Running head: IMAGERY-ENHANCED CBGT FOR SOCIAL ANXIETY DISORDER

Imagery enhancements increase the effectiveness of cognitive behavioural group therapy for social anxiety disorder: a benchmarking study

Peter M. McEvoy[a, b], David M. Erceg-Hurn[a, c], Lisa M. Saulsman[a], & Michel A. Thibodeau [d]

[a] Centre for Clinical Interventions, Perth, Australia

[b] School of Psychology and Speech Pathology, Curtin University, Perth, Australia

[c] School of Psychology, University of Western Australia, Perth, Australia

[d] Psychology Department, University of Regina, Saskatchewan, Canada

Accepted Manuscript, *Behaviour Research and Therapy*

Tables: 3

Figures: 3

Supplementary Tables: 3

Correspondence concerning this article should be addressed to Peter M. McEvoy, Ph.D., School of Psychology and Speech Pathology, GPO Box U1987, 6845, Australia. Phone: +61 8 9266 5110. Fax: +618 9226 2464.
Email: peter.mcevoy@curtin.edu.au

Abstract

Emerging evidence suggests that imagery-based techniques may enhance the effectiveness of traditional verbal-linguistic cognitive interventions for emotional disorders. This study extends an earlier pilot study by reporting outcomes from a naturalistic trial of an imagery-enhanced cognitive behavioural group therapy (IE-CBGT, n = 53) protocol for social anxiety disorder (SAD), and comparing outcomes to historical controls who completed a predominantly verbally-based group protocol (n = 129). Patients were consecutive referrals from health professionals to a community clinic specialising in anxiety and mood disorders. Both treatments involved 12, two-hour group sessions plus a one-month follow-up. Analyses evaluated treatment adherence, predictors of dropout, treatment effect sizes, reliable and clinically significant change, and whether self-reported tendencies to use imagery in everyday life and imagery ability predicted symptom change. IE-CBGT patients were substantially more likely to complete treatment than controls (91% vs. 65%). Effect sizes were very large for both treatments, but were significantly larger for IE-CBGT. A higher proportion of the IE-CBGT patients achieved reliable change, and better imagery ability was associated with larger symptom change. Outcomes compared very favourably to published group and individual treatments for SAD, suggesting that IE-CBGT may be a particularly effective and efficient mode of treatment delivery.

Key Words: imagery; cognitive behaviour therapy; social anxiety disorder; effectiveness; group therapy

Social anxiety disorder (SAD) is characterised by significant and persistent anxiety when exposed to possible scrutiny by others (American Psychiatric Association, APA, 2013). SAD is one of the most common anxiety disorders, has an early age of onset (median 12 years), and can be highly debilitating (Andrews, Henderson, & Hall, 2001; McEvoy, Grove, & Slade, 2011). Cognitive behavioural therapy (CBT) is highly efficacious within research settings (Clark et al., 2003, 2006; Rapee, Gaston, & Abbott, 2009) and effective within community clinics (McEvoy, Nathan, Rapee, & Campbell, 2012). However, a significant minority of patients remain symptomatic after CBT (McEvoy et al., 2012), so further treatment innovations are required. A recent pilot of an imagery-enhanced group CBT protocol (IE-CBGT) found high attendance rates and very large effect sizes (McEvoy & Saulsman, 2014). The aims of this study were to extend these pilot data by including a larger clinical sample, examining weekly trajectories of change, and examining whether general imagery use and imagery ability moderate outcomes from IE-CBGT.

Cognitive theories of emotion (Holmes & Mathews, 2010) and SAD in particular (Heimberg, Brozovich, & Rapee, 2014; Ng, Abbott, & Hunt, 2014) suggest that negative imagery contributes to the maintenance of emotional disorders. Imagery has been defined as multisensory-perceptual representations that may have visual, somatic, auditory, olfactory, and/or gustatory elements, and which have particularly strong links to both positive and negative emotions (Holmes & Mathews, 2010). Rapee and Heimberg's (1997) cognitive behavioural model suggests that individuals with SAD construct a mental representation of the self as seen by others (i.e., the observer perspective), which is guided by a "pre-existing image, stored in long-term memory and based on feedback from others, actual images of the self (e.g., from mirrors, photographs, etc.), and prior experiences in a given situation. (p. 744)." An individual's mental representation of the self is also guided by preferential attention on both internal experiences (e.g., physical sensations of blushing) and perceived external indicators of evaluation (e.g., others' non-verbal and verbal behaviour). The model

argues that individuals with SAD attempt to formulate the audience's performance standard and then determine whether or not this standard is being met, with any discrepancy being used to guide the perceived likelihood and costs of evaluation. The anticipation of evaluation then results in a range of physiological, cognitive, and behavioural effects that reinforce the individual's negative mental representation of the self (Heimberg et al., 2014). Rapee and Heimberg (1997) argue that the process is the same whether the situation is being experienced, anticipated or reflected on. Therefore, images about the self and the consequences of evaluation can be present before, during, and after social situations.

Research has shown that negative imagery is common in high socially anxious individuals (Moscovitch, Garvric, Merrifield, Bielak, & Moscovitch, 2011), ubiquitous in individuals with SAD (Hackmann, Clark, & McManus, 2000), and features in both anticipatory and post-event processing in relation to social stressors (Chiupka, Moscovitch, & Bielak, 2012). Importantly, these negative social images reflect the individual's feared outcomes rather than being reality based, and they serve to reinforce negative self-appraisals and expectations of negative evaluation from others (Hackmann et al., 2000). Experimental studies have found that negative imagery exacerbates anxiety, increases the use of avoidant behaviours, increases self-focused attention, results in more negative self-appraisals, and interferes with social performance (Hirsch, Clark, Mathews, & Williams, 2003; Hirsch, Meynen, & Clark, 2004). These findings suggest that negative imagery is an important treatment target for SAD.

There is also evidence that imagery is a more potent facilitator of cognitive and affective change more generally compared to verbal-linguistic activity (Holmes, Lang, & Shah, 2009; Holmes & Mathews, 2010). Cognitive interventions within the imagery mode may therefore potentiate greater affective shifts in treatment compared to verbal techniques. McEvoy and Saulsman's (2014) pilot study (N = 19) found an imagery-enhanced cognitive behavioural group therapy protocol (IE-CBGT) for SAD to be associated with high retention

(95%), large effect sizes, and a high proportion of patients achieving reliable improvement. The main components in the IE-CBGT protocol were based on those in existing efficacious and effective group (Rapee et al., 2009) and individual (Clark et al., 2003) treatments, with adaptations to ensure each component was delivered in imagery mode. For example, imagery was used prior to cognitive restructuring and behavioural experiments to elicit specific beliefs, and afterwards to envisage more realistic conclusions. Metaphorical ('coping') imagery was developed to assist patients with tolerating anxiety during behavioural experiments, and positive imagery was used to develop and road-test new core beliefs. Video-feedback was used to modify negative self-images (Harvey, Clark, Ehlers, & Rapee, 2000), and past imagery rescripting was used to modify negative core beliefs.

Past imagery rescripting involves revisiting and modifying recurrent memories, associated images, and meanings of past traumas within imagery to alter the encapsulated meanings of the original event. Several small clinical trials with SAD patients have found that imagery rescripting alone is associated with significant improvements in negative social beliefs, the vividness and distress of negative images and early memories, fear of negative evaluation, and social anxiety symptoms (e.g., Nilsson, Lundh, & Viborg, 2012; Wild & Clark, 2011). To our knowledge, other than the pilot study (McEvoy & Saulsman, 2014), past imagery rescripting has not previously been evaluated within a group format.

Identifying treatment moderators can aid clinical decision-making. It is plausible that individuals who naturally tend to operate within an imagery mode in their day-to-day life, and who have a greater capacity to elicit vivid imagery, would benefit more from an imagery-enhanced treatment. However, a consistent relationship between imagery ability and treatment outcomes has not been reported in the literature. Hunt and Fenton (2007) found that whilst imagery ability was associated with avoidance during an imagery induction procedure, it was unrelated to the efficacy of imagery rescripting for snake phobias. The authors note that their measure of imagery ability was suboptimal in terms of

non-standard administration and only marginally acceptable internal consistency. In a small sample with SAD (N = 23), Lee and Kwon (2013) also failed to find a significant association between mental imagery ability and outcomes from imagery rescripting. Although poor imagery ability may not necessarily be an impediment to benefiting from imagery-based techniques, it is plausible that a comprehensive treatment protocol emphasising imagery within each component would be more beneficial for individuals who are able to evoke more vivid images.

The first aim of this naturalistic benchmarking study was to evaluate an IE-CBGT protocol for SAD. The current study extends an earlier pilot (McEvoy & Saulsman, 2014) by including larger IE-CBGT (N = 53, henceforth referred to as the 'imagery-enhanced') and historical control (N = 129) samples, and by more comprehensively assessing treatment completion and symptom change. Imagery-enhanced outcomes were compared to historical controls who attended a predominantly verbally-based CBGT protocol in three ways: (a) the proportion of patients completing treatment and mean number of sessions attended, (b) trajectories of weekly change and effect sizes, and (c) the proportion of patients achieving reliable and clinically significant change. It was hypothesised that the imagery-enhanced group would demonstrate higher attendance rates, more rapid change, larger effect sizes, and higher rates of reliable and clinically significant change, compared to the historical controls.

The second aim was to examine whether (a) self-reported tendencies to operate within an imagery mode and (b) imagery ability were associated with outcomes in the imagery-enhanced group. It was hypothesised that individuals with a stronger natural tendency to operate within the visual mode and with greater ability to evoke vivid images would benefit more from imagery-enhanced treatment.

Method

Participants

Inclusion criteria were (a) a Diagnostic and Statistical Manual of Mental Disorders (DSM-IV; APA, 1994) SAD diagnosis, (b) no current suicidal intent, (c) no psychotic illness, and (d) a level of substance use judged by the assessing clinician as unlikely to significantly interfere with engagement in treatment. The Mini International Neuropsychiatric Interview (MINI PLUS 5.0; Sheehan et al., 2001) was administered by masters- or doctorate-level clinical psychologists to establish Axis I disorders. The MINI has good validity and converges with other structured interviews (e.g., Sheehan et al., 1997). A maximum of three diagnoses were coded in the database. Patients and assessing clinicians made a collaborative decision for the patient to attend the social anxiety program if SAD was the most debilitating problem. Written informed consent was provided from all patients for de-identified data to be used for evaluation purposes and approval for this study was received from the health service's Human Research Ethics Committee (QI 2014_05).

Imagery-Enhanced Group. Participants comprised 53 consecutive referrals by health professionals (general medical practitioners, psychiatrists, psychologists) in 2013 and 2014 with a diagnosis of SAD to a specialist community mental health clinic. Demographic and clinical characteristics of the sample are presented in Table 1. Most (74%) of the patients had at least one comorbid disorder, and 34% had at least two additional disorders. The most common comorbid disorders were major depression (n = 22), generalised anxiety disorder (n = 16), and dysthymia (n = 5). More than half the patients were unemployed and most were not in a relationship. About one-fifth of the sample had a past history of self-harm, suicide attempts, and psychiatric hospitalization, respectively. Most patients were born in Australia or New Zealand (n=41), with the remainder from Europe/United Kingdom (n = 6), Asia (n = 5), North America (n = 1).

Historical Control Group. Historical controls comprised of 129 participants referred to the same community clinic from September 2007 to September 2012. The demographic and clinical characteristics of historical control sample were similar to those of

the imagery-enhanced group (see Table 1). The most common comorbid disorders were major depression (n = 63), generalised anxiety disorder (n = 38), and dysthymia (n = 18). Most participants were born in Australia or New Zealand (n=101), with the remainder from Europe/United Kingdom (n=14), Asia (n = 6), South America (n = 1), and Other (n = 7).

Outcome Measures

Social Phobia Scale (SPS) & Social Interaction Anxiety Scale (SIAS). The SPS and SIAS (Mattick & Clarke, 1998) are 20-item measures of performance and interaction anxiety, respectively. The SPS describes situations in which the person is the focus of attention and observed by others, such as eating, drinking, and writing. The SIAS contains items reflecting cognitive, affective, and behavioural reactions to interaction situations, such as nervousness when speaking to authority or mixing with people. The 5-point response scale for both scales is *Not at all*, *Slightly*, *Moderately*, *Very*, or *Extremely* characteristic of me. These scales have demonstrated high 12-week test-retest reliabilities (SIAS r = .92; SPS r = .93, Mattick & Clarke, 1998) and sensitivity to change (Cox, Ross, Swinson, & Direnfeld, 1998). Internal consistencies were high for the SIAS (αs > .81) and the SPS (αs > .92) across all time points in the current sample.

Brief Fear of Negative Evaluation-Straightforwardly Worded (BFNE-S, Rodebaugh et al., 2004**).** The BFNE-S is an 8-item self-report measure of fear of negative evaluation that excludes four negatively worded (i.e., reverse-scored) items from the original 12-item version (Leary, 1983). The eight straight-forwardly worded items (BFNE-S) have demonstrated superior psychometric properties and less bias associated with education level than the 12-item version, and has excellent internal consistency, factorial validity and construct validity in analogue and clinical samples (Rodebaugh et al., 2004; Weeks et al., 2005). Respondents indicate how characteristic each statement is of them on a 5-point response scale, *not at all, slightly, moderately, very,* or *extremely* characteristic of me. Excellent internal consistency was demonstrated in the current sample (αs > .84).

Depression, Anxiety and Stress Scale (DASS-21). The DASS-21 (Lovibond & Lovibond, 1995) is a shorter version of the original 42-item version. The DASS-21 has excellent psychometric properties in psychiatric settings (Page, Hooke, & Morrison, 2007), including a robust factor structure, internal consistency, and sensitivity to change during treatment. Respondents indicate the degree to which each item applies to them: 0 = *did not apply to me at all*, 1 = *applied to me to some degree or some of the time*, 2 = *applied to me to a considerable degree or a good part of the time*, and 4 = *applied to me very much or most of the time*. Internal consistency was very high in the current sample (αs > .91).

Spontaneous Use of Imagery Questionnaire (SUIS). The SUIS (Reisberg, Pearson, & Kosslyn, 2003) is a 12-item measure of habitual imagery use. An example item is "When going to a new place, I prefer directions that include detailed descriptions of landmarks (such as the size, shape, and colour of a petrol station) in addition to their names." Respondents indicate the degree to which each item is appropriate for them using a 5-point scale: 5 = *completely appropriate*, 3 = *appropriate about half of the time*, and 1 = *never appropriate*. Internal consistency was very high in the current sample (α = .84).

Vividness of Visual Imagery Questionnaire (VVIQ). Consistent with previous literature (e.g., Reisberg et al., 2003), we used an abbreviated 4-item version of the 16-item VVIQ (Marks, 1977). The VVIQ has demonstrated acceptable split-half and internal consistency (McKelvie, 1986) and is unifactorial (Richardson, 1999). Instructions were, "For the visualisation task below, consider carefully the picture that comes before your mind's eye." The VVIQ asks respondents to visualise a rising sun and then rate how clearly and vividly they imagined: (a) the sun is rising above the horizon into a hazy sky, (b) the sky clears and surrounds the sun with blueness, (c) clouds, a storm blows up, with flashes of lightning, and (d) a rainbow appears. The items are rated on a 5-point scale: 1 = *Perfectly clear and as vivid as normal vision*, 2 = *clear and reasonably vivid*, 3 = *moderately clear*

and vivid, 4 = *vague and dim*, 5 = *no image at all, you only 'know' that you are thinking of the object*. In the current study, internal consistency was excellent (α = .89).

Procedure and Treatment

Both treatments comprised 12 weekly, 2-hour sessions plus a one-month follow-up. The SIAS and SPS were administered at pre-treatment, post-treatment, and at one-month follow-up. The BFNE-S and DASS were administered prior to each treatment session. The SUIS and VVIQ were administered at pre-treatment for the imagery-enhanced group only. Treatment integrity was encouraged by the use of a detailed treatment manual with therapist instructions, patient handouts, and worksheets. All groups were co-facilitated by two masters- or doctoral-level clinical psychologists or a clinical psychologist and intern. Groups comprised of between 3 and 12 participants (M = 7.0, SD = 2.5).

The imagery-enhanced protocol was modified from a manual demonstrated to be efficacious (Rapee et al., 2009) and effective (McEvoy et al., 2012). The modifications are detailed in McEvoy and Saulsman (2014). Many of the imagery-based strategies were modified for use in a group setting from those described in Hackmann, Bennet-Levy, and Holmes (2011). The protocol was designed to target six main mechanisms: negative social images, avoidance, safety behaviours, negative self-images, self-focused attention, and negative core beliefs. Session 1 involved socialising patients to the cognitive behavioural model of SAD and presenting the rationale for working with negative past, present, and future social *images* rather than negative thoughts. Patients were encouraged to transform negative thoughts into images to encourage specificity and because evidence suggests that imagery has stronger associations with emotions than verbal-linguistic thoughts. Session 2 introduced imagery challenging and Session 3 introduced coping imagery and provided a rationale for using behavioural experiments to target avoidance. Session 4 involved a within-session group behavioural experiment (i.e., the group walking down the street in a straight line) after which patients began developing individualised behavioural experiment

hierarchies. Session 5 involved psychoeducation about the importance of dropping safety behaviours followed by a behavioural experiment with and without safety behaviours. Session 6 involved video-feedback from a spontaneous speech task, Session 7 involved within-session behavioural experiments (i.e., shame-attacking), and Session 8 involved attention training and focusing. Session 9 was the imagery rescripting session and Session 10 involved a series of within-session individual shame-attacking behavioural experiments. Session 11 focused on the development of new core beliefs using positive imagery which formed the basis for developing future action plans. Session 12 involved a review of treatment components, relapse prevention, and a future-oriented imagery exercise. All sessions involved a homework review, new content including in-session skills practice, summary of three take home messages, and homework for the following week. The one-month follow-up session involved a review of progress, treatment principles, relapse prevention plans, and future goal-setting.

The historical control protocol targeted the same six mechanisms using predominantly verbal-linguistic methods (except negative social *thoughts* rather than images). Imagery was not mentioned during any of the treatment components with the exception of the video-feedback session. All other components were completed within the verbal mode (see McEvoy & Saulsman, 2014, and Rapee et al., 2009). Idiosyncratic cognitive content (imagery or verbal thoughts) for all strategies across both treatments was selected by patients based on their own individual experiences, although the process of working on the content was structured within the manualised tasks.

Data analysis

All analyses were conducted using the statistical software *R* version 3.1.1 (R Development Core Team, 2014). The imagery-enhanced and historical control groups were compared on a range of pre-treatment clinical and demographic variables using Chi-square tests with the Wilson confidence interval (Newcombe, 2013, binary variables), or

independent samples t-tests and confidence intervals for mean differences using the Welch-Satterthwaite approach (Cumming, 2012, continuous variables). Treatment acceptability was compared in terms of: (i) the mean number of sessions completed, and (ii) the proportion of patients who received a high dose of treatment (≥ 9 sessions). Pearson correlations were used to explore whether the dropout was associated with clinical or demographic variables.

Mixed-effects model repeated measures (MMRM) analyses were used to compare treatments on the SIAS, SPS, BFNE-S and DASS. MMRM or multiple imputation (MI) were used to correct for potential biases caused by missing data (National Research Council Panel on Handling Missing Data in Clinical Trials, 2010). These methods estimate how the symptoms of patients who dropped out would have changed had they stayed in treatment using an intent-to-treat approach. In the current paper, MMRM was used to handle missing data for the primary analyses of treatment efficacy and for computing effect sizes. For other analyses where it was not feasible to use MMRM, such as baseline comparisons of binary or ordinal variables, MI was used. One hundred imputed datasets were generated using the random-forest multiple imputation by chained equations algorithm (Shah, Bartlett, Carpenter, Nicholas, & Hemingway, 2014). Analyses of the 100 imputed datasets were combined using Rubin's rules (Van Buuren, 2012). MI analyses were conducted using the R packages mice (Van Buuren, 2012) and CALIBERrfimpute (Shah, Bartlett, Hemingway, Nicholas, & Hingorani, 2014).

Primary analyses tested whether changes from (i) pre-treatment to post-treatment, and (ii) pre-treatment to follow up, were greater for imagery-enhanced than the control treatment. Secondary analyses explored trajectories of change across each treatment session on the BFNE-S and DASS. MMRM analyses were conducted using R's *nlme* (Pinherio & Bates, 2000) and *lsmeans* (Lenth, 2014) packages. All MMRM analyses were intent-to-treat, two-tailed, and an unstructured (co)variance matrix was used to model the within-subject errors. Within-treatment effect sizes (standardised mean change scores) were calculated for

each intervention using the following formula: $d = (M_{pre} - M_{post\ or\ fu}) / SD$. Between-treatment effect sizes were calculated by first computing the mean change from pre-to-post (or follow-up) treatment for each intervention, and then using the formula: d = (Mean Change $_{IECBGT}$ – Mean Change $_{CONTROL}$) / SD. The standardiser for each effect size was the pre-treatment standard deviation, pooled across the two treatments (Morris, 2008).

Reliable change (RC) criteria were identical to those reported in McEvoy et al. (2012) and McEvoy and Saulsman (2014), which were based on Jacobson and Truax's (1991) method. The magnitude of change required to achieve RC on the SIAS and SPS were 8.84 and 10.66, respectively. The cutoff for achieving Clinically Significant Change (CSC) was defined as the mid-point between the means of clinical and normative samples from Carleton et al. (2014), which corresponded to 40.56 for the SIAS and 31.61 for the SPS. To be classified as having achieved CSC an individual must have scored above the CSC cutoff before treatment, achieved RC, and scored below the CSC cutoff after treatment. Individuals who scored below the CSC cutoff at pre-treatment were excluded from these analyses. Pearson correlations were used to examine whether pre-treatment VVIQ and SUIS scores were associated with symptom change for the imagery-enhanced treatment.

For CSC analyses our goal was to determine the proportion of patients in each group who definitely achieved CSC. Patients who discontinue treatment, particularly early in treatment, are unlikely to achieve CSC. Using imputed data would therefore overestimate the likelihood of patients meeting CSC criteria, and disproportiately so for the control group in this study where a substantially larger proportion of patients discontinued earlier in treatment. Therefore, CSC analyses were based on all available observed data at post-treatment and follow up.

Results

Baseline Characteristics

Pre-treatment demographic and clinical characteristics for the imagery-enhanced and historical control groups are compared in Table 1. The majority of differences were small and not statistically significant, including age, educational achievement, employment status, and clinical features such as the proportion of patients with a comorbid diagnosis. There were no pre-treatment differences on the SIAS, SPS, or BFNE-S (see Table 2). There was a small but statistically significant difference on the DASS, with imagery-enhanced patients being slightly more severe. Overall, the findings suggest that the composition of the imagery-enhanced and control samples was similar at pre-treatment.

Adherence to Treatment

Adherence to treatment in the imagery-enhanced group was high. The mean number of attended sessions was 10.9 (median = 12), and nearly all patients (n = 48, 91%) completed a high dose of treatment. One patient discontinued after one session, and another after two sessions, due to scheduling clashes with their University timetables. Both patients were participating in subsequent groups at the time of manuscript preparation. The three other patients who dropped out completed between 5 and 8 sessions. One patient discontinued due to a life crisis unrelated to treatment, one due to difficulties with CBT, and no reason was recorded for the other patient.

Adherence to treatment was poorer in the control group. The mean number of sessions completed was 9.1 (median = 10) and only 65% received a high dose. Imagery-enhanced patients completed a mean of 1.8 sessions more than the controls, 95% CI [0.8, 2.7], $p < .001$), and substantially more imagery-enhanced patients received a high dose of treatment (difference = 24.7%, CI [11.4, 34.7], $p = .001$).

The strongest demographic and clinical predictors (listed in Tables 1 and 2) of attendance were variables related to depression. These included having a diagnosis of major depression or dysthymia ($r = -.24$, $p = .001$), the number of previous suicide attempts ($r = -.24$, $p = .002$), prior psychiatric hospitalization ($r = -.21$, $p = .006$), the number of previous

self-harm attempts (r = .19, p = .012), number of comorbid diagnoses (r = -.17, p = .023), and having a tertiary education (r = .17, p = .024). Other variables were only weakly related, or not related at all, to the number of sessions completed (rs < |.16|, all ps > .05).

There was evidence of an interaction between depression and treatment type, as can be seen in Figure 1. A diagnosis of depression was unrelated to adherence amongst imagery-enhanced patients, with the mean number of sessions attended being practically identical for those with (M = 10.7) and without (M = 11.3) depression. In contrast, control patients with depression (M = 8.3) completed, on average, two less sessions than those without depression (M = 10.3). Depressed patients in the imagery-enhanced group completed 2.4 more sessions than those in the control group, 95% CI [1.0, 3.8], p = .001. Nearly all (88%) of the depressed imagery-enhanced patients received a high dose of treatment, compared to just over half (57%) of controls (difference = 31%, 95% CI [9, 51], p = .004).

Mean Changes in Symptoms

Table 2 displays MMRM estimates of pre-treatment, post-treatment, and follow-up means for the SIAS, SPS, BFNE-S and DASS. Table 3 contains unstandardised and standardised effects sizes, and p values for tests of differences between the treatments. Observed means and standard deviations, along with MMRM and MI estimates for every session, are provided in the supplementary tables. Patients in both treatments experienced very large reductions in social anxiety symptoms. The mean reduction in interaction anxiety (SIAS) during treatment was three points larger for the imagery group, which increased to six points by follow up. The difference was statistically significant at follow up. There were no differences in performance anxiety (SPS). There were large and statistically significant differences between the treatments on both the BFNE-S and DASS. Scores decreased more for the imagery-enhanced than controls during treatment.

On the DASS, significant differences were first detected at session four (2.8 point difference, p = .007), and outcomes for patients in imagery-enhanced treatment remained

superior to those of controls at all subsequent sessions except session eight (Figure 2). On the BFNE-S, a significant difference was first detected at session nine (2.3 point difference, $p = .029$) and outcomes continued to be superior at all subsequent sessions except session ten. Differences were nearly significant at sessions three ($p = .065$) and five ($p = .054$).

Standardised Effect Sizes

Within- and between-group standardised effect sizes are displayed in Table 3. All within-group effects exceeded 0.8, which is *large* (Cumming, 2012; Sawilowsky, 2009). For patients in the imagery-enhanced group, the largest change in symptoms from pre-to-post treatment was on the SIAS ($d = 1.93$), followed by the BFNE-S ($d = 1.41$), DASS ($d = 1.38$), and SPS ($d = 1.21$). By follow-up, the effect size exceeded two standard deviations on the SIAS and approached two standard deviations on the BFNE-S. According to Sawilowsky's (2009) an effect of 1.2 is *very large* and an effect size of 2.0 is *huge*. The between group effect sizes indicated an advantage for IE-CBGT over the control except on the SPS, on which there were no differences (Table 2). The size of the difference during treatment (i.e,. pre to post) was moderate on the BFNE-S and DASS. The pre-to-post effect was somewhat smaller on the SIAS, but the size of the difference had doubled by follow up.

Clinical Significance

Four imagery-enhanced patients were excluded from RC and CSC analyses on the SIAS because they were missing scores at pre-treatment. None of the imagery-enhanced patients scored below the CSC cutoff for the SIAS at pre-treatment. Of the patients included in the analysis (n = 49), 73% (n = 36) had reliably improved and 37% (n = 18) achieved CSC by post-treatment, and 78% (n = 38) had reliably improved and 41% (n = 20) achieved CSC at follow-up. Twelve control patients were excluded due to missing pre-treatment SIAS data (n = 4) or pre-treatment scores below the CSC cut off (n = 8). Of the patients included in the analysis (n = 117), 48% (n = 56) reliably improved and 29% achieved CSC by post-treatment, and 50% (n = 58) reliably improved and 27% (n = 32) achieved CSC by follow-

up. Substantially more imagery-enhanced patients achieved reliable change than control patients, both at post-treatment (Diff = 26%, 95% CI [9, 39], p = .002) and at follow up (Diff = 28%, 95% CI [12, 41], p = .001). In terms of CSC, imagery-enhanced treatment was numerically but not statistically superior to the control at both post-treatment (Diff = 8%, 95% CI [-7, 24], p = .331) and follow up (Diff = 13%, 95% CI [-2, 29], p = .088).

Eleven imagery-enhanced patients were excluded from RC and CSC analyses on the SPS because they were missing scores at pre-treatment (n = 3) or had pre-treatment scores below the CSC cutoff (n = 8). Of the patients in the analysis (n = 42), 71% (n = 30) achieved reliable change and 50% (n = 21) achieved CSC at post-treatment, and 83% (n = 35) achieved reliable change and 69% (n = 29) achieved CSC at follow-up. Forty-five control patients were excluded due to missing pre-treatment SPS data (n = 16) or pre-treatment scores below the CSC cutoff (n = 29). Of the patients included in the analysis (n = 84), 52% (n = 44) achieved reliable change and 48% (n = 40) achieved CSC at post-treatment, and 57% (n = 48) achieved reliable change and 45% (n = 38) achieved CSC at follow-up. Substantially more imagery-enhanced patients achieved reliable change at post-treatment (Diff = 19%, 95% CI [1, 35], p = .041), and at follow up (Diff = 26%, 95% CI [9, 40], p = .003). There was no difference in rates of CSC at post-treatment (Diff = 3%, 95% CI [-16, 20], p = .801), but by follow up considerably more imagery-enhanced patients had achieved CSC (Diff = 24%, 95% CI [5, 39], p = .012).

Imagery Use and Imagery Ability as Predictors of Change

The mean scores on the VVIQ and SUIS were 10.6 (SD = 4.1) and 35.1 (SD = 8.3), and the correlation between them was r = -.13 (p = .36). The VVIQ (all rs <.27, ps > .05) and the SUIS (all rs < .17, ps>.28) were not significantly associated with pre-treatment scores on the SIAS, SPS, BFNE-S or DASS. There were modest correlations between the VVIQ and symptom *change* during treatment. The VVIQ was most strongly associated with change in BFNE-S scores (r = .32, p = .03), followed by the SIAS (r = .29, p = .04), SPS (r

= .26, *p* =.06) and DASS (*r* = .21, *p* =.16). The DASS correlation was attenuated by a single outlier; once it was removed, the correlation was .34 (*p* = .02). SUIS scores were not predictive of symptom change on the SIAS, SPS, BFNE-S or DASS (*r*s = .01, .00, -.02 and -.22, respectively, all *p*s > .16).

Because imagery ability was associated with symptom change, we explored whether outcomes for imagery-enhanced patients with *low* imagery ability were worse than those of control group patients. Low imagery ability was defined as a VVIQ score at the 25th percentile (i.e., 8). For each symptom measure, we used regression to estimate the mean change during treatment for imagery-enhanced patients with low imagery ability. These estimated changes were then compared to the mean changes for control patients. Imagery-enhanced patients with low imagery ability had similar or higher mean change scores compared to controls on the SIAS (17.00 vs. 16.30), BFNE-S (7.68 vs. 6.28), and DASS (10.00 vs. 7.10), but somewhat lower on the SPS (16.70 vs. 19.10).

Discussion

Social anxiety disorder (SAD) is an early onset and debilitating disorder that tends to be unremitting without treatment (DeWit, Ogborne, Offord, & McDonald, 1999). CBT is efficacious and effective for SAD but a substantial minority of sufferers do not respond to treatment in community clinics (McEvoy, 2007; McEvoy et al., 2012). In an attempt to improve outcomes, the first aim of this study was to compare the acceptability and effectiveness of an imagery-enhanced CBGT protocol to a large sample of historical controls who completed predominantly verbally-based CBGT. Importantly, both interventions were delivered in the same community mental health clinic and were the same length, patients were recruited via the same referral pathways, and the groups did not significantly differ on a range of clinical and demographic variables, suggesting that both samples were drawn from a similar population.

The imagery-enhanced protocol was associated with very high retention and 26% more patients received a high dose of the imagery-enhanced protocol compared to historical controls. The imagery-enhanced completion rate of 91% compared very favourably to previous effectiveness trials in community clinics (Hofmann & Suvak, 2006; Lincoln et al., 2005). Furthermore, while comorbid depression was associated with poorer attendance in the verbally-based intervention this was not the case for the imagery-enhanced intervention. The imagery-enhanced protocol therefore appeared to be highly acceptable within this severe and highly comorbid sample.

Both treatments were highly effective and demonstrated large effect sizes on all outcome measures. As hypothesised, between-groups effect sizes demonstrated that the imagery-enhanced group improved more on the BFNE-S and DASS at post-treatment (small to medium effect sizes), and on the SIAS, BNE and DASS at follow-up (medium to large effect sizes). The magnitude of these differences was striking given that both treatments targeted the same key mechanisms, albeit via predominantly different modes (verbal vs. imagery). Within-groups effect sizes were largest for the imagery-enhanced group and compared very favourably to previous group and individual cognitive behavioural therapy trials (Clark et al., 2003; McEvoy et al., 2012; Rapee et al., 2009). Weekly comparisons between the imagery and control groups demonstrated that scores on the BFNE-S and DASS significantly diverged in favour of the imagery-enhanced group. A majority of patients receiving the imagery intervention achieved reliable change, and around one-third and two-thirds achieved clinically significant change on social interaction anxiety and performance anxiety, respectively. A higher proportion of the imagery-enhanced group achieved reliable change on social interaction anxiety and performance anxiety at post-treatment and follow-up, and a significantly higher proportion of the imagery-enhanced group achieved CSC on the SPS at follow-up. Overall, the imagery-enhanced protocol appeared to be more acceptable and effective than the predominantly verbal-linguistic control intervention.

The one exception to this pattern was that both groups improved to a similar degree on performance anxiety, for which there are at least two potential explanations. First, all patients scored within the clinical range on the social interaction anxiety scale but not on performance anxiety, which is suggestive of a predominance of generalised social interaction anxiety. The exclusion of the DSM-IV (APA, 1994) generalised specifier, but retention of the performance-only specifier, in DSM-5 (APA, 2013) reflects the fact that most patients in clinical practice experience anxiety in a broad array of social situations. Performance situations may have been only one of numerous social anxiety triggers and the SIAS therefore may have been a more comprehensive and sensitive measure of the sample's broad spectrum of social fears. The substantially larger pre-treatment standard deviation on the SPS compared to the SIAS is consistent with performance anxiety being prominent for a smaller proportion of the sample compared to interaction anxiety. Second, both programs included an identical number of formal within-session performance exposure tasks (i.e., one video-taped presentation) and a larger number of social interaction tasks. Given the generalised nature of the patients' social fears they may have also prioritised social interactions for homework tasks. Thus, during treatment there was likely to have been a larger 'dose' of imagery enhancements within social interaction situations relative to performance situations. Individualised hierarchies were not retained for analysis, so this explanation is speculative. Future studies directly comparing imagery- to verbally-based interventions for performance anxiety in particular are required to more clearly determine whether one mode is superior to the other for this subtype.

The second aim of this study was to examine whether (a) self-reported tendencies to operate within an imagery mode and (b) imagery ability were associated with outcomes in the imagery-enhanced group. Patients' self-reported tendency to use imagery in daily life was not associated with symptom change, suggesting that those who naturally operate within an imagery mode do not necessarily do better in the imagery-enhanced intervention than

those who do not have this tendency. This finding is consistent with previous studies that have failed to find this association (e.g., Lee & Kwon, 2013). It is important to note that patients were not discouraged from applying their skills to negative verbal-linguistic thoughts throughout the imagery-enhanced protocol, but they were encouraged to transform thoughts into specific vivid images, for example, before and after cognitive restructuring and behavioural experiments. Few patients reported an inability to transform their thoughts into images, although some reported somewhat fragmented images or a 'felt sense' rather than vivid images. Greater imagery awareness emerged for some patients during the group. For instance, one patient reported that she always thinks in verbal thoughts and then immediately described it as being "like a whirlwind of thoughts above her head." This vivid metaphorical image could then be transformed in therapy and, in turn, the meaning of the image as an uncontrollable and chaotic spiral could be modified.

The association between self-reported imagery ability and symptom improvement was modest but significant, suggesting that those who reported being able to elicit more vivid (benign) images before treatment achieved greater symptom relief during treatment. It is plausible that the imagery-enhanced protocol was more effective for those with greater imagery ability, but this study cannot definitively support this proposition given that this measure was only administered to the imagery-enhanced group. It is important for future research to demonstrate that poor imagery ability is not a proxy for cognitive inflexibility more generally, a core feature of emotional disorder (e.g., Hayes, Strosahl, & Wilson, 2011), which could also predict outcomes from verbally-based interventions. However, if imagery ability was simply a proxy for cognitive flexibility it would be expected to negatively correlate with symptom severity at pre-treatment, but this was not the case. It is also important to note that those with poorer imagery ability achieved comparable change scores to the historical controls, suggesting that they did not do more poorly in the imagery-enhanced protocol but rather that those with superior imagery ability did better.

Additionally, 69% of the sample scored below the mid-point on the VVIQ, suggesting that relatively poor imagery ability overall did not adversely affect outcomes. Therefore, regardless of imagery ability the imagery-enhanced protocol may still prove to be the treatment of choice.

A large body of evidence suggests that similar emotional, cognitive, and physiological responses are observed in both imagined and actual events (Kosslyn, Ganis, & Thompson, 2001), and recent evidence has demonstrated that imagery has a particularly powerful relationship to emotions compared to verbal-linguistic activity (Holmes & Mathews, 2010). These literatures suggest that imagery-based therapeutic techniques may facilitate greater emotional activation and modification than verbal-linguistic techniques alone. Imagery in psychotherapy is not new (Beck, Rush, Shaw, & Emery, 1979; Edwards, 2007) and it has been incorporated into treatments for a large range of disorders. However, the contribution of imagery-based techniques to clinical outcomes has recently garnered considerable interest within the emotional disorder literature in general and the SAD literature in particular. Studies investigating the impact of imagery rescripting alone (e.g., Nilsson et al., 2012) and within CBT packages (Clark et al., 2006) delivered individually have been promising.

This study builds on these earlier studies in several ways. First, other than the pilot study (McEvoy & Saulsman, 2014) we are unaware of any previous study incorporating past imagery rescripting within a group context. Second, imagery was incorporated into all treatment components in an attempt to more comprehensively exploit the strong imagery-emotion association. Third, our study found that the imagery-enhanced protocol was associated with low attrition, including amongst those with comorbid depression. Emerging evidence suggests that dysphoria and anxiety are associated with a deficit of positive imagery (Holmes, Lang, Moulds, & Steele, 2008; Morina, Deeprose, Pusowski, Schmid, & Holmes, 2011). Although our finding requires replication, it is tempting to speculate that the

imagery-enhanced intervention better engaged patients with and without comorbid depression by explicitly and repeatedly prompting more realistic and positive imagery to correct this deficit. Fourth, the finding that greater imagery ability was associated with greater symptom change during the imagery-enhanced protocol was also novel and may suggest that, to the degree that imagery ability is trainable, techniques designed to enhance imagery ability prior to an imagery-enhanced intervention could increase treatment potency. This intriguing possibility requires further research.

This study's findings must be considered in light of several limitations. Patients were not randomised to treatments, which may have introduced systematic differences across the samples. This concern was somewhat mitigated by the similarities in the samples in terms of clinical and sociodemographic variables, treatment setting, recruitment procedures and clinicians. It is noteworthy that MMRM estimates outcomes assuming all patients had completed treatment. Attrition was low for the imagery group, so we can be confident that the MMRM estimates for this group closely approximated those likely to be seen in clinical practice. In contrast, the higher attrition rate in the control group was likely to have overestimated treatment effectiveness in practice, because those who discontinued treatment were unlikely to have achieved the same degree of benefit than if they had attended the whole program (as is assumed from MMRM). The estimated between-treatment effect sizes are therefore more likely to be underestimates than overestimates.

The lack of a waitlist or placebo control group also means that we cannot definitively conclude that patients would not have improved without intervention or that the specific content of the treatments caused the symptom changes. However, these possibilities are unlikely given that (a) pre-treatment means were amongst the most severe in the published literature, (b) SAD is chronic and unremitting without treatment (DeWit et al., 1999), (c) the effect sizes were amongst the largest in the published literature, including trials that have included control groups (Clark et al., 2003, 2006), and (d) the historical control treatment

used in this study has been demonstrated to be more efficacious than more traditional CBT and stress management control groups (Rapee et al., 2009) and equally effective across research and community mental health settings (McEvoy et al., 2012).

Another limitation was that treatment adherence was not independently assessed, but was instead supported by the use of detailed therapist notes, handouts, and worksheets. Effectiveness trials emphasise external validity and the important contribution they make to transporting efficacious treatments into real world clinics has been well recognised (Stewart & Chambless, 2009). It is critical for the dissemination of evidence-supported treatments that efficacious treatments are shown to be effective in clinically representative samples and settings. Nonetheless, the superiority of imagery-enhanced group treatment found in this study needs to be verified in a randomised controlled trial. Our reliance on self-report measures was also a limitation. Finally, the standard procedure within the community clinic is to provide a one-month follow-up, but longer-term assessment of trajectories of change would be informative.

This study found that an imagery-enhanced CBGT program in a community clinic was associated with high retention and large effect sizes, and compared well to historical controls who completed a predominantly verbally-based intervention. The ability to evoke vivid imagery was associated with superior outcomes from imagery-enhanced CBGT but there was no evidence that those without this ability did more poorly than those completing the comparison treatment. Future randomised controlled trials investigating whether imagery-based techniques can enhance outcomes for SAD and other emotional disorders would be informative, as would studies investigating differential mechanisms of change across verbally- and imagery-based interventions.

References

American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.). Washington DC: APA.

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). Washington DC: APA.

Andrews, G., Henderson, S., & Hall, W. (2001). Prevalence, comorbidity, disability and service utilisation: Overview of the Australian National Mental Health Survey. *British Journal of Psychiatry, 178*, 145-153.

Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). Cognitive therapy of depression. New York: The Guilford Press.

Carleton, R. N., Thibodeau, M. A., Weeks, J. W., Teale Sapach, M. J. N., McEvoy, P. M., Horswill, S. C., & Heimberg, R. G. (in press). Comparing Short Forms of the Social Interaction Anxiety Scale and the Social Phobia Scales. *Psychological Assessment*. Accepted 19/04/2014.

Chiupka, C. A., Moscovitch, D. A., & Bielak, T. (2012). In vivo activation of anticipatory vs. post-event autobiographical images and memories in social anxiety. *Journal of Social and Clinical Psychology, 31*, 783-809.

Clark, D. M., Ehlers, A., Hackmann, A., McManus, F., Fennell, M., Grey, N., et al. (2006). Cognitive therapy versus exposure and applied relaxation in social phobia: A randomized controlled trial. *Journal of Consulting and Clinical Psychology,* 74, 568–578.

Clark, D. M., Ehlers, A., McManus, F., Hackmann, A., Fennell, M., Campbell, H., et al. (2003). Cognitive therapy versus fluoxetine in generalized social phobia: A randomized placebo-controlled trial. *Journal of Consulting and Clinical Psychology, 71*, 1058–1067.

Cox, B. J., Ross, L., Swinson, R. P., & Direnfeld, D. M. (1998). A comparison of social phobia outcome measures in cognitive-behavioral therapy. *Behavior Modification, 22*, 285-297.

Cumming, G. (2012). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. New York: Routledge.

DeWit, D. J., Ogborne, A., Offord, D. R., & MacDonald, K. (1999). Antecedents of the risk of recovery from DSM-III-R social phobia. *Psychological Medicine, 29*, 569-582.

Edwards, D. (2007). Restructuring implicational meaning through memory-based imagery: some historical notes. *Journal of Behavior Therapy and Experimental Psychiatry, 38*, 306-316.

Hackmann, A., Bennett-Levy, J. & Holmes, E. A. (2011). *Oxford Guide to Imagery in Cognitive Therapy.* Oxford: Oxford University Press.

Hackmann, A., Clark, D. M., & McManus, F. (2000). Recurrent images and early memories in social phobia. *Behaviour Research and Therapy, 38*, 601-610.

Harvey, A. G., Clark, D. M., Ehlers, A., & Rapee, R. M. (2000). Social anxiety and self-impression: cognitive preparation enhances the beneficial effects of video feedback following a stressful social task. *Behaviour Research and Therapy, 38*, 1183-1192.

Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (2011). Acceptance and Commitment Therapy: The Process and Practice of Mindful Change (2nd Ed.). New York: The Guilford Press.

Heimberg, R. G., Brozovich, F. A., & Rapee, R. M. (2014). A cognitive behavioural model of social anxiety disorder: update and extension. In S. G. Hofmann & P. M. DiBartolo (Eds), *Social Anxiety: Clinical, Developmental, and Social Perspectives (3rd edition, pp. 706-729).* New York: Elsevier.

Hirsch, C. R., Clark, D. M., Mathews, A., & Williams, R. (2003). Self-images play a causal role in social phobia. *Behaviour Research and Therapy, 41*, 909-921.

Hirsch, C. R., Meynen, T., & Clark, D. M. (2004). Negative self-imagery in social anxiety contaminates social interactions. *Memory, 12*, 496-506.

Hofmann, S. G., & Suvak, M. (2006). Treatment attrition during group therapy for social phobia. *Anxiety Disorders, 20*, 961-972.

Holmes, E. A., Lang, T. J., Moulds, M. L., & Steele, A. M. (2008). Prospective and positive mental imagery deficits in dysphoria. *Behaviour Research and Therapy, 46*, 976-981.

Holmes, E. A., Lang, T. J., & Shah, D. M. (2009). Developing interpretation bias modification as a 'cognitive vaccine' for depressed mood . *Journal of Abnormal Psychology, 118*, 76-88.

Holmes, E. A., & Mathews, A. (2010). Mental imagery in emotion and emotional disorders. *Clinical Psychology Review, 30*, 349-362.

Hunt, M., & Fenton, M. (2007). Imagery rescripting versus in vivo exposure in the treatment of snake fear. *Journal of Behavior Therapy and Experimental Psychiatry, 38*, 329-344.

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.

Kosslyn, S. M., Ganis, G., & Thompson, W. L. (2001). Neural foundations of imagery. Nature Reviews: Neuroscience, 2, 635-642.

Leary, M. R. (1983). A brief version of the Fear of Negative Evaluation Scale. Personality and *Social Psychology Bulletin, 9*, 371-375.

Lee, S. W., & Kwon, J-H. (2013). The efficacy of imagery rescripting (IR) for social phobia: a randomized controlled trial. *Journal of Behavior Therapy and Experimental Psychiatry, 44*, 351-360.

Lenth, R. V. (2014). lsmeans: Least Squares Means Version 2.11. Retrieved from http://cran.r-project.org/web/packages/lsmeans

Lincoln, T. M., Rief, W., Hahlweg, K., Frank, M., von Witzleben, I., Schroeder, B., & Fiegenbaum, W. (2005). Who comes, who stays, who profits? Predicting refusal, dropout, success, and relapse in a short intervention for social phobia. *Psychotherapy Research, 15,* 210-225.

Lovibond, S.H. & Lovibond, P.F. (1995). *Manual for the Depression Anxiety Stress Scales*. (2nd. Ed.) Sydney: Psychology Foundation.

Marks, D. F. (1977). Imagery and consciousness: a theoretical review from an individual differences perspective. *Journal of Mental Imagery 12,* 275–290.

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.

McEvoy, P. M. (2007). Effectiveness of cognitive behavioural group therapy for social phobia in a community clinic: a benchmarking study. *Behaviour Research and Therapy, 45,* 3030-3040.

McEvoy, P. M., Grove, R., & Slade, T. (2011). Epidemiology of Anxiety Disorders in the Australian General Population: Findings of the 2007 Australian National Survey of Mental Health and Wellbeing. *Australian and New Zealand Journal of Psychiatry, 45,* 957-967.

McEvoy, P. M., Nathan, P., Rapee, R. M., & Campbell, B. N. C. (2012). Cognitive behavioural group therapy for social phobia: Evidence of transportability to community clinics. *Behaviour Research and Therapy, 50,* 258-265.

McEvoy, P. M., & Saulsman, L. (2014). Imagery-enhanced cognitive behavioural group therapy for social anxiety disorder: A pilot study. *Behaviour Research and Therapy, 55,* 1-6.

McKelvie, S. J. (1986). Effects of format of the Vividness of Visual Imagery Questionnaire on content validity, split-half reliability, and the role of memory in test–retest reliability. *British Journal of Psychology, 77*, 229–236.

Morina, N., Deeprose, C., Pusowski, C., Schmid, M., & Holmes, E. A. (2011). Prospective mental imagery in patients with major depressive disorder or anxiety disorders. *Journal of Anxiety Disorders, 25*, 1032-1037.

Morris, S. B. (2008). Estimating Effect Sizes From the Pretest-Posttest-Control Group Designs. *Organizational Research Methods, 11*, 364–386.

Moscovitch, D. A., Garvric, D. L., Merrifield, C., Bielak, T., & Moscovitch, D. (2011). Retrieval properties of negative vs. positive mental images and autobiographical memories in social anxiety: Outcomes with a new measure. *Behaviour Research and Therapy, 49*, 505-17.

National Research Council Panel on Handling Missing Data in Clinical Trials. (2010). *The Prevention and Treatment of Missing Data in Clinical Trials*. Washington, DC: The National Academies Press.

Newcombe, R. G. (2013). *Confidence intervals for proportions and related measures of effect size.* Boca Raton, FL, CRC Press.

Ng, A. S., Abbott, M. J., & Hunt, C. (2014). The effect of self-imagery on symptoms and processes in social anxiety: a systematic review. *Clinical Psychology Review, 34*, 620-633.

Nilsson, J., Lundh, L., & Viborg, G. (2012). Imagery rescripting of early memories in social phobia: An experimental study. *Behaviour Research and Therapy, 50*, 387-92.

Page, A. C., Hooke, G. R., & Morrison, D. L. (2007). Psychometric properties of the Depression Anxiety Stress Scales (DASS) in depressed clinical samples. *British Journal of Clinical Psychology, 46*, 283-297.

Pinheiro, J. C., & Bates, D. M. (2000). Mixed-Effects Models in S and S-PLUS. New York: Springer-Verlag.

R Development Core Team. (2014). R: A language and environment for statistical computing. Vienna, Australia: R Foundation for Statistical Computing.

Rapee, R. M., Gaston, J. E., & Abbott, M. J. (2009). Testing the efficacy of theoretically derived improvements in the treatment of social phobia. *Journal of Consulting and Clinical Psychology, 77*, 317-327.

Rapee, R. M., & Heimberg, R. G. (1997). A cognitive-behavioral model of anxiety in social phobia. *Behaviour Research and Therapy, 35*,741–756.

Reisberg, D., Pearson, D. G., & Kosslyn, S. M. (2003). Intuitions and introspections about imagery: the role of imagery experience in shaping an investigator's theoretical views. *Applied Cognitive Psychology, 17*, 147-160.

Richardson, J. T. E. (1999). *Imagery.* Psychology Press: Hove.

Rodebaugh, T. L., Woods, C. M., Thissen, D. M., Heimberg, R. G., Chambless, D. L., & Rapee, R. M. (2004). More information from fewer questions: the factor structure and item properties of the original and Brief Fear of Negative Evaluation scale. *Psychological Assessment, 16*, 169–181.

Sawilowsky, S. S. (2009). Very large and huge effect sizes. *Journal of Modern Applied Statistical Methods, 8*, 597-599.

Shah, A. D., Bartlett, J. W., Carpenter, J., Nicholas, O., & Hemingway, H. (2014). Comparison of Random Forest and Parametric Imputation Models for Imputing Missing Data Using MICE: A CALIBER Study. *American Journal of Epidemiology, 179*, 764-774.

Shah, A. D., Bartlett, J., Hemingway, H., Nicholas, O., & Hingorani, H. (2014). *CALIBERrfimpute: Imputation in MICE using Random Forest. (Version 0.1-6).* [Software].

Sheehan, D., Janavus, J., Baker, R., Harnett-Sheehan, K., Knapp, E., & Sheehan, M. (2001) *M.I.N.I. Plus. (Mini International Neuropsychiatric Interview).* Version 5.0.

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Janavs, J., Weiller, E., Keskiner, A.,et al. (1997). The validity of the Mini International Neuropsychiatric Interview (MINI) according to the SCID-P and its reliability. *European Psychiatry, 12*, 232-241.

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.

van Buuren, S. (2012). *Flexible Imputation of Missing Data*. New York: Chapman and Hall/CRC.

Weeks, J. W., Heimberg, R. G., Fresco, D. M., Hart, T. A., Turk, C. L., Schneier, F. R., et al. (2005). Empirical validation and psychometric evaluation of the Brief Fear of Negative Evaluation scale in patients with social anxiety disorder. *Psychological Assessment, 17*, 179-190.

Wild, J., & Clark, D. M. (2011). Imagery rescripting of early traumatic memories in social phobia. *Cognitive and Behavioral Practice, 18*, 433-443.

Table 1

Comparison of pre-treatment demographic and clinical characteristics for the imagery enhanced and historical control groups

	IE-CBGT	CONTROL	Difference		
Measure	(*n* = 53)	(*n* = 129)	Est	95% CI	*p*
Demographics					
Age (years)	29 (11)	31 (11)	-2.0	[-6, 1]	.22
Women	53	40	13	[-2, 29]	.10
Single	77	70	7	[-7, 21]	.32
Employed	42	42	-1	[-16, 15]	.93
Tertiary educated	23	27	-4	[-18, 10]	.61
Diagnoses					
≥ 2 diagnoses	74	74	-1	[-13, 16]	.91
≥ 3 diagnoses	34	32	2	[-12, 17]	.77
MDD / Dysthymia	47	59	-12	[-27, 4]	.15
GAD	30	29	1	[-13, 16]	.92
Other clinical features					
Self-harmed	19	24	-5	[-18, 8]	.45
Attempted suicide	22	30	-8	[-22, 6]	.24
Hospitalised	21	26	-6	[-19, 8]	.42
Medicated	60	65	-5	[-21, 11]	.53

Note. The numbers are percentages except for age, which is a mean and standard deviation. IE-CBGT = imagery enhanced cognitive behavioural group therapy, Est = estimate, MDD = major depressive disorder, GAD = generalised anxiety disorder.

Table 2
MMRM estimated means (and standard errors) at pre-treatment, post-treatment, and follow up

	IE-CBGT		CONTROL	
Measure and time	M	SE	M	SE
SIAS				
Pre	58.7	1.4	57.6	0.9
Post	39.4	2.0	41.3	1.5
Follow Up	36.9	2.2	41.9	1.7
SIAS				
Pre	42.8	2.3	43.0	1.5
Post	23.7	2.0	23.9	1.5
Follow Up	21.8	2.0	22.9	1.6
BFNE-S				
Pre	32.2	0.8	31.8	0.6
Post	23.3	1.1	25.5	0.8
Follow Up	21.5	1.1	24.6	0.8
DASS				
Pre	20.8	1.1	18.0	0.7
Post	9.9	1.1	11.0	0.8
Follow Up	9.4	1.3	10.4	0.9

Note. IE-CBGT = imagery enhanced cognitive behavioural group therapy, SIAS = Social Interaction Anxiety Scale, SPS = Social Phobia Scale, BFNE-S = Brief Fear of Negative Evaluation scale, DASS = Depression, Anxiety, and Stress Scales.

Table 3

MMRM estimates of unstandardised and standardised within-group and between-group treatment effects

	Mean Changes					Standardized Effect Sizes			
Measure and Time	IE-CBGT	Control	Diff	95% CI		IE-CBGT	Control	Diff	*p*
SIAS									
Pre to post	19.3	16.3	3.0	[-1.8,	7.9]	1.93	1.63	.30	.224
Post to follow up	21.8	15.7	6.1	[0.8,	11.3]	2.18	1.57	.61	.024
SPS									
Pre to post	19.0	19.1	0.0	[-5.3,	5.3]	1.20	1.21	.00	.991
Post to follow up	21.0	20.0	1.0	[-4.6,	6.5]	1.33	1.27	.06	.737
BFNE-S									
Pre to post	8.8	6.3	2.5	[0.2,	4.9]	1.41	1.01	.41	.033
Post to follow up	10.6	7.3	3.4	[0.9,	5.9]	1.70	1.16	.54	.008
DASS									
Pre to post	10.8	7.1	3.8	[1.2,	6.3]	1.39	0.90	.48	.004
Post to follow up	11.4	7.6	3.8	[0.8,	6.8]	1.45	0.97	.48	.014

Note. IE-CBGT = Imagery enhanced cognitive behavioural group therapy; Diff = Difference between the IE-CBGT and Control groups; CI = Confidence Interval. SIAS = Social Interaction Anxiety Scale, SPS = Social Phobia Scale, BFNE-S = Brief Fear of Negative Evaluation scale, DASS = Depression, Anxiety, and Stress Scales. *p* value is for the difference between the mean change scores. Pooled pre-treatment standard deviations used to compute the standardised mean changes were 9.98 (SIAS), 15.80 (SPS), 6.25 (BFNE-S), and 7.83 (DASS).

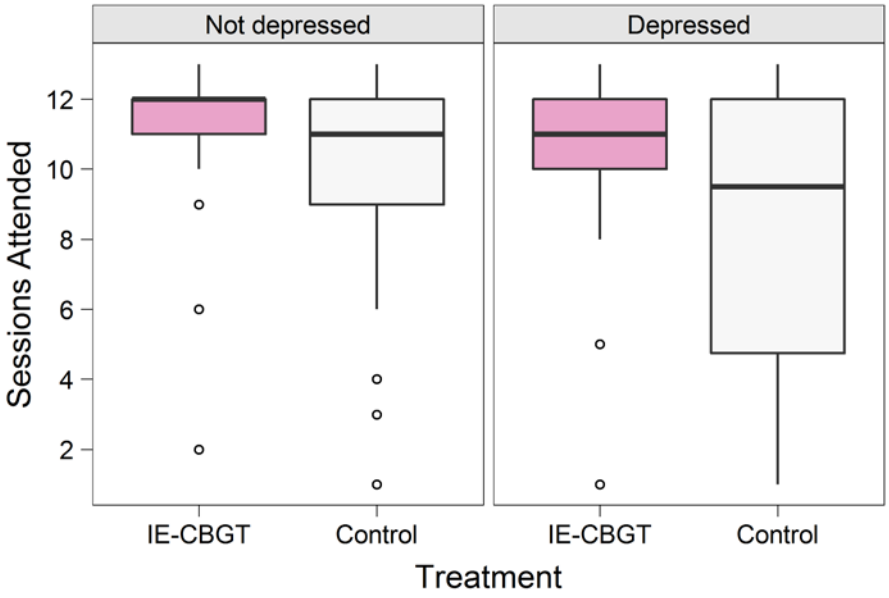

Figure 1. Boxplot showing the number of treatment sessions completed, stratified by treatment type and depressive diagnosis. Most IE-CBGT patients completed a high number of sessions irrespective of whether they were depressed or not. Amongst control patients, session attendance was much more variable. This was particularly the case for depressed patients, as is evident from the massive spread of the box and whiskers for the depressed controls.

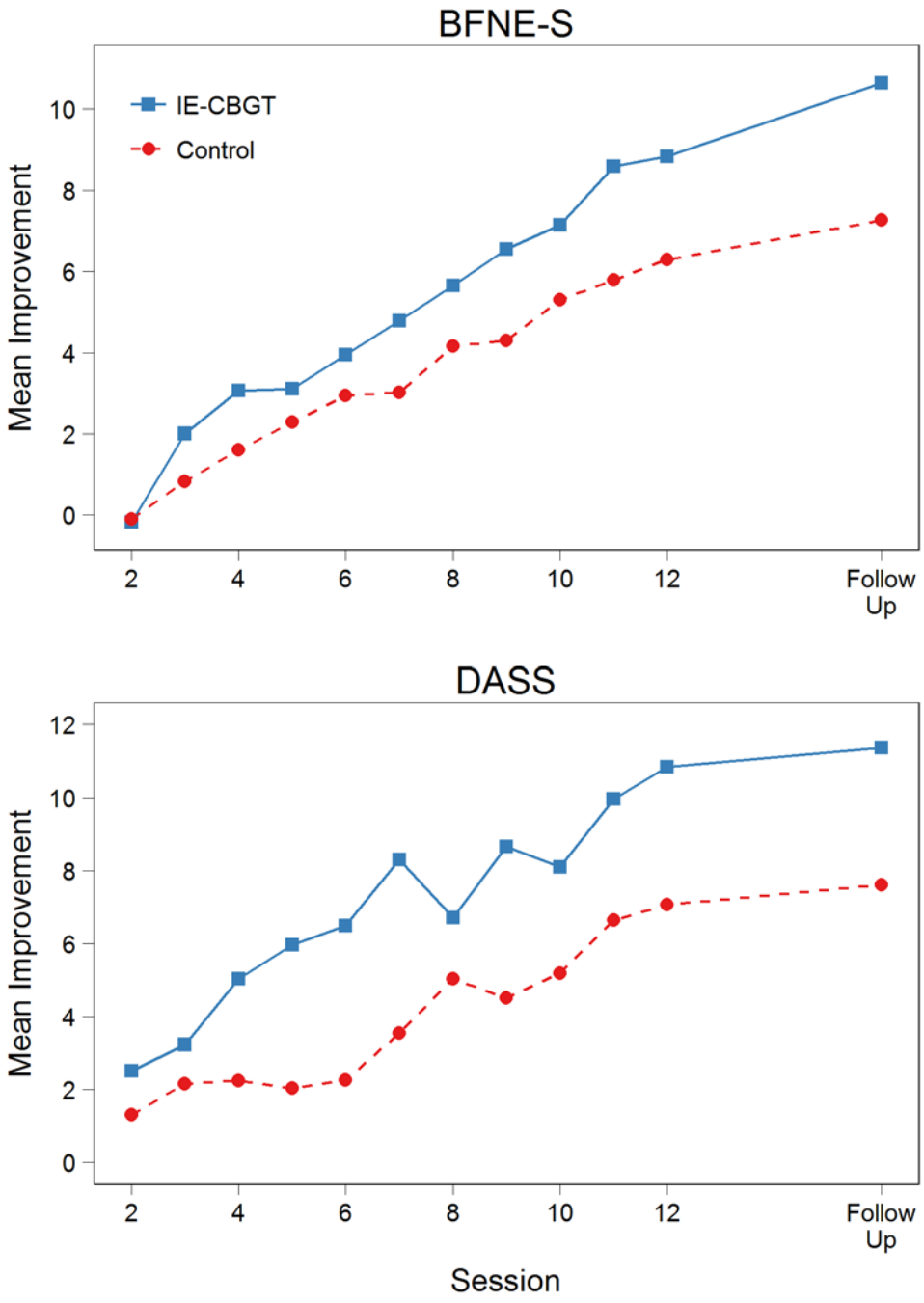

Figure 2. Trajectories of symptom improvement on the Brief Fear of Negative Evaluation Scale (BFNE-S) and Depression, Anxiety and Stress Scale (DASS) across treatments sessions, and at one month follow up. Mean improvements in symptoms since the first session (T1) are plotted.

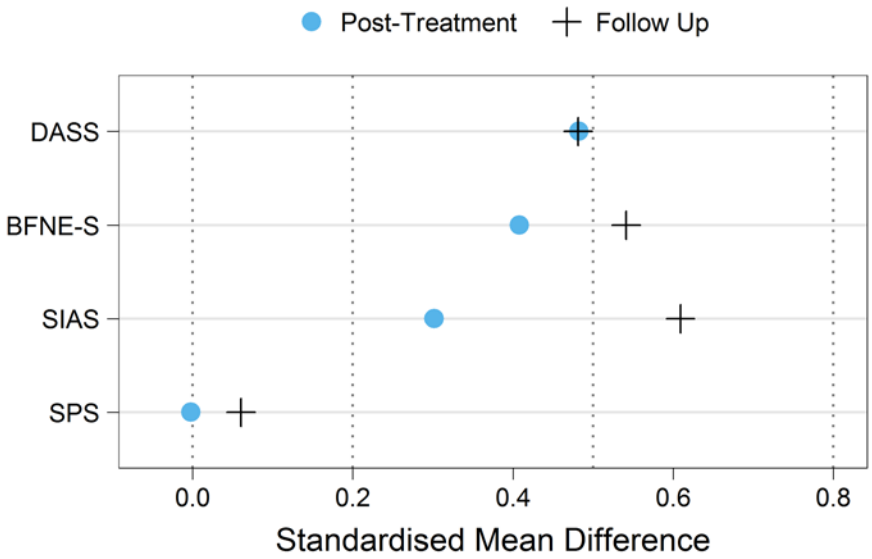

Figure 3. Dotplot of between-group pre-to-post treatment, and pre-treatment to follow up standardised effect sizes. A positive effect size indicates that symptoms reduced more for IE-CBGT patients than controls. Vertical lines correspond to what are often considered to be nil (0.0), small (0.2), medium (0.5) and large (0.8) effects.