton, W. W., Wittchen, H.-U., McGonagle, K. A., & Kessler, R. C. (1996). Agoraphobia, simple phobia, and social phobia in the National Comorbidity Survey. *Archives of General Psychiatry, 53,* 159–168. http://dx.doi.org/10.1001/archpsyc.1996.01830020077009

Mattick, R. P., & Clarke, J. C. (1998). Development and validation of measures of social phobia scrutiny fear and social interaction anxiety. *Behaviour Research and Therapy, 36,* 455–470. http://dx.doi.org/10.1016/S0005-7967(97)10031-6

Mayo-Wilson, E., Dias, S., Mavranezouli, I., Kew, K., Clark, D. M., Ades, A. E., & Pilling, S. (2014). Psychological and pharmacological interventions for social anxiety disorder in adults: A systematic review and network meta-analysis. *The Lancet Psychiatry, 1,* 368–376. http://dx.doi.org/10.1016/S2215-0366(14)70329-3

McFarquhar, T., Luyten, P., & Fonagy, P. (2018). Changes in interpersonal problems in the psychotherapeutic treatment of depression as measured by the Inventory of Interpersonal Problems: A systematic review and meta-analysis. *Journal of Affective Disorders, 226,* 108–123. http://dx.doi.org/10.1016/j.jad.2017.09.036

Mellings, T. M., & Alden, L. E. (2000). Cognitive processes in social anxiety: The effects of self-focus, rumination and anticipatory processing. *Behaviour Research and Therapy, 38,* 243–257. http://dx.doi.org/10.1016/S0005-7967(99)00040-6

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

Olfson, M., Guardino, M., Struening, E., Schneier, F. R., Hellman, F., & Klein, D. F. (2000). Barriers to the treatment of social anxiety. *The American Journal of Psychiatry, 157,* 521–527. http://dx.doi.org/10.1176/appi.ajp.157.4.521

Proudfoot, J., Clarke, J., Birch, M.-R., Whitton, A. E., Parker, G., Manicavasagar, V., . . . Hadzi-Pavlovic, D. (2013). Impact of a mobile phone and web program on symptom and functional outcomes for people with mild-to-moderate depression, anxiety and stress: A randomised controlled trial. *BMC Psychiatry, 13,* 312. http://dx.doi.org/10.1186/1471-244X-13-312

R Core Team. (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rizvi, S. L., Dimeff, L. A., Skutch, J., Carroll, D., & Linehan, M. M. (2011). A pilot study of the DBT coach: An interactive mobile phone application for individuals with borderline personality disorder and substance use disorder. *Behavior Therapy, 42,* 589–600. http://dx.doi.org/10.1016/j.beth.2011.01.003

Saraceno, B., van Ommeren, M., Batniji, R., Cohen, A., Gureje, O., Mahoney, J., . . . Underhill, C. (2007). Barriers to improvement of mental health services in low-income and middle-income countries. *Lancet, 370,* 1164–1174. http://dx.doi.org/10.1016/S0140-6736(07)61263-X

Schulz, A., Stolz, T., Vincent, A., Krieger, T., Andersson, G., & Berger, T. (2016). A sorrow shared is a sorrow halved? A three-arm randomized controlled trial comparing internet-based clinician-guided individual versus group treatment for social anxiety disorder. *Behaviour Research and Therapy, 84,* 14–26. http://dx.doi.org/10.1016/j.brat.2016.07.001

Spek, V., Cuijpers, P., Nyklícek, I., Riper, H., Keyzer, J., & Pop, V. (2007). Internet-based cognitive behaviour therapy for symptoms of depression and anxiety: A meta-analysis. *Psychological Medicine, 37,* 319–328. http://dx.doi.org/10.1017/S0033291706008944

Stangier, U., & Heidenreich, T. (2004). Die Liebowitz Soziale Angst-Skala (LSAS) [The Liebowitz Social Anxiety Scale (LSAS)]. In Collegium Internationale Psychiatriae Scalarum (Ed.), *Internationale skalen für psychiatrie* (pp. 299–306). Göttingen, Germany: Hogrefe-Verlag.

Stangier, U., Heidenreich, T., Berardi, A., Golbs, U., & Hoyer, J. (1999). Die Erfassung Sozialer Phobie durch die Social Interaction Anxiety Scale (SIAS) und die Social Phobia Scale (SPS). *Zeitschrift für Klinische Psychologie, 28,* 28–36. http://dx.doi.org/10.1026//0084-5345.28.1.28

Stangier, U., Heidenreich, T., & Peitz, M. (2009). *Soziale Phobien*. Weinheim, Germany: Beltz.

Stangier, U., Heidenreich, T., Peitz, M., Lauterbach, W., & Clark, D. M. (2003). Cognitive therapy for social phobia: Individual versus group treatment. *Behaviour Research and Therapy, 41,* 991–1007. http://dx.doi.org/10.1016/S0005-7967(02)00176-6

Tracey, T. J., Rounds, J., & Gurtman, M. (1996). Examination of the general factor with the interpersonal circumplex structure: Application to the Inventory of Interpersonal Problems. *Multivariate Behavioral Research, 31,* 441–466. http://dx.doi.org/10.1207/s15327906mbr3104_3

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software, 45,* 1–67.

Wang, P. S., Berglund, P., Olfson, M., Pincus, H. A., Wells, K. B., & Kessler, R. C. (2005). Failure and delay in initial treatment contact after first onset of mental disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry, 62,* 603–613. http://dx.doi.org/10.1001/archpsyc.62.6.603

Watts, S., Mackenzie, A., Thomas, C., Griskaitis, A., Mewton, L., Williams, A., & Andrews, G. (2013). CBT for depression: A pilot RCT comparing mobile phone vs. computer. *BMC Psychiatry, 13,* 49. http://dx.doi.org/10.1186/1471-244X-13-49

Whittaker, R., McRobbie, H., Bullen, C., Rodgers, A., & Gu, Y. (2016). Mobile phone-based interventions for smoking cessation. *Cochrane Database of Systematic Reviews, 4*(CD006611), CD006611.

Wittchen, H.-U., Zaudig, M., & Fydrich, T. (1997). *Skid. Strukturiertes Klinisches Interview für DSM–IV Achse I und II. Handanweisung*. Göttingen, Germany: Hogrefe.

Yuen, E. K., Goetter, E. M., Herbert, J. D., & Forman, E. M. (2012). Challenges and opportunities in internet-mediated telemental health. *Professional Psychology: Research and Practice, 43,* 1–8. http://dx.doi.org/10.1037/a0025524

Received February 14, 2017
Revision received February 9, 2018
Accepted February 14, 2018 ■