



# Manualized cognitive therapy versus cognitive-behavioral treatment-as-usual for social anxiety disorder in routine practice: A cluster-randomized controlled trial



Jürgen Hoyer<sup>a,\*</sup>, Jasmin Čolić<sup>a</sup>, Andre Pittig<sup>a</sup>, Stephen Crawcour<sup>a</sup>, Manuela Moeser<sup>a</sup>, Denise Ginzburg<sup>b</sup>, Jihong Lin<sup>b</sup>, Joerg Wiltink<sup>c</sup>, Eric Leibing<sup>d,1</sup>, Ulrich Stangier<sup>b,1</sup>

<sup>a</sup> Institute of Clinical Psychology and Psychotherapy, Technische Universität Dresden, Germany

<sup>b</sup> Institute of Psychology, Department of Clinical Psychology and Psychotherapy, Goethe University Frankfurt, Germany

<sup>c</sup> Clinic of Psychosomatic Medicine and Psychotherapy, University Medical Center of the Johannes Gutenberg University of Mainz, Germany

<sup>d</sup> Department of Psychosomatic Medicine and Psychotherapy, University Medicine, Georg-August-University Göttingen, Germany

## ARTICLE INFO

### Article history:

Received 11 August 2016

Received in revised form

3 May 2017

Accepted 22 May 2017

Available online 26 May 2017

### Keywords:

Cognitive therapy

Naturalistic psychotherapy

Randomized controlled trial

Social anxiety disorder

Cognitive-behavioral therapy

## ABSTRACT

**Objective:** This study examined the effectiveness of manualized cognitive therapy (mCT) following the Clark-Wells approach versus non-manualized cognitive-behavioral treatment-as-usual (CBTAU) for social anxiety disorder (SAD) in routine practice.

**Methods:** Forty-eight private practitioners were recruited within a multi-center trial and either received training in manualized CT for SAD or no such training. Practitioners treated 162 patients with SAD in routine practice ( $N = 107$  completers,  $n = 57$  for mCT,  $n = 50$  for CBTAU). Social anxiety symptoms (Liebowitz Social Anxiety Scale; LSAS) and secondary measures were assessed before treatment, at treatment-hour 8, 15, and 25, at end of treatment, as well as 6 and 12 months after treatment.

**Results:** Patients in both groups showed significant reductions of SAD severity after treatment ( $d = 1.91$  [mCT] and  $d = 1.80$  [CBTAU], within-group effect sizes, intent-to-treat analyses, LSAS observer ratings), which remained stable at follow-up. There were no differences between groups in terms of symptom reduction and treatment duration.

**Conclusions:** The present trial confirms the high effectiveness of CBTAU and mCT for SAD when practitioners conduct the treatments in routine practice. Additional training in the CT manual did not result in significant between-group effects on therapy outcome. Explanations for this unexpected result are discussed.

© 2017 Elsevier Ltd. All rights reserved.

## 1. Introduction

Social anxiety disorder (SAD) is characterized by persistent fears of one or more social situations, in which the individual is exposed to other people and expects to be scrutinized. The individual is afraid of acting (or showing anxiety symptoms) in a way that could cause embarrassment, humiliation, or rejection (APA, 2013). Estimates of lifetime and 12-month prevalence are high, with 12.1% and 7.1% in the US (Ruscio et al., 2008) and 6.7% (lifetime prevalence) and 2.0% (12-month prevalence) in Europe (Fehm, Beesdo,

Jacobi, & Fiedler, 2008). SAD is characterized by an early onset, a chronic course, severe psychosocial impairments, and high socio-economic costs (Aderka et al., 2012; Egger et al., 2015; Fehm et al., 2008; Kessler, 2003; Stuhldreher et al., 2014). Thus, there is a high demand for effective treatments, particularly in routine practice.

Cognitive-behavioral therapy (CBT) has consistently been shown to be highly efficacious for SAD. A recent meta-analysis (Mayo-Wilson et al., 2014) showed that, among the various types of interventions, CBT based on the Clark-Wells approach (referred to as cognitive therapy, CT; Clark, 2001; Clark & Wells, 1995) yielded the highest effect sizes in randomized controlled trials (Clark et al., 2003, 2006). Furthermore, there is evidence that CT is significantly more effective than standard pharmacological treatment (Clark et al., 2003), interpersonal therapy (Stangier, Schramm, Heidenreich, Berger, & Clark, 2011), manualized short-term

\* Corresponding author. Technische Universität Dresden, Klinische Psychologie und Psychotherapie, Hohe Str. 53, 01187, Dresden, Germany.

E-mail address: [juergen.hoyer@tu-dresden.de](mailto:juergen.hoyer@tu-dresden.de) (J. Hoyer).

<sup>1</sup> Shared senior authorship.

psychodynamic therapy (Leichsenring et al., 2013), and also yielded the highest effect sizes compared to other forms of CBT in recent meta-analyses (Mayo-Wilson et al., 2014). Thus, it may be expected that the dissemination of CT may substantially improve the treatment effectiveness of SAD in routine practice, even in the experienced CBT therapists.

However, the dissemination of evidence-based treatments into routine practice has been identified as a major problem within mental health care delivery (Gillihan, Conklin, & Foa, 2014; Gunter & Whittal, 2010; McHugh & Barlow, 2010). In routine practice, patients often do not receive empirically supported psychological treatment (Shafraan et al., 2009). Barriers to the dissemination of evidence-based psychological therapy include commonly held beliefs among clinicians about the limited applicability of findings from RCTs in clinical practice and lack of knowledge about empirically validated treatments, as well as their proper delivery (Shafraan et al., 2009; Weisz, Ng, & Bearman, 2014). In addition, there is evidence that clinicians apply and slightly modify techniques learned in their initial psychotherapeutic training, while feeling reluctant to adopt new evidence-based techniques (Barlow, 1981; Gunter & Whittal, 2010; Weisz et al., 2014). Thus, even therapists previously trained in CBT may not adopt approaches such as CT, though well validated. Providing additional training for these CBT therapists may increase the effectiveness of treatments into routine practice.

As an answer to these dissemination problems, the Improving Access to Psychological Therapies (IAPT) programme was implemented in the National Health Service in England to increase the availability of evidence-based psychological treatment for depression and anxiety disorders in specialised stepped-care services (Clark, 2011). Preliminary results show that, across depression and anxiety disorders, most patients received therapies recommended by the National Institute of Health and Care Excellence (NICE) Treatment Guidelines, with observed recovery rates at 40.3% and improvement rates at 63.7% (Gyani, Shafraan, Layard, & Clark, 2013). As part of such dissemination programmes training practitioners to use treatment manuals seems to be crucial for the transference of evidence-based treatments into practice (McHugh & Barlow, 2010; Shafraan et al., 2009). Manual training may help to acquire new evidence-based techniques and provide a framework of how to utilize these techniques in routine practice. A manualized treatment may ensure that all key components of the manual are delivered in a temporal sequence, which is typically shorter than treatments based on individualized case formulation. Summarized, the integration of potent treatment techniques in a shorter time period may increase the effectiveness of treatment in routine practice.

For SAD, McEvoy, Nathan, Rapee, and Campbell (2012) demonstrated that a CBT protocol for group treatment that had previously proven efficacious in a research unit could be successfully implemented in a community clinic. Treatment outcome did not differ across the treatment settings (research clinic vs. community clinic) for either completer or intention-to-treat analyses. Despite this notable example, some of the central questions concerning the science-practice transfer of CBT remain unanswered. Most importantly, there are only very few studies investigating the effects of training experienced practitioners in empirically validated manuals. It remains unclear whether practitioners will truly accept and adhere to a standardized manual or rather continue to use their individually preferred interventions (Shafraan et al., 2009; see also Hoefler, Gloster, & Hoyer, 2010). A recent meta-analysis indicated little evidence for a relationship between manual adherence and treatment outcome (Webb, DeRubeis, & Barber, 2010). However, treatments were mostly delivered within randomized controlled trial settings, which aim to minimize non-adherence and thereby

limit the magnitude of process–outcome relations (Webb et al., 2010). The same is true for therapist competence, which has been found to be associated with better treatment outcome of CT in a randomized trial (Ginzburg et al., 2012). It thus remains unclear whether similar results would hold true for experienced CBT therapists in routine practice. In addition, it is unclear whether the replacement of a complex case formulation by a straight-forward manualized therapy will shorten overall treatment duration (Eifert, Schulte, Zvolensky, Lejuez, & Lau, 1997). The findings of smaller effects of CBT for SAD in naturalistic settings than in other anxiety disorders (Hans & Hiller, 2013) additionally emphasize the need for answers to these questions.

The present multicenter randomized clinical trial aimed at implementing the Clark-Wells CT manual (German version: Stangier, Clark, & Ehlers, 2006) into routine practice in Germany. We examined how the CT manual was accepted by certified and well-experienced practitioners and whether training in manualized CT applied in routine practice would augment treatment outcome as compared to cognitive-behavioral treatment-as-usual in terms of symptom reduction and treatment duration. Certified practitioners were recruited and randomized to either receive training in the CT manual or no such training. Subsequently, they treated patients with SAD in their own practice.

## 2. Methods

### 2.1. Study design

The design of this study included a randomized controlled trial in routine practice. Psychotherapeutic practitioners were randomized either to a training in CT for SAD based on the German version of the treatment manual by Clark and Wells (Stangier, Ehlers & Clark, 2006), or to no such additional training. Practitioners in the training group (referred to as manualized CT, mCT) received two separate blocks of training, each block with a mean duration of 12 h. Adherence to the CT manual and the techniques used in each session were documented afterwards.

All patients and therapists provided written informed consent to the procedures approved by both the responsible ethics committee of the Technische Universität Dresden (EK 183062010) and the Data Safety Monitoring Committee (DSMC). The study was monitored by the Coordination Center for Clinical Trials (KKS Dresden), which is independent of the participating research centers, and registered at [ClinicalTrials.gov](http://ClinicalTrials.gov) (Identifier: NCT01388231). More detailed information on the study protocol can be found in [Crawcour et al. \(2012\)](#).

### 2.2. Recruitment and randomization

#### 2.2.1. CBT therapists

CBT therapists were licensed psychotherapists who were trained in and regularly conducted cognitive-behavioral therapy in their private practices. Information about this study was posted on the website of the Chambers of Psychotherapists (in which membership is required for all licensed psychotherapists in Germany). After being informed about the study in discussion rounds, 49 CBT therapists reporting neither having formal training in CT by Clark and Wells nor using such manual in the last five years decided to partake in the study. One therapist had to cancel participation due to a severe negative life event. The remaining 48 CBT therapists were randomly assigned to either the training group (mCT) or the control group, which did not receive additional mCT training (referred to as cognitive-behavioral treatment-as-usual, CBTAU) by KKS Dresden. The private practitioners were block randomized, i.e., randomization into one of the two treatment arms was stratified

for each trial center using the nQueryAdvisor 6.01 software program (Elashoff, 2005).

Mean age of therapists in the mCT group ( $n = 24$ ) was 42.4 years ( $SD = 7.76$ ), and 45.8 years ( $SD = 7.9$ ) in the CBTAU group ( $n = 24$ ). Average general clinical experience was 5.7 years ( $SD = 4.11$ ) for mCT therapists, and 8.9 years ( $SD = 3.95$ ) for CBTAU therapists. With regard to the number of CBT treatments conducted prior to the present trial, 65% (31/48) of therapists indicated more than 100 treatments (50% of mCT therapists, 12/24; 79% of CBTAU therapists, 19/24,  $\chi^2(1) = 4.46$ ,  $p = 0.035$ ), 8% (4/48) of therapists indicated 50–100 treatments (4% of mCT therapists, 1/24; 12% of CBTAU therapists, 3/24, Fisher's exact = 0.61), while 29% (14/48) indicated less than 50 treatments (46% of mCT therapists, 11/24; 13% of CBTAU therapists, 3/24,  $\chi^2(1) = 6.45$ ,  $p = 0.011$ ). There was no significant difference between therapists of both groups regarding age,  $t(46) = 1.50$ ;  $p = 0.14$ , but with regard to clinical experience,  $t(46) = 2.75$ ;  $p = 0.009$ . As a compensation for their contribution, therapists received 200–€ per each documented patient.

### 2.2.2. Patients

Patients were either recruited through the outpatient clinics of the universities of Dresden, Frankfurt and Goettingen or directly by their CBT therapists. All patients were formally informed by the study team about the intentions of the study and about the fact that there were two versions of CBT of potentially equal effectiveness, while remaining blind regarding who among the CBT therapists had been trained and who had not.

The following inclusion criteria were applied: age between 18 and 70 years, a primary diagnosis of social anxiety disorder according to the Diagnostic and Statistical Manual of Mental Disorders (4th ed., text rev.; DSM-IV-TR; APA, 2000; German Version: Saß, Wittchen, Zaudig, & Houben, 2003) determined with the Structured Clinical Interview for DSM-IV (SCID; First, Spitzer, Gibbon, & Williams, 1997; German Version: Wittchen, Wunderlich, Gruschwitz, & Zaudig, 1997) and a total score of 30 or more on the Liebowitz Social Anxiety Scale (Liebowitz, 1987; German Version: Stangier & Heidenreich, 2003). To define the primary diagnoses, we used the clinician's severity rating from the Anxiety Disorders Interview Schedule for DSM-IV (ADIS-IV; Brown, DiNardo, & Barlow, 1994; German Version: Margraf, Schneider, Soeder, Neumer, & Becker, 1996). The primary diagnosis was the diagnosis with the highest degree of interference in daily functioning as compared to other diagnoses. Whether social anxiety disorder represented the primary disorder was determined by use of the respective item of the Anxiety Disorders Interview Schedule for DSM-IV. The exclusion criteria were as follows: psychotic and acute substance-related disorders, cluster A and B personality disorders, prominent risk of self-harm and suicidal behavior, organic mental disorders, severe medical conditions, and concurrent psychotherapeutic or psychopharmacological treatments. Patients on anti-depressive medication were included if the dosage had been stable prior inclusion and kept on a stable dose during treatment.

As participants in a routine practice, patients were not randomized to CBT therapists. Instead, patients selected their preferred therapist without knowing whether or not their chosen CBT therapist had been trained in mCT. A total of 202 patients were recruited for the initial clinical interview. Of these,  $n = 40$  had to be excluded due to diverse reasons (see Fig. 1). Of the remaining  $n = 162$  patients who fulfilled inclusion criteria,  $n = 77$  chose a therapist from the CBTAU group and  $n = 85$  a therapist from the mCT group (for the full flowchart see Fig. 1).

In terms of gender ratio, age, education and comorbidity, there were no significant differences between patients of the two randomized groups (see Table 1).

### 2.3. Manual training

CBT therapists randomized to the mCT group were trained in this treatment approach by the first or last author (US or JH). The training was based on the German version of the Clark and Wells manual (Clark & Wells, 1995), authored by Stangier et al. (2006) and comprised two weekend workshops, with up to 16 training sessions (à 45 min) each. Attendance of therapists was mandatory to be included in the mCT group of the study.

Training workshops were evaluated using a self-developed brief questionnaire including 6-point rating scales (0 = *not at all*, 5 = *completely*), which covered the overall satisfaction with the workshops, as well as the acquisition of additional knowledge about social phobia and its treatment. Feedback was especially positive in terms of high satisfaction with the workshops ( $M = 4.63$ ,  $SD = 0.55$ ) and in terms of positive expectations regarding benefit in therapy ( $M = 4.34$ ,  $SD = 0.86$ ). Workshops were further evaluated as having met expectations ( $M = 4.38$ ,  $SD = 0.75$ ). Therapists also indicated that knowledge about social phobia ( $M = 3.44$ ,  $SD = 1.19$ ) and about the structure of the therapeutic process in the treatment of SAD ( $M = 3.78$ ,  $SD = 1.1$ ) had moderately increased.

In order to maintain treatment fidelity during the trial, mCT therapists received regular site level supervision. At each center, supervision of therapists was performed as a group supervision of 90 min conducted by the first or last author. During the first six months of the trial, supervision was performed fortnightly and monthly thereafter. In order to prevent any biased treatment delivery by CBTAU therapists, no specific supervision or instruction on how supervision should be done was given. Instead, therapists were requested to complete supervision "as usual", i.e., supervision of the CBTAU group was performed as an unrestricted intervision among clinicians.

### 2.4. Treatments

CBT therapists either applied CBT as usual or CT according to Clark and Wells (Clark, 2001; Clark & Wells, 1995). This approach includes several components: a) establishing an idiosyncratic version of the model using the patient's own thoughts, images and behaviors; b) role-play-based behavioral experiments to demonstrate the adverse effects of self-focused attention; c) practicing external focus of attention; d) restructuring distorted self-image by video-feedback; e) behavioral experiments in anxiety-provoking social situations, while dropping safety behaviors and focusing the attention externally; f) identification and modification of anticipatory and post-event processing; and g) identification and modification of dysfunctional assumptions through behavioral experiments and cognitive restructuring.

CBT therapists in the CBTAU group were instructed to treat the patients as usually. The most frequent techniques applied in the CBTAU group were as follows: an individual case formulation, role plays, activation of resources, social competence training, cognitive restructuring, exposure, or relaxation techniques.

In both conditions, up to 25 individual treatment-hours (à 50 min) were possible, which could also be conducted in the form of double- or triple-hour sessions (100–150 min, combining two to three treatment-hours). In addition, up to 5 preparatory sessions (à 50 min) were conducted which are required within the German health care system to cover diagnostic and administrative issues.

### 2.5. Treatment adherence and competence

#### 2.5.1. Session protocols

In order to assess manual adherence, we examined 2055 available session protocols (mCT = 1117, CBTAU = 938) of 88 patients

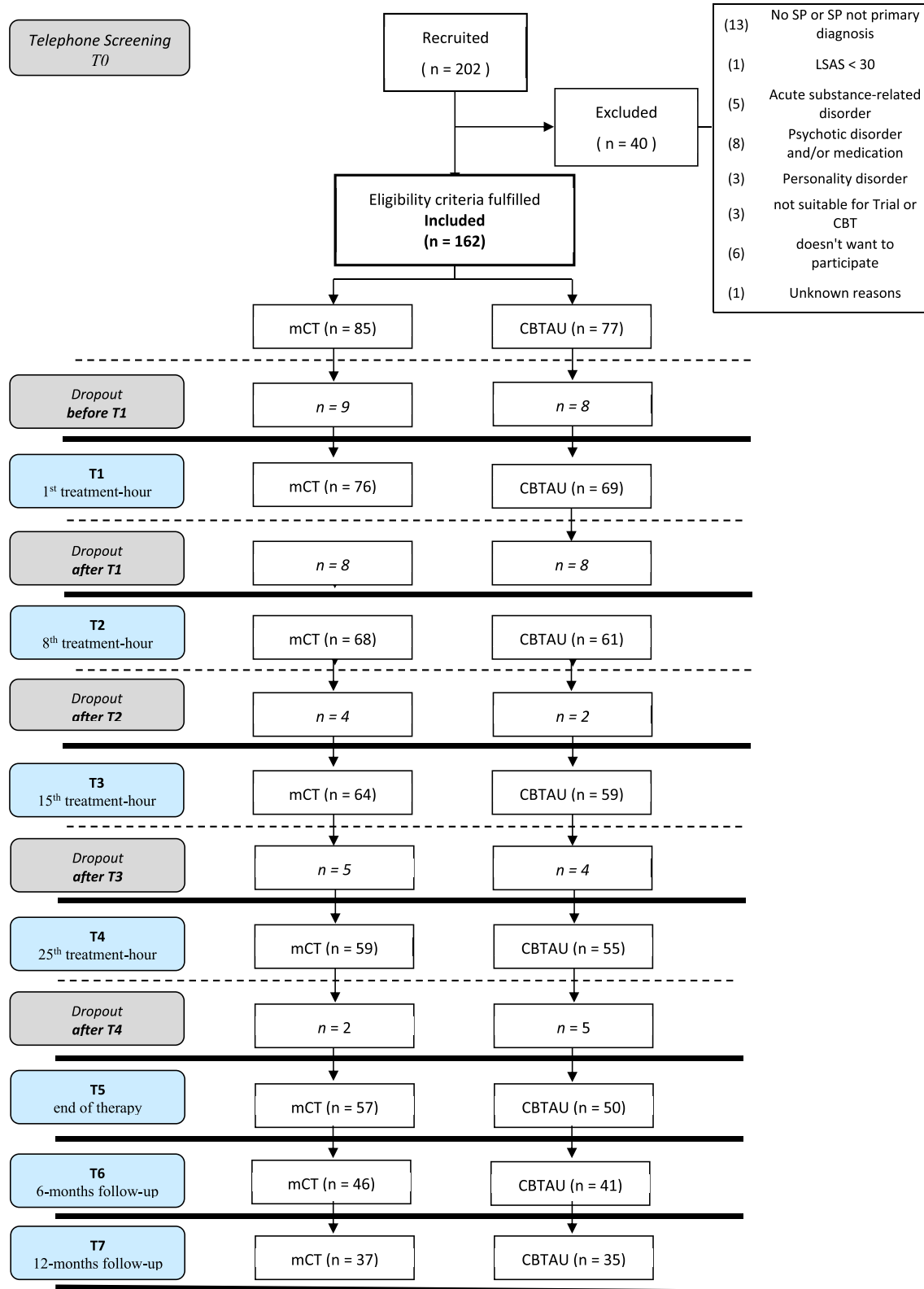


Fig. 1. Flowchart of participants' progress through the study. Please note that more than one treatment-hour (á 50 min) could be combined in a session (e.g., double- or triple-hour sessions of 100–150 min, combining two to three treatment-hours).

who completed their therapies ( $N = 50$  from mCT;  $N = 38$  from CBTAU). For the analysis of adherence, six key components of mCT were derived from the manual by Stangier et al. (2006): 1.) prolonged session duration of at least 75 min in more than 50% of the

sessions; 2.) formulation of an individual cognitive model; 3.) at least one experiment on self-focused attention and safety behaviors; 4.) at least one in-session behavioral experiment to test dysfunctional beliefs; 5.) at least one application of video-feedback;

**Table 1**  
Clinical and sociodemographic characteristics of participants.

| Characteristic               | mCT (N = 85) | CBTAU (N = 77) | Statistic and p-value                  |
|------------------------------|--------------|----------------|--|
| Female (%)                   | 43 (50.6%)   | 39 (50.6%)     | $\chi^2(1) = 0.001, p = 0.99$          |
| Age (SD)                     | 29.21 (6.43) | 31.12 (10.14)  | $t(160) = -1.44, p = 0.15$             |
| Marital Status (%)           |              |                |  |
| Single                       | 70 (82.35%)  | 58 (75.32%)    | $\chi^2(1) = 1.20, p = 0.27$           |
| Married                      | 13 (15.29%)  | 15 (19.48%)    | $\chi^2(1) = 0.49, p = 0.48$           |
| Divorced                     | 2 (2.35%)    | 4 (5.19%)      | $p = 0.42, \text{Fisher's exact test}$ |
| Professional degree (%)      |              |                |  |
| No degree                    | 21 (24.7%)   | 18 (23.4%)     | $\chi^2(1) = 0.11, p = 0.74$           |
| Non-University degree        | 30 (35.3%)   | 32 (41.6%)     | $\chi^2(1) = 0.67, p = 0.41$           |
| University degree            | 33 (38.8%)   | 24 (31.2%)     | $\chi^2(1) = 1.04, p = 0.31$           |
| Other                        | 1 (1.2%)     | 3 (3.8%)       | $p = 0.35, \text{Fisher's exact test}$ |
| Treated earlier (%)          | 37 (43.5%)   | 26 (33.8%)     | $\chi^2(1) = 1.62, p = 0.20$           |
| Comorbid Diagnoses (%)       |              |                |  |
| No comorbidities             | 46 (54.1%)   | 42 (54.5%)     | $\chi^2(1) = 0.003, p = 0.96$          |
| One comorbidity              | 23 (27.1%)   | 16 (20.8%)     | $\chi^2(1) = 0.87, p = 0.35$           |
| Two comorbidities            | 10 (11.8%)   | 13 (16.9%)     | $\chi^2(1) = 0.87, p = 0.35$           |
| Three and more comorbidities | 6 (7%)       | 6 (7.8%)       | $p = 1, \text{Fisher's exact test}$    |
| Personality disorders (%)    | 16 (18.8%)   | 17 (22.1%)     | $\chi^2(1) = 0.26, p = 0.61$           |

Note. mCT = manualized Cognitive therapy, CBTAU = Cognitive-behavioral treatment-as-usual,  $\chi^2$  = chi-square test. The most common comorbid disorders include: avoidant personality disorder (n = 27), major depressive disorder (n = 12), agoraphobia with/without panic disorder (n = 8), recurrent depressive disorder (n = 6), dysthymia (n = 6), specific phobia (n = 6).

6.) at least one application of cognitive restructuring to modify anticipatory and post-event processing. In addition, the following four components were categorized as interventions which are related to CBT in general: 7.) modification of dysfunctional assumptions; 8.) application of guided discovery; 9.) assignment and review of homework; 10.) relapse prevention. Subsequently, we checked for every patient how many of these criteria were followed by the treating therapist. For differences between groups per criterion see Table 2.

2.5.2. Audio recordings

Additionally, treatment adherence and competence were assessed via audio recordings of therapy sessions. Recordings of 41 therapy sessions (n = 18 for mCT, n = 23 for CBTAU) from 13 therapists (n<sub>mCT</sub> = 7, n<sub>CBTAU</sub> = 6) were available. The assessments were carried out by two trained and independent raters, who were blind to the treatment conditions of the treated patients. Both raters were postgraduate students, who were in advanced stages of their clinical training (i.e., at least two years of therapeutic experience and a minimum of two therapies completed under supervision). Adherence was rated on the Cognitive Therapy Adherence Scale for Social Phobia (CTAS-SP; von Consbruch, Heinrich, & Engelhardt, 2008), while competence was rated on the Cognitive Therapy Competence Scale for Social Phobia (CTCS-SP; von

Consbruch, Clark, & Stangier, 2012). The CTAS-SP consists of 17 items, which are rated on a scale from zero (not adherent) to three (very adherent), while the CTCS-SP consists of 16 items, rated on a scale from zero (poor implementation of interventions) and six (excellent implementation of interventions; von Consbruch et al., 2008). The CTCS-SP was further divided into global (e.g., establishing a therapeutic alliance), (manual-)specific (e.g., modification of anticipatory and post-event processing) and additional therapeutic competences (e.g., social skills training).

2.6. Measures

2.6.1. Assessment and masking

After the telephone screening, questionnaires were filled out by patients at eight time points in the study (see Table 3): prior to treatment onset (T0), at treatment-hour one (T1), treatment-hour eight (T2), treatment-hour 15 (T3), treatment-hour 25 (T4), at end of therapy (T5), as well as six months (T6) and 12 months after treatment finalization (T7). For the measurement of social phobia symptoms, German versions of the following measures were used: the Liebowitz Social Anxiety Scale (Liebowitz, 1987; Stangier & Heidenreich, 2003) in its observer-rating version (LSAS-OR) and in its self-rating version (LSAS-SR), a rating of social anxiety severity using a German Version of the Clinical Global Impression

**Table 2**  
Use of interventions as reported by the therapists and tests for differences.

|  | Groups       |      |                |        | Statistic           | p      |
|--|--------------|------|----------------|--------|---------------------|--------|
|  | mCT (n = 50) |      | CBTAU (n = 38) |        |                     |        |
|  | n            | (%)  | n              | (%)    |                     |        |
| prolonged session duration (>75 min) in more than 50% of sessions          | 1            | (2)  | 1              | (2.6)  | Fisher's exact test | 1      |
| formulation of an individual cognitive model                               | 48           | (96) | 5              | (13.2) | Fisher's exact test | <0.001 |
| experiment on self-focused attention and safety behaviors *                | 38           | (76) | 12             | (31.6) | $\chi^2(1) = 5.08$  | 0.02   |
| in-session behavioral experiment to test dysfunctional beliefs *           | 28           | (56) | 12             | (31.6) | $\chi^2(1) = 2.01$  | 0.16   |
| video-feedback *   | 44           | (88) | 15             | (39.5) | $\chi^2(1) = 4.83$  | 0.03   |
| cognitive restructuring to modify anticipatory and post-event processing * | 47           | (94) | 3              | (7.9)  | Fisher's exact test | <0.001 |
| assignment and review of homework  | 32           | (64) | 17             | (44.7) | $\chi^2(1) = 0.94$  | 0.33   |
| guided discovery   | 38           | (76) | 27             | (71.1) | $\chi^2(1) = 0.04$  | 0.84   |
| modification of dysfunctional assumptions                                  | 36           | (72) | 2              | (5.3)  | Fisher's exact test | <0.001 |
| relapse prevention   | 34           | (68) | 9              | (23.7) | $\chi^2(1) = 6.22$  | 0.01   |

Note: \* at least one intervention was reported.

**Table 3**  
Estimated means and standard error of scores on social anxiety, depression and psychopathology measures (Intent-to-treat analysis, N = 162).

| Measure                | Groups       |              | Between-group differences |         |                       |
|------------------------|--------------|--------------|---------------------------|---------|-----------------------|
|                        | mCT          | CBTAU        | z-test                    | p-value | ES [95% CI]           |
|                        | M (SE)       | M (SE)       |                           |         |                       |
| <b>LSAS-OR</b>         |              |              |                           |         |                       |
| Baseline (T0)          | 70.12 (4.27) | 74.51 (5.01) | −1.31                     | 0.19    | 0.21 [−0.10, 0.52]    |
| End of Therapy (T5)    | 29.56 (6.11) | 36.01 (7.31) | −1.33                     | 0.18    | 0.30 [−0.14, 0.75]    |
| FU (12-months; T7)     | 29.11 (5.37) | 35.74 (6.79) | −1.50                     | 0.13    | 0.31 [−0.10, 0.72]    |
| <b>LSAS-SR</b>         |              |              |                           |         |                       |
| Treatment-hour 1 (T1)  | 67.54 (4.89) | 72.32 (6.25) | −1.18                     | 0.24    | 0.20 [−0.13, 0.53]    |
| Treatment-hour 8 (T2)  | 55.77 (5.77) | 59.29 (6.32) | 0.80                      | 0.42    | 0.15 [−0.21, 0.50]    |
| Treatment-hour 15 (T3) | 47.49 (5.85) | 55.21 (6.85) | 1.68                      | 0.09    | 0.33 [−0.05, 0.70]    |
| Treatment-hour 25 (T4) | 37.09 (5.81) | 45.57 (7.21) | 1.77                      | 0.08    | 0.36 [−0.03, 0.74]    |
| FU (6-months; T6)      | 38.34 (7.39) | 41.30 (7.74) | −0.59                     | 0.56    | 0.12 [−0.29, 0.54]    |
| <b>SPAI</b>            |              |              |                           |         |                       |
| Baseline (T0)          | 3.65 (0.22)  | 3.78 (0.22)  | −0.75                     | 0.45    | 0.13 [−0.17, 0.44]    |
| End of Therapy (T5)    | 2.44 (0.34)  | 2.25 (0.32)  | 0.82                      | 0.41    | −0.19 [−0.66, 0.28]   |
| FU (12-months; T7)     | 2.36 (0.35)  | 2.03 (0.33)  | 1.27                      | 0.21    | −0.33 [−0.82, 0.15]   |
| <b>BSPS</b>            |              |              |                           |         |                       |
| Baseline (T0)          | 34.89 (3.32) | 36.96 (3.39) | −0.86                     | 0.39    | 0.14 [−0.18, 0.48]    |
| End of Therapy (T5)    | 19.59 (4.53) | 15.96 (4.22) | 1.15                      | 0.25    | −0.25 [−0.69, 0.18]   |
| <b>SCQ</b>             |              |              |                           |         |                       |
| Treatment-hour 1 (T1)  | 54.54 (4.04) | 55.08 (3.56) | −0.20                     | 0.85    | 0.03 [−0.30, 0.37]    |
| Treatment-hour 15 (T3) | 45.37 (1.98) | 49.03 (2.27) | 1.21                      | 0.22    | 0.23 [−0.13, 0.59]    |
| End of Therapy (T5)    | 40.18 (4.15) | 39.34 (4.22) | 0.28                      | 0.78    | −0.05 [−0.42, 0.31]   |
| FU (12-months; T7)     | 35.77 (4.60) | 45.90 (5.58) | −2.75                     | 0.006** | 0.63** [0.18, 1.08]   |
| <b>BDI</b>             |              |              |                           |         |                       |
| Baseline (T0)          | 13.95 (0.86) | 13.84 (0.91) | 0.09                      | 0.93    | −0.04 [−0.37, 0.28]   |
| Treatment-hour 25 (T4) | 6.55 (1.00)  | 6.59 (0.95)  | −0.03                     | 0.98    | 0.01 [−0.01, −0.84]   |
| End of Therapy (T5)    | 8.21 (1.42)  | 4.88 (0.92)  | 1.97                      | 0.05*   | −0.43* [−0.87, −0.01] |
| FU (6-months; T6)      | 7.57 (1.16)  | 6.02 (0.96)  | 1.03                      | 0.30    | −0.15 [−0.55, 0.24]   |
| FU (12-months; T7)     | 5.59 (1.24)  | 5.27 (1.07)  | 0.20                      | 0.84    | −0.05 [−0.47, 0.37]   |
| <b>CGI</b>             |              |              |                           |         |                       |
| Baseline (T0)          | 4.62 (0.20)  | 4.65 (0.17)  | −0.03                     | 0.97    | 0.02 [−0.30, 0.34]    |
| End of Therapy (T5)    | 2.34 (0.35)  | 2.26 (0.40)  | 0.36                      | 0.72    | −0.07 [−0.57, 0.43]   |
| FU (12-months; T7)     | 2.06 (0.42)  | 2.07 (0.71)  | −0.02                     | 0.98    | 0.01 [−0.76, 0.79]    |

Note. LSAS-OR = Liebowitz Social Anxiety Scale – observer rating, LSAS-SR = Liebowitz Social Anxiety Scale – self rating, SPAI = Social Phobia and Anxiety Inventory, BSPS = Brief Social Phobia Scale, SCQ = Social Cognitions Questionnaire, BDI = Beck Depression Inventory, CGI = Clinical Global Impression, mCT = Manualized cognitive therapy, CBTAU = Cognitive-behavioral treatment-as-usual, ES = Effect Size.

\*p < 0.05 \*\*p < 0.01.

scale (CGI; [Zaider, Heimberg, Fresco, Schneier, & Liebowitz, 2003](#)), the Social Phobia and Anxiety Inventory (SPAI; [Turner, Beidel, Dancu, & Stanley, 1989](#)) in a shortened German 22-item-form ([Fydrich, 2002](#)), the Brief Social Phobia Scale (BSPS; [Davidson et al., 1991](#); German Version: [Chaker, Hausteiner, Hoyer, & Davidson, 2011](#)), and the Social Cognitions Questionnaire (SCQ; [Wells, Stopa, & Clark, 1993](#); German Version: [Stangier, Heidenreich, Ehlers, & Clark, 1996](#)). Depression was measured with the Beck Depression Inventory-II (BDI-II; [Beck, Steer, & Brown, 1996](#); German Version: [Hautzinger, Keller, & Kühner, 2006](#)). For more details on the assessments conducted at each time point please see [Crawcour et al. \(2012\)](#).

Clinical ratings and diagnostic interviews (SCID, LSAS-OR, CGI) were conducted at the three time points (T0, T5, and T7; see [Table 3](#)) by independent and trained interviewers, who were blind to the treatment of the assessed patient. Masking was ensured by instructing patients not to discuss their treatment with the assessors during assessment. All assessors had advanced clinical training and experience in diagnosing and treating mental disorders in university outpatient clinics and received a 2-day workshop on the SCID. Furthermore, personnel involved in handling study documentation and scheduling interviews were required to keep all treatment-specific documents separate from other patient information.

Additionally, prior to treatment therapists completed the 4-item Reaction to Treatment Questionnaire developed by [Holt and Heimberg \(1990\)](#) assessing treatment allegiance for every treated

patient. Items are scaled from 1 (not at all confident or logical) to 10 (very logical or confident), resulting in a total score with a possible range of 1–40. Also, therapists were instructed to complete session protocols after every therapy session. These session protocols were used to derive data on various aspects of treatment duration: In addition to overall treatment duration, which we defined as the mean number of days from treatment onset (T1) to end of treatment (T5), we examined for all remitted patients the mean number of days from treatment start to the assessment point at which remission occurred (i.e., time to remission). Finally, with respect to the specific prescriptions of prolonged sessions in the manual (see section about treatments), we also analyzed session protocols for indicators of treatment infidelity, such as session duration (in minutes) and number of double sessions (i.e., longer sessions of 100 min or more) per patient in each group.

## 2.6.2. Outcomes

The primary outcome measure was the Liebowitz Social Anxiety Scale in the observer-rating version. In addition to the raw reduction of LSAS scores, remission and response rates were analyzed. Remission was defined by an LSAS score  $\leq 30$  ([Bandelow, 2006](#); [Liebowitz, Mangano, Bradwejn, & Asnis, 2005](#)). Response was defined by a 31% reduction (or more) in the LSAS scores from baseline ([Bandelow, Baldwin, Dolberg, Andersen, & Stein, 2006](#)). Additionally, remission was determined on the basis of the SCID conducted after treatment-hour 25 as well, while response was also analyzed using the reliable change index ([Jacobson & Truax, 1991](#)).

Secondary outcomes for this report were the CGI for social anxiety symptom severity, the LSAS in its self-rating version, other measures of social phobia (SPAI, BSPS, SCQ), the BDI-II for depression, as well as the overall treatment duration and time to remission.

### 2.7. Sample size calculation

For the estimation of sample size, we used Campbell et al.'s formula for cluster-randomized controlled trials, which accounts for the intra-cluster correlation coefficient (ICC; Campbell, Elbourne, & Altman, 2004; Campbell, Mollison, & Grimshaw, 2001). The ICC was assumed to be of moderate size (0.15). To detect a between-group difference with medium effect size between mCT and CBTAU ( $d = 0.50$ ,  $\alpha = 0.05$ , power of 0.80),  $n = 58$  patients per group were required. Given an assumed drop-out rate of 25%, a total of  $N = 162$  patients were included.

### 2.8. Statistical analyses

All statistical analyses were carried out using Stata Statistical Software Version 14.1. (StataCorp, 2015). For the analysis of differences in sociodemographic and clinical variables chi-square tests of independence and Fisher's exact tests were used. Primary and secondary continuous outcomes were analyzed using multilevel mixed effects linear regression models (Rabe-Hesketh & Skrondal, 2012). To address selective dropouts and missings<sup>2</sup> according to treatment (mCT and CBTAU) and time (different time points from T0 to T7 across various measures), as well as their combinations, we specified the models as saturated for the combined effects of treatment and time by using dummy variables for the associated main effects and interactions (Wood, White, Hillsdon, & Carpenter, 2005). Hereby, because of non-equidistant assessment points, we modelled time as discrete rather than continuous. The intercept parameter was specified as random, while the other parameters were defined as fixed (random effects for the influences of time and treatment did not yield different results). Because of significant differences in therapists' clinical experience between the groups, this variable was adjusted for as a covariate. Additionally, robust standard errors were calculated (Royall, 1986), to account for non-normality and unequal variances. The effect sizes (Cohen's  $d$ ; Cohen, 1988) for between-group and within-group differences were derived as differences in means divided by the pooled standard deviation at baseline (standardizing the outcome variables accordingly before entering them into the models). According to Cohen (1988), Effect sizes (ES) of  $d = 0.2$  were interpreted as small, ES of  $d = 0.5$  as medium and ES of  $d = 0.8$  as large.

For complete-case analysis of remission and response chi-square tests and Fischer's exact tests were used, while between-group differences in duration measures were estimated via t-tests and nonparametric U-tests. For the associated intent-to-treat analysis, missing data were imputed via multiple imputation by chained equations (MICE; Royston & White, 2011). In our dataset, using the aforementioned criteria we first identified remissions and responses within the available data. Trial site, sociodemographic variables (age, gender, marital status, education, and current life situation), baseline LSAS-scores, as well as some therapist variables (age, sex, and years of experience) were then employed as predictors in the MI-algorithm. Because approximately 30% of data were missing ( $N = 52$  missing values) and based on the suggestion of Graham, Olchowski, and Gilreath (2007), 30 datasets were

imputed. Random effects logistic regression models were used to adjust for therapists' sex and years of experience. Differences between proportions in the intent-to-treat analyses of remission and response were quantified with Cohen's  $h$  (Cohen, 1988).

MICE were also calculated for parameters of treatment duration. Trial site and therapist variables (age, sex, and years of experience) were included as predictors. Thirty datasets were imputed. Based on the combined values of the imputed datasets, differences between groups were estimated using linear regression models.

For the analyses of social phobia outcomes (LSAS-OR, LSAS-SR, SPAI, BSPS, SCQ), a Bonferroni-adjustment of the significance level was applied and alpha was set to 0.0083 ( $\alpha = 0.05/6$ ; Tabachnick & Fidell, 2007). For other outcomes, alpha was set at 0.05. In all analyses two-sided testing was applied.

### 2.9. Additional analyses

As a result of the cluster randomization and due to this study being conducted in routine practice, there was a variance in the number of patients treated by each therapist (ranging from ten to one treated patient per therapist). To account for uncertainty arising from this variance, we conducted different sensitivity analyses. For example, we compared the outcomes of patients who were the fourth or later patients of a therapist with the outcomes of patients who were among the first few in order for their therapists. Also, we allocated different weights to patients depending on the number of patients treated by their therapist and then repeated the analyses for primary outcomes. At last, we calculated Spearman's rho coefficients between the number of patients treated by a therapist and the primary outcomes. Results from all approaches led to identical conclusions as analyses without considering this variance (results available upon request).

## 3. Results

### 3.1. CBT therapists' adherence, competence and allegiance

CBT therapists from the mCT group showed significantly higher adherence to the manual ( $M = 1.22$ ,  $SD = 0.58$ ) than CBT therapists from the CBTAU group ( $M = 0.43$ ,  $SD = 0.29$ ),  $F(1,39) = 32.33$ ,  $p < 0.001$ ,  $\eta = 0.45$ . As for therapeutic competences, there were no significant differences between the groups for global competence,  $F(1,39) = 0.79$ ,  $p = 0.38$ ,  $\eta = 0.02$ . The mCT group exhibited a mean global competence of  $M = 2.68$  ( $SD = 0.59$ ), while the CBT group had a mean rating of  $M = 2.52$  ( $SD = 0.58$ ). The values for specific competences were  $M = 2.72$  ( $SD = 0.77$ ) for the mCT group, and  $M = 1.72$  ( $SD = 0.84$ ) for the CBTAU group. This difference was significant,  $F(1,39) = 15.47$ ,  $p < 0.001$ ,  $\eta = 0.28$ .

CBT therapists of the mCT group ( $n = 62$  patients) reported an average allegiance score of  $M = 33.24$  ( $SD = 4.49$ ), the CBTAU therapists ( $n = 60$  patients) of  $M = 33.4$  ( $SD = 3.82$ ). There was no significant difference in the subjective allegiance to treatment,  $t(120) = 0.21$ ,  $p = 0.834$ .

Since no significant differences in the use of in-session behavioral experiments (isBEs) were found between both groups, possible implications in the mCT group were further analyzed. In 24 treatments (46.2%) no isBE was reported, in 12 treatments (23.1%) one isBE, in 7 treatments (13.5%) two isBE, whereas in three treatments (5.8%) three, four or five isBEs were reported by the therapists. None of the therapists reported the recommended number of six BEs. The correlation between the number of isBEs and the pre-post LSAS difference score was not significant,  $r = 0.12$ ,  $p = 0.60$ . There was a non-significant trend for a negative association between the pretreatment level of LSAS and the number of isBEs,  $r = -0.25$ ,  $p = 0.07$ , which may point to fewer isBEs being

<sup>2</sup> For completers, the percent of missing data for primary and secondary measure were: LSAS-OR = 19.2%, LSAS-SR = 20.3%, SPAI = 24.7%, BSPS = 24.9%, SCQ = 28.4%, BDI = 25.7%, CGI = 28.6%.

conducted in more severe cases.

### 3.2. Drop-out analysis

After accounting for non-starters in each group (mCT = 10,6%, 9/85; CBTAU = 10,4%, 8/77), drop-out rates were 25.0% for mCT (19/76) and 27.5% for CBTAU (16/69). The drop-out rates did not statistically differ between groups (see also Fig. 1). The reasons for drop-out were as follows: withdrawal of consent ( $n = 2$ ), the specification of the wrong diagnosis by the diagnostician ( $n = 1$ ), hospital stay because of illness ( $n = 1$ ), intake of medication ( $n = 2$ ), move to another town ( $n = 2$ ), therapy too time-consuming ( $n = 1$ ), dissatisfaction with therapy ( $n = 1$ ). Twenty-five patients discontinued their therapy for unknown reasons and could not be contacted by their therapists.

Drop-outs across both groups did not differ from completers in regard to sex, education, comorbidity and baseline symptom severity on the LSAS and BDI. However, there was a difference regarding family status between drop-outs ( $n = 49$  single,  $n = 5$  married,  $n = 0$  divorced) and completers ( $n = 78$  single,  $n = 22$  married,  $n = 6$  divorced),  $\chi^2(2) = 7.18, p = 0.02$ . Also, there was a difference in age,  $t(158) = 2.49, p = 0.01$ , as drop-outs ( $M = 27.81$  years,  $SD = 0.94$ ) were significantly younger than completers ( $M = 31.30$  years,  $SD = 0.87$ ).

### 3.3. Primary outcomes

There were no significant differences between the groups at baseline or end of therapy on the primary outcome, the LSAS-OR (see Table 3). Also, there were no between-group differences in slopes from baseline to end of therapy,  $z = 0.41, p = 0.683$ . However, symptoms, as assessed by the LSAS-OR, showed a strong decline from baseline to end of therapy in both groups. Intent-to-treat analyses revealed significant pre-post-differences within both groups with large within-group effect sizes (see Table 4).

### 3.4. Secondary outcomes

Again, there were no between-group differences in the estimated mean scores at any of the assessment points for the LSAS-SR (see Table 3 and Fig. 2). The between-group effect sizes for differences at treatment-hour 15 (T3) and treatment-hour 25 (T4) were

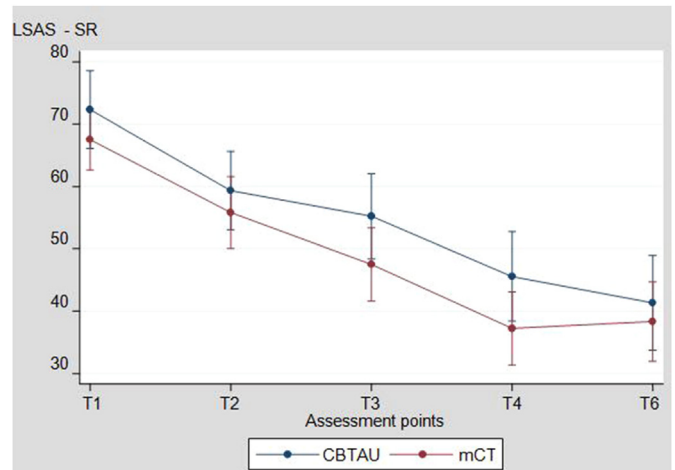


Fig. 2. Scores on the Liebowitz Social Anxiety Scale – self rating across groups and assessment points. T1 = Treatment-hour one; T2 = Treatment-hour eight; T3 = Treatment-hour 15; T4 = Treatment-hour 25; T6 = Six-month-follow-up. Error bars represent confidence intervals of the estimated means.

small to medium (Cohen, 1988; see Table 3; please note that results for T3 and T4 were not considered statistical trends due to Bonferroni correction). Group by time interactions could not be observed either (to treatment-hour 8,  $z = -0.34, p = 0.731$ ; to treatment-hour 15,  $z = 0.69, p = 0.492$ ; to treatment-hour 25,  $z = 0.72, p = 0.472$ ). However, within-group differences were evident in the LSAS self-rating scores (see Fig. 2 and Table 4).

Regarding other self-rating measures, there were no significant between-group differences in any of the time points, neither for the SPAI, nor for the BSPS (Table 3). As shown in Table 4, the estimated means at the end of therapy (on SPAI and BSPS) indicated significant differences from pre-treatment scores in both groups. In regard to the trajectories from baseline, there were no between-group differences for the SPAI (to end of therapy,  $z = -1.23, p = 0.219$ ) or for the BSPS (to end of therapy,  $z = -1.57, p = 0.115$ ). In regard to the SCQ, there were no significant differences between the groups at treatment-hour 1, treatment-hour 15, and end of therapy. Significant reductions in socially anxious cognitions compared to baseline were evident for both groups (Table 4). There

Table 4  
Within-group differences of scores on social anxiety, depression and psychopathology measures (Intent-to-treat analyses, N = 162).

| Measures | Groups | Within-group differences |         |                   |                      |         |                   |
|----------|--------|--------------------------|---------|-------------------|----------------------|---------|-------------------|
|          |        | Baseline - Post          |         |                   | Baseline - Follow-up |         |                   |
|          |        | z-test                   | p-value | ES [95% CI]       | z-test               | p-value | ES [95% CI]       |
| LSAS-OR  | mCT    | 12.70                    | <0.001  | 1.91 [1.62, 2.21] | 14.84                | <0.001  | 1.93 [1.68, 2.19] |
|          | CBTAU  | 9.92                     | <0.001  | 1.82 [1.46, 2.17] | 11.53                | <0.001  | 1.83 [1.52, 2.14] |
| LSAS-SR  | mCT    | 11.59                    | <0.001  | 1.28 [1.06, 1.50] | 10.43                | <0.001  | 1.23 [1.00, 1.47] |
|          | CBTAU  | 6.34                     | <0.001  | 1.13 [0.78, 1.48] | 6.99                 | <0.001  | 1.31 [0.94, 1.68] |
| SPAI     | mCT    | 6.68                     | <0.001  | 1.29 [0.91, 1.67] | 8.82                 | <0.001  | 1.74 [1.34, 2.14] |
|          | CBTAU  | 7.79                     | <0.001  | 1.65 [1.23, 1.97] | 7.55                 | <0.001  | 1.53 [1.13, 1.93] |
| BSPS     | mCT    | 6.06                     | <0.001  | 1.09 [0.73, 1.44] | –                    | –       | –                 |
|          | CBTAU  | 8.11                     | <0.001  | 1.49 [1.13, 1.85] | –                    | –       | –                 |
| SCQ      | mCT    | 5.83                     | <0.001  | 0.89 [0.59, 1.20] | 6.32                 | <0.001  | 1.17 [0.80, 1.53] |
|          | CBTAU  | 5.75                     | <0.001  | 0.98 [0.65, 1.31] | 3.01                 | 0.003   | 0.57 [0.19, 0.94] |
| BDI      | mCT    | 4.30                     | <0.001  | 0.72 [0.39, 1.05] | 8.35                 | <0.001  | 1.05 [0.69, 1.42] |
|          | CBTAU  | 7.41                     | <0.001  | 1.12 [0.82, 1.42] | 8.56                 | <0.001  | 1.07 [0.75, 1.39] |
| CGI      | mCT    | 12.12                    | <0.001  | 2.15 [1.79, 2.51] | 11.51                | <0.001  | 2.39 [1.93, 2.85] |
|          | CBTAU  | 10.37                    | <0.001  | 2.24 [1.80, 2.68] | 6.53                 | <0.001  | 2.44 [1.74, 3.14] |

Note. LSAS-OR = Liebowitz Social Anxiety Scale – observer rating, LSAS-SR = Liebowitz Social Anxiety Scale – self rating, SPAI = Social Phobia and Anxiety Inventory, BSPS = Brief Social Phobia Scale, SCQ = Social Cognitions Questionnaire, BDI = Beck Depression Inventory, CGI = Clinical Global Impression, mCT = Manualized cognitive therapy, CBTAU = Cognitive-behavioral treatment-as-usual, ES = Effect Size.



were no differences between mCT and CBTAU in slopes from baseline to treatment-hour 15,  $z = 1.03$ ,  $p = 0.304$ , and from baseline to end of therapy,  $z = -0.37$ ,  $p = 0.708$ .

On the BDI-II, the groups did not differ significantly from each other at pre-treatment (T0), but there was a significant difference at the end of therapy,  $z = -1.97$ ,  $p = 0.05$ , indicating lower depression scores in the CBTAU group, with a medium between-group effect size of  $d = 0.43$ . In addition, patients who had 25 treatment-hours or more rated their depression at treatment-hour 25 (T4) as well. For patients in both groups, there was a significant symptom reduction from pre-treatment to T4 on the BDI with large within-group effect sizes, while between-group differences were not observed (Tables 3 and 4). There were no group differences in slopes from baseline to treatment-hour 25,  $z = 0.10$ ,  $p = 0.924$ . There was, however, a non-significant trend regarding the slope from baseline to end of therapy,  $z = -1.79$ ,  $p = 0.074$ . Assessor ratings of global social anxiety severity (CGI) showed no differences between the groups at pre-treatment (T0) and at end of therapy (T5). Similar to other measures, there were large within-group differences (Table 4), while no group differences could be observed in slopes from baseline to end of therapy,  $z = -0.36$ ,  $p = 0.717$ .

### 3.5. Follow-up

In terms of primary outcome (LSAS-OR), there was no difference between mean scores in mCT and mean scores in CBTAU achieved at 12-month-follow-up (FU-12; see Table 3), and there was no difference in slopes from baseline to FU-12,  $z = 0.51$ ,  $p = 0.608$ . There were significant differences between baseline and follow-up scores in both groups (as evident in Table 4).

The follow-up measurement on the LSAS-SR was taken six months after T4. The mean score achieved in mCT did not significantly differ from the mean score in the CBTAU group (Table 3). As shown in Table 4, the secondary outcome measures yielded within-group differences of large magnitude. Large differences between the baseline and follow-up scores were apparent in both groups, while there was no group difference in slopes from baseline to six-month-follow-up (FU-6),  $z = -0.34$ ,  $p = 0.732$ . On the SPAI, there was no difference between the groups on the follow-up scores, but these scores significantly differed from baseline in both groups. Also, there was no between-group difference in slopes from baseline to FU-12,  $z = 0.76$ ,  $p = 0.449$ . In regard to the SCQ, there was a difference between follow-up scores achieved in the mCT group and scores in the CBTAU group,  $z = -2.75$ ,  $p = 0.006$ , indicating a medium effect size in favor of mCT ( $d = 0.63$ ; Table 3). Also, there were significant within-group differences and there was a between-group difference in the slope from baseline to 12-month-follow-up as well,  $z = 2.25$ ,  $p = 0.024$ .

Regarding the BDI, at six-month-follow-up and 12-month-follow-up, there were no differences between the mean scores achieved in mCT and the mean scores achieved in CBTAU (Table 3). Scores at both follow-up measures differed significantly from baseline for both groups. However, there were no group differences in slopes from baseline to FU-6,  $z = -0.93$ ,  $p = 0.352$ , and from baseline to FU-12,  $z = -0.11$ ,  $p = 0.912$ . Further, there were no differences between the groups regarding general social anxiety severity (CGI; Table 3) and there was also no between-group difference in slopes from baseline,  $z = -0.09$ ,  $p = 0.927$ . The change in general social anxiety severity (CGI) from baseline to 12-months-follow-up was significant for both groups (Table 4).

### 3.6. Remission and response

Remission rates were 57% after mCT and 54% after CBTAU. The

difference was non-significant,  $OR = 1.10$ , 95% CI [0.61, 1.97],  $h = 0.06$ . The difference between estimated response rates out of the intent-to-treat sample (79% in mCT, 74% in CBTAU) was non-significant as well,  $OR = 1.16$ , 95% CI [0.44, 3.10],  $h = 0.12$ . In the completer sample ( $n_{mCT} = 25$ ,  $n_{CBTAU} = 28$ ), remission rates were 60% ( $n = 15$ , mCT) and 48% ( $n = 14$ , CBTAU), while the response rates were 88% ( $n = 22$ , mCT) and 76% ( $n = 22$ , CBTAU), respectively. There were no between-group differences regarding remission,  $\chi^2(1) = 0.742$ ,  $p = 0.39$ , or response, Fisher's Exact test = 0.31, in the complete-case-analysis.

When response was based on the clinically significant change in LSAS-OR scores as analyzed using the reliable change index (Jacobson & Truax, 1991), response rates also did not differ between groups (mCT = 56%, CBTAU = 62%;  $\chi^2(1) = 0.204$ ,  $p = 0.651$ ). When remission was based on information from the SCID, there was no significant difference between remission rates in mCT (65.12%, 28/43) and remission rates in CBTAU (51.22%, 21/41),  $\chi^2(1) = 1.67$ ,  $p = 0.197$ .

### 3.7. Treatment duration

We hypothesized that mCT would require significantly less days to achieve remission and would have a shorter treatment duration. Tests of these hypotheses based on the complete cases showed no differences between the groups regarding these two variables (see Table 5). Moreover, the groups did not differ in a fully imputed dataset, neither regarding treatment duration,  $M = 342.76$ ,  $SD = 128.67$  in mCT;  $M = 294.24$ ,  $SD = 165.65$  in CBTAU;  $t(105) = 1.64$ ,  $p = 0.107$ , nor in respect to time to remission,  $M = 188.79$ ,  $SD = 191.06$  in mCT;  $M = 199.84$ ,  $SD = 198.22$  in CBTAU;  $t(63) = 0.20$ ,  $p = 0.841$ .

## 4. Discussion

The present study compared treatments of patients with SAD conducted by experienced CBT practitioners trained in the CT manual for SAD (Clark & Wells, 1995) with those of experienced practitioners who did not receive additional mCT training. Acceptance of the CT manual was high as indicated by an overall high satisfaction with the quality and format of the training workshops. However, contrary to our expectations, training in manualized cognitive therapy based on the Clark & Wells model did not significantly improve the effectiveness of treatment. Furthermore, although the effectiveness of manualized treatment for SAD was high, with within-group effect sizes ranging from  $d = 1.29$  (self-report) to 1.91 (clinical observer ratings), CBT therapists without additional mCT training also achieved high within-group effect sizes, with  $d = 1.13$  (self-report) to 1.82 (clinical observer ratings). Thus, both manualized cognitive therapy and non-manualized CBT-as-usual produced significant improvements, with no difference between conditions.

Analysis of the session protocols and audio tapes revealed that adherence to the manual in the mCT group was lower than expected. Although the therapists in this condition reported to apply significantly more often specific interventions from the manual, including formulation of an individual cognitive model, behavioral experiment on self-focused attention and safety behaviors, video-feedback, modification of dysfunctional assumptions, and relapse prevention, in-session behavioral experiments to test dysfunctional beliefs were not implemented in an adequate number. In line with this finding, CBT therapists did not administer a general change of the session format toward prolonged session duration (at least 75 min in more than 50% of the sessions). Obviously, both groups of practitioners remained in their usual schedule of sessions with a duration of less than one hour, which prevented the adequate

**Table 5**  
Means, standard deviations, U-tests and p-values of the interval between sessions, number of treatment-hours, treatment duration and time to remission (N = 88 cases with complete session protocols).

|  | Groups       |          |                |          | U test   | t-test | p     |
|--|--------------|----------|----------------|----------|----------|--------|-------|
|  | mCT (n = 50) |          | CBTAU (n = 38) |          |          |        |       |
|  | M            | (SD)     | M              | (SD)     |          |        |       |
| Number of treatment-hours                  | 24.24        | (7.19)   | 25.7           | (7.8)    | 15251.18 |        | 0.710 |
| Number of double sessions, per participant | 0.66         | (1.33)   | 1.54           | (3.13)   | 899.50   |        | 0.351 |
| Session duration, minutes                  | 56.38        | (4.96)   | 69.11          | (5.89)   | 1725     |        | 0.402 |
| Treatment duration, days                   | 341.27       | (122.35) | 299.66         | (111.12) |          | 1.65   | 0.104 |
| Time to Remission, days                    | 181.28       | (126.66) | 210.29         | (125.90) |          | 0.64   | 0.524 |

Note. mCT = Manualized Cognitive Therapy, CBTAU = Cognitive-behavioral treatment-as-usual.  
\* < 0.05 \*\* < 0.01 \*\*\* < 0.001.

application of behavioral experiments in the mCT condition. Although there was no significant correlation between the use of in-session behavioral experiments and treatment outcome, no clear conclusions can be drawn about the possible implications for the outcome of mCT, due to the *low rates of BEs conducted*. In sum, this may indicate that the practitioners participating in the training transferred some components of the treatment manual into clinical practice, but this was not true for an essential component that requires change of the session format, the use of behavioral experiments.

Although the analysis of audio tapes as well as the clinician-reported session protocols showed that the mCT compared to CBTAU therapists more frequently delivered key components of the CT manual, the mCT therapists adherence to the manual was rather low (1.22 on a scale ranging from 0 to 3). In particular, in-session behavioral experiments to test dysfunctional beliefs were not implemented as specified in the manual for cognitive therapy, which can be explained by the regulations of the German health care system for psychotherapy, favoring short sessions and interfering with prolonged interventions such as exposure and in-session behavioral experiments.

An alternative explanation for the failure to find significant differences between therapist groups may be provided by the fact that treatments in CBTAU achieved unexpected high within-group effect sizes. Although the effect sizes are far lower than in the RCTs of the Clark group (Clark et al., 2003, 2006), they seem to be higher than in the meta-analysis by Stewart and Chambless (2009), who found an within-group effect size of 1.04 for CBT in social anxiety disorder. Furthermore, it may be possible that CBT therapists were equally effective due to a high level of clinical experience and general competence in both groups of CBT therapists, which may have compensated in the CBTAU group for the use of highly effective techniques provided by the manual. This is supported by the analyses of audio tapes which indicate that although independent ratings for specific competence were significantly lower, this was not the case for general competence. Given the high efforts to meet the requirements for the participation in the trial, the practitioners in both conditions did not only show a high level of general competence and clinical experience, but also were highly motivated to meet these requirements. Thus, it can be assumed that another factor explaining the high within-group effect sizes might also be a high level of allegiance, which was documented by high mean scores on the Reaction to Treatment Questionnaire in both groups. Although there are no systematic data on the effects of the therapists' allegiance on outcome (Falkenström, Markowitz, Jonker, Philips, & Holmqvist, 2013), Wampold (2001) found moderate to high effects of this factor on treatment outcome in his meta-analysis.

The study has a number of methodological limitations. First, in accordance with a naturalistic approach in routine practice, we did

not enforce treatment differentiation. Besides the fact that some of the CBT therapists in the mCT condition did not adhere to the use of behavioral experiments, CBT therapists in "usual CBT" were also not proscribed to use specific techniques as proposed by Clark and Wells (1995). In addition, treatment fidelity assessment was mainly based on clinician-reported session protocols, as the number of audio recordings was limited due to low quality and therapists not complying with being recorded. Although therapists deliberately reported not to adhere to some essential components of the mCT manual, suggesting openness to reporting non-adherence, it cannot be ruled out that subjective reports differed from actual procedures. Altogether, there might be a considerable overlap between both conditions, which makes valid conclusions regarding the effectiveness of the implementation of this manualized approach in clinical practice difficult. Second, we cannot discard the possibility of a selection bias regarding the CBT therapists in both conditions. In the present study, levels of age and experience of the therapists were more representative than in previous studies conducted in Germany (e.g., Leichsenring et al., 2013; Stangier et al., 2011). However, it is not clear whether our results can be generalized to psychotherapists in Germany with respect to competence, adherence to evidence-based treatments or allegiance. Third, although the large overall treatment duration is in line with the specific treatment regulations of the German health care system, it remains unclear whether our results can be generalized to other national health care systems favouring brief treatment settings (Shafraan et al., 2009). Fourth, the naturalistic approach in routine practice was more prone to missing data. In addition, drop-out rates were slightly higher compared to averaged drop-out in randomized clinical trials for anxiety disorders in adults (Hofmann & Smits, 2008). However, higher rates ( $\geq 35\%$ ) have been reported in effectiveness studies for SAD (Haug et al., 2016; van Velzen, Emmelkamp & Scholing, 1997). Despite these weaknesses, several methodological strengths of the study should be noted, including a high statistical power, randomization of CBT therapists to treatment conditions, use of independent and blind assessments, and high external validity with respect to practice settings.

In conclusion, this study compared the effectiveness of non-manualized cognitive-behavioral treatment-as-usual to a training in manualized cognitive therapy, which has been shown to be superior to other treatment approaches for social anxiety disorder (Mayo-Wilson et al., 2014). Whereas both treatments showed high within-group effectiveness, the additional training in manualized treatment failed to demonstrate a higher effectiveness. To improve the effectiveness of training methods, the application of standardized role plays and internet resources targeting specific competencies may be of significant benefit (Craske et al., 2011; Fairburn & Cooper, 2011). In addition, and contrary to our expectations, we found evidence that CBT resulted in high success rates in the

treatment of SAD, irrespectively of the training in manualized CT. Further investigation of moderators and mediators of change is needed to identify the factors contributing to the improvement of training methods in the dissemination of effective psychological treatments into clinical practice.

## Acknowledgements

This work (including data collection, analysis, interpretation of the data, and writing of the report) was supported by a grant from the German Federal Ministry of Education and Science (BMBF, 01GV1001) to the first author, and by a grant from the German Research Foundation to the last author (Deutsche Forschungsgemeinschaft, STA 512/11-1). The study was registered at ClinicalTrials.gov, Identifier: NCT01388231. The authors would like to thank all therapists and patients who contributed to the completion of the study.

## References

- Aderka, I. M., Hofmann, S. G., Nickerson, A., Hermesh, H., Gilboa-Schechtman, E., & Marom, S. (2012). Functional impairment in social anxiety disorder. *Journal of Anxiety Disorders*, 26, 393–400. <http://dx.doi.org/10.1016/j.janxdis.2012.01.003>.
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.
- Bandelow, B. (2006). Defining response and remission in anxiety disorders: Toward an integrated approach. *CNS Spectrums*, 11, 21–28. <http://dx.doi.org/10.1017/S1092852900025815>.
- Bandelow, B., Baldwin, D. S., Dolberg, O. T., Andersen, H. F., & Stein, D. J. (2006). What is the threshold for symptomatic response and remission for major depressive disorder, panic disorder, social anxiety disorder, and generalized anxiety disorder? *The Journal of Clinical Psychiatry*, 67, 1428–1434. Retrieved from <http://www.psychiatrist.com/JCP/article/Pages/2006/v67n09/v67n0914.aspx>.
- Barlow, D. H. (1981). On the relation of clinical research to clinical practice: Current issues, new directions. *Journal of Consulting and Clinical Psychology*, 49, 147–155. <http://dx.doi.org/10.1037/0022-006X.49.2.147>.
- Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Beck Depression Inventory-II (BDI-II)*. San Antonio, TX: Psychological Corporation.
- Brown, T. A., DiNardo, P. A., & Barlow, D. H. (1994). *Anxiety Disorders Interview Schedule for DSM-IV (ADIS-IV)*. San Antonio, TX: Psychological Corporation.
- Campbell, M. K., Elbourne, D. R., & Altman, D. G. (2004). CONSORT statement: Extension to cluster randomised trials. *BMJ*, 328, 702–708. <http://dx.doi.org/10.1136/bmj.328.7441.702>.
- Campbell, M. K., Mollison, J., & Grimshaw, J. M. (2001). Cluster trials in implementation research: Estimation of intracluster correlation coefficients and sample size. *Statistics in Medicine*, 20, 391–399. [http://dx.doi.org/10.1002/1097-0258\(20010215\)20:3<391::AID-SIM800>3.0.CO;2-Z](http://dx.doi.org/10.1002/1097-0258(20010215)20:3<391::AID-SIM800>3.0.CO;2-Z).
- Chaker, S., Haustein, E., Hoyer, J., & Davidson, J. R. T. (2011). Ein Interview zur Erfassung sozialer Ängste unter Einbeziehung von Körpersymptomen. *Verhaltenstherapie*, 21, 194–196. <http://dx.doi.org/10.1159/000330926>.
- Clark, D. M. (2001). A cognitive perspective on social phobia. In W. R. Crozier, & L. E. Alden (Eds.), *The essential handbook of social anxiety for clinicians* (pp. 405–430). Chichester, England: John Wiley & Sons.
- Clark, D. M. (2011). Implementing NICE guidelines for the psychological treatment of depression and anxiety disorders: The IAPT experience. *International Review of Psychiatry*, 23, 318–327. <http://dx.doi.org/10.3109/09540261.2011.606803>.
- Clark, D. M., Ehlers, A., Hackmann, A., McManus, F., Fennell, M., Grey, N., ... Wild, J. (2006). Cognitive therapy versus exposure and applied relaxation in social phobia: A randomized controlled trial. *Journal of Consulting and Clinical Psychology*, 74, 568–578. <http://dx.doi.org/10.1037/0022-006X.74.3.568>.
- Clark, D. M., Ehlers, A., McManus, F., Hackmann, A., Fennell, M., Campbell, H., ... Louis, B. (2003). Cognitive therapy versus fluoxetine in generalized social phobia: A randomized placebo-controlled trial. *Journal of Consulting and Clinical Psychology*, 71, 1058–1067. <http://dx.doi.org/10.1037/0022-006X.71.6.1058>.
- Clark, D. M., & Wells, A. (1995). A cognitive model of social phobia. 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: The Guilford Press.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.
- von Consbruch, K., Clark, D. M., & Stangier, U. (2012). Assessing therapeutic competence in cognitive therapy for social phobia: Psychometric properties of the Cognitive Therapy Competence Scale for Social Phobia (CTCS-SP). *Behavioural and Cognitive Psychotherapy*, 40, 149–161. <http://dx.doi.org/10.1017/S1352465811000622>.
- von Consbruch, K., Heinrich, K., & Engelhardt, K. (2008). Psychometrische Eigenschaften der Cognitive Therapy Competence Scale for Social Phobia (CTCS-SP) und der Cognitive Therapy Adherence Scale for Social Phobia (CTAS-SP). In E. Leibing, S. Salzer, & F. Leichsenring (Eds.), *SOPHONET – Forschungsverbund Zur Psychotherapie Der Sozialen Phobie* (pp. 14–21). Göttingen, Germany: Cuvillier Verlag.
- Craske, M. G., Stein, M. B., Sullivan, G., Sherbourne, C., Bystritsky, A., Rose, R. D., ... Roy-Byrne, P. (2011). Disorder-specific impact of coordinated anxiety learning and management treatment for anxiety disorders in primary care. *Archives of General Psychiatry*, 68, 378–388. <http://dx.doi.org/10.1001/archgenpsychiatry.2011.25>.
- Crawcour, S., Leibing, E., Ginzburg, D., Stangier, U., Wiltink, J., & Hoyer, J. (2012). Transfer of manualized CBT for social phobia into clinical practice (SOPHO-PRAX): A study protocol for a cluster-randomized controlled trial. *Trials*, 13, 70. <http://dx.doi.org/10.1186/1745-6215-13-70>.
- Davidson, J. R., Potts, N. L., Richichi, E. A., Ford, S. M., Krishnan, K. R., Smith, R. D., et al. (1991). The Brief Social Phobia Scale. *Journal of Clinical Psychiatry*, 52, 48–51. Retrieved from <http://psycnet.apa.org/psycinfo/1992-11301-001>.
- Egger, N., Konnopka, A., Beutel, M. E., Herpertz, S., Hiller, W., Hoyer, J., ... Wiltink, J. (2015). Short-term cost-effectiveness of psychodynamic therapy and cognitive-behavioral therapy in social anxiety disorder: Results from the SOPHO-NET trial. *Journal of Affective Disorders*, 180, 21–28. <http://dx.doi.org/10.1016/j.jad.2015.03.037>.
- Eifert, G. H., Schulte, D., Zvolensky, M. J., Lejuez, C. W., & Lau, A. W. (1997). Manualized behavior therapy: Merits and challenges. *Behavior Therapy*, 28, 499–509. [http://dx.doi.org/10.1016/S0005-7894\(97\)80005-6](http://dx.doi.org/10.1016/S0005-7894(97)80005-6).
- Elashoff, J. D. (2005). *nQuery Advisor. Version 6.01 User's guide* (Los Angeles, CA).
- Fairburn, C. G., & Cooper, Z. (2011). Therapist competence, therapy quality, and therapist training. *Behaviour Research and Therapy*, 49, 373–378. <http://dx.doi.org/10.1016/j.brat.2011.03.005>.
- Falkenström, F., Markowitz, J. C., Jonker, H., Philips, B., & Holmqvist, R. (2013). Can psychotherapists function as their own controls? Meta-analysis of the “crossed therapist” design in comparative psychotherapy trials. *Journal of Clinical Psychiatry*, 74, 482–491. <http://dx.doi.org/10.4088/JCP.12r07848>.
- Fehm, L., Beesdo, K., Jacobi, F., & Fiedler, A. (2008). Social anxiety disorder above and below the diagnostic threshold: Prevalence, comorbidity and impairment in the general population. *Social Psychiatry and Psychiatric Epidemiology*, 43, 257–265. <http://dx.doi.org/10.1007/s00127-007-0299-4>.
- First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. (1997). *User's guide for the structured clinical interview for DSM-IV axis I disorders SCID-I: Clinician version*. Washington, DC: American Psychiatric Publishing.
- Fydrich, T. (2002). SPAL - Soziale Phobie und Angst Inventar. In E. Brähler, J. Schumacher, & B. Straub (Eds.), *Diagnostische Verfahren in der Psychotherapie* (pp. 335–338). Göttingen, Germany: Hogrefe.
- Gillihan, S. J., Conklin, P., & Foa, E. B. (2014). Dissemination and implementation of evidence-based treatments. In P. Emmelkamp, & T. Ehring (Eds.), *The wiley handbook of anxiety disorders* (pp. 1228–1241). Chichester, United Kingdom: John Wiley & Sons.
- Ginzburg, D. M., Bohn, C., Höfling, V., Weck, F., Clark, D. M., & Stangier, U. (2012). Treatment specific competence predicts outcome in cognitive therapy for social anxiety disorder. *Behaviour Research and Therapy*, 50, 747–752. <http://dx.doi.org/10.1016/j.brat.2012.09.001>.
- Graham, J. W., Olchowski, A. E., & Gilreath, T. D. (2007). How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prevention Science*, 8, 206–213. <http://dx.doi.org/10.1007/s11121-007-0070-9>.
- Gunter, R. W., & Whittal, M. L. (2010). Dissemination of cognitive-behavioral treatments for anxiety disorders: Overcoming barriers and improving patient access. *Clinical Psychology Review*, 30, 194–202. <http://dx.doi.org/10.1016/j.cpr.2009.11.001>.
- Gyani, A., Shafiran, R., Layard, R., & Clark, D. M. (2013). Enhancing recovery rates: Lessons from year one of IAPT. *Behaviour Research and Therapy*, 51, 597–606. <http://dx.doi.org/10.1016/j.brat.2013.06.004>.
- Hans, E., & Hiller, W. (2013). A meta-analysis of nonrandomized effectiveness studies on outpatient cognitive behavioral therapy for adult anxiety disorders. *Clinical Psychology Review*, 33, 954–964. <http://dx.doi.org/10.1016/j.cpr.2013.07.003>.
- Haug, T., Nordgreen, T., Öst, L.-G., Tangen, T., Kvale, G., Hovland, O. J., ... Havik, O. E. (2016). Working alliance and competence as predictors of outcome in cognitive behavioral therapy for social anxiety and panic disorder in adults. *Behaviour Research and Therapy*, 77, 40–51. <http://dx.doi.org/10.1016/j.brat.2015.12.004>.
- Hautzinger, M., Keller, F., & Kühner, C. (2006). *Beck Depressions-inventar (BDI-II)*. Frankfurt am Main, Germany: Harcourt Test Services.
- Hoefer, M., Gloster, A. T., & Hoyer, J. (2010). Causal effects in psychotherapy: Counterfactuals counteract overgeneralization. *Psychotherapy Research*, 20, 668–679. <http://dx.doi.org/10.1080/10503307.2010.501041>.
- Hofmann, S. G., & Smits, J. A. J. (2008). Cognitive-behavioral therapy for adult anxiety disorders: A meta-analysis of randomized placebo-controlled trials. *Journal of Clinical Psychiatry*, 69, 621–632. <http://dx.doi.org/10.4088/JCP.v69n0415>.
- Holt, C. S., & Heimberg, R. G. (1990). The Reaction to Treatment Questionnaire: Measuring treatment credibility and outcome expectancies. *The Behavior Therapist*, 13(213–214), 222. Retrieved from <http://www.abct.org/Journals/?m=mJournal&fa=TB>.
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting*

- and *Clinical Psychology*, 59, 12–19. <http://dx.doi.org/10.1037/0022-006X.59.1.12>.
- Kessler, R. C. (2003). The impairments caused by social phobia in the general population: Implications for intervention. *Acta Psychiatrica Scandinavica*, 108, 19–27. <http://dx.doi.org/10.1034/j.1600-0447.108.s417.2.x>.
- Leichsenring, F., Salzer, S., Beutel, M. E., Herpertz, S., Hiller, W., Hoyer, J., ... Leibing, E. (2013). Psychodynamic therapy and cognitive-behavioral therapy in social anxiety disorder: A multicenter randomized controlled trial. *American Journal of Psychiatry*, 170, 759–767. <http://dx.doi.org/10.1176/appi.ajp.2013.12081125>.
- Liebowitz, M. R. (1987). Social phobia. In D. F. Klein (Ed.), *Modern trends in pharmacopsychiatry* (pp. 141–173). Basel, Switzerland: Karger Publishers.
- Liebowitz, M. R., Mangano, R. M., Bradwejn, J., & Anisn, G. (2005). A randomized controlled trial of venlafaxine extended release in generalized social anxiety disorder. *The Journal of Clinical Psychiatry*, 66, 238–247. Retrieved from <http://www.psychiatrist.com/Pages/home.aspx>.
- Margraf, J., Schneider, S., Soeder, U., Neumer, S., & Becker, E. (1996). *Diagnostisches Interview bei psychischen Störungen—Forschungsversion (F-DIPS)*. Unpublished manuscript. Germany: Technische Universität Dresden.
- Mayo-Wilson, E., Dias, S., Mavranzouli, I., Kew, K., Clark, D. M., Ades, A. E., et al. (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](http://dx.doi.org/10.1016/S2215-0366(14)70329-3).
- McEvoy, P. M., Nathan, P., Rapee, R. M., & Campbell, B. N. (2012). Cognitive behavioural group therapy for social phobia: Evidence of transportability to community clinics. *Behaviour Research and Therapy*, 50, 258–265. <http://dx.doi.org/10.1016/j.brat.2012.01.009>.
- McHugh, R. K., & Barlow, D. H. (2010). The dissemination and implementation of evidence-based psychological treatments: A review of current efforts. *American Psychologist*, 65, 73–84. <http://dx.doi.org/10.1037/a0018121>.
- Rabe-Hesketh, S., & Skrondal, A. (2012). *Multilevel and longitudinal modeling using Stata* (3rd rev. ed.). College Station, TX: Stata Press.
- Royall, R. M. (1986). Model robust confidence intervals using maximum likelihood estimators. *International Statistical Review/Revue Internationale de Statistique*, 221–226. <http://dx.doi.org/10.2307/1403146>.
- Royston, P., & White, I. R. (2011). Multiple imputation by chained equations (MICE): Implementation in Stata. *Journal of Statistical Software*, 45, 1–20.
- Ruscio, A. M., Brown, T. A., Chiu, W. T., Sareen, J., Stein, M. B., & Kessler, R. C. (2008). Social fears and social phobia in the USA: Results from the National Comorbidity Survey Replication. *Psychological Medicine*, 38, 15–28. <http://dx.doi.org/10.1017/S0033291707001699>.
- Saß, H., Wittchen, H.-U., Zaudig, M., & Houben, I. (2003). *Diagnostisches und Statistisches Manual Psychischer Störungen – textrevision – DSM-IV-TR*. Göttingen, Germany: Hogrefe.
- Shafraan, R., Clark, D. M., Fairburn, C. G., Arntz, A., Barlow, D. H., Ehlers, A., ... Wilson, G. T. (2009). Mind the gap: Improving the dissemination of CBT. *Behaviour Research and Therapy*, 47, 902–909. <http://dx.doi.org/10.1016/j.brat.2009.07.003>.
- Stangier, U., Clark, D. M., & Ehlers, A. (2006). *Soziale phobie*. Göttingen, Germany: Hogrefe.
- Stangier, U., & Heidenreich, T. (2003). Die Liebowitz Soziale Angst-Skala (LSAS). In *Collegium Internationale Psychiatriae Scalarum* (Ed.), *Skalen Für Psychiatrie*. Göttingen, Germany: Beltz Test.
- Stangier, U., Heidenreich, T., Ehlers, A., & Clark, D. (1996). *Fragebogen zu sozial-phobischen Kognitionen (SPK)*. Unpublished manuscript. Frankfurt am Main, Germany: Johann-Wolfgang-Goethe-Universität.
- Stangier, U., Schramm, E., Heidenreich, T., Berger, M., & Clark, D. M. (2011). Cognitive therapy vs interpersonal psychotherapy in social anxiety disorder: A randomized controlled trial. *Archives of General Psychiatry*, 68, 692–700. <http://dx.doi.org/10.1001/archgenpsychiatry.2011.67>.
- StataCorp. (2015). *Stata statistical software: Release 14*. College Station, TX: Stata Corp LP.
- Stewart, R. E., & Chambless, D. L. (2009). Cognitive-behavioral therapy for adult anxiety disorders in clinical practice: A meta-analysis of effectiveness studies. *Journal of Consulting and Clinical Psychology*, 77, 595–606. <http://dx.doi.org/10.1037/a0016032>.
- Stuhldreher, N., Leibing, E., Leichsenring, F., Beutel, M. E., Herpertz, S., Hoyer, J., ... König, H. H. (2014). The costs of social anxiety disorder: The role of symptom severity and comorbidities. *Journal of Affective Disorders*, 165, 87–94. <http://dx.doi.org/10.1016/j.jad.2014.04.039>.
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (3rd ed.). Essex, England: Pearson Education Limited.
- Turner, S. M., Beidel, D. C., Dancu, C. V., & Stanley, M. A. (1989). An empirically derived inventory to measure social fears and anxiety: The Social Phobia and Anxiety Inventory. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, 1, 35–40. <http://dx.doi.org/10.1037/1040-3590.1.1.35>.
- van Velzen, C. J., Emmelkamp, P. M., & Scholing, A. (1997). The impact of personality disorders on behavioural treatment outcome for social phobia. *Behaviour Research and Therapy*, 35, 889–900. [http://dx.doi.org/10.1016/S0005-7967\(97\)00052-1](http://dx.doi.org/10.1016/S0005-7967(97)00052-1).
- Wampold, B. E. (2001). *The great psychotherapy debate: Models, methods, and findings*. New York: Lawrence Erlbaum Associates.
- Webb, C. A., DeRubeis, R. J., & Barber, J. P. (2010). Therapist adherence/competence and treatment outcome: A meta-analytic review. *Journal of Consulting and Clinical Psychology*, 78, 200–211. <http://dx.doi.org/10.1037/a0018912>.
- Weisz, J. R., Ng, M. Y., & Bearman, S. K. (2014). Odd couple? Reen visioning the relation between science and practice in the dissemination-implementation era. *Clinical Psychological Science*, 2, 58–74. <http://dx.doi.org/10.1177/2167702613501307>.
- Wells, A., Stopa, L., & Clark, D. M. (1993). *The social cognitions questionnaire*. Unpublished manuscript. UK: Department of Psychiatry, University of Oxford.
- Wittchen, H.-U., Wunderlich, U., Gruschwitz, S., & Zaudig, M. (1997). *SKID I. Strukturiertes Klinisches Interview für DSM-IV. Achse I: Psychische Störungen. Interviewheft und Beurteilungsheft*. Göttingen, Germany: Hogrefe.
- Wood, A. M., White, I. R., Hillsdon, M., & Carpenter, J. (2005). Comparison of imputation and modelling methods in the analysis of a physical activity trial with missing outcomes. *International Journal of Epidemiology*, 34, 89–99. <http://dx.doi.org/10.1093/ije/dyh297>.
- Zaider, T. I., Heimberg, R. G., Fresco, D. M., Schneier, F. R., & Liebowitz, M. R. (2003). Evaluation of the Clinical Global Impression Scale among individuals with social anxiety disorder. *Psychological Medicine*, 33, 611–622. <http://dx.doi.org/10.1017/S0033291703007414>.