**Supplement**

**Imagery-enhanced versus verbally-based group cognitive behavior therapy for social anxiety disorder: a randomized controlled trial**

*Description of CONSORT diagram*

A CONSORT diagram showing the flow of participants through the study can be found in Figure 1. Ten participants randomized to treatment did not commence therapy due to a clash with other commitments such as study, work, of family responsibilities (*n* = 7) or for unknown reasons (*n* = 3). These participants were unaware of the condition to which they had been assigned, therefore their failure to commence therapy could not have been related to treatment condition. Patients who started therapy but failed to complete it or provide follow-up data also nominated a clash with other commitments as the primary reason (*n* = 6 imagery-enhanced cognitive behaviour therapy, IE-CBT, *n* = 12 verbally-based, VB-CBT). Three patients (*n* = 1 IE-CBT, *n* = 2 VB-CBT) stopped treatment due to losing interest in therapy. In the IE-CBT condition, one patient was removed from the study due to a significant increase in suicidal ideation at the 1-month follow-up, although clinicians were informed that this elevated risk resolved within one week. Another imagery patient reported being too anxious to continue with the group and was put on a wait list for individual therapy; and one ceased treatment after it became apparent that a personality disorder, rather than SAD, was the patient’s primary problem. In the VB-CBT arm, one patient was referred to another service after it became apparent depression rather than SAD was the principal problem, and patients did not provide follow-up data due to medical illness (*n* = 1), moving away (*n* = 1) or being too anxious (*n* = 1). All patients randomized to treatment were included in analyses, except for two IE-CBT patients who were excluded for intentionally providing invalid responses to questionnaires, such as repeatedly rating outside the maximum of a scale and continually using the same response option to rate every item in a scale. The likely implausible responses were raised by treating clinicians and reviewed by a member of the research team (DEH) blind to treatment condition, who confirmed that they were consistent with invalid responding. Data from these two patients were excluded to prevent biasing the analyses. This decision was made before undertaking analyses and documented in the trial’s statistical analysis plan, which was uploaded to the Open Science Foundation (https://osf.io/msq9w/).

**eTable 1: Interrater Reliability for DSM-5 Diagnoses and Clinician-Rated Severity**

|  |  |  |
| --- | --- | --- |
| Rating | Gwet’s AC1 | Cohen’s Kappaa |
| SAD | 1 | 1 |
| SAD Severity | .46 | .23 |
| MDD (Current) | .45 | .46 |
| GAD | .48 | .40 |

*Note.* SAD = social anxiety disorder, MDD = major depressive disorder, GAD = generalized anxiety disorder. aMcHugh (2012).

*Clinician training*

All treating clinicians had a masters degree or doctorate in clinical psychology. In Australia, this means that all therapists received at least two or three years of full-time training in evidence-based treatments, respectively, and CBT is required to be the predominant therapeutic orientation. In addition, to be fully registered as a clinical psychologist, clinicians are required to undertake two years of post-university supervision with a certified supervisor. Clinical supervisors are also required to undertake additional training to be certified as a clinical supervisor. The lead author (PM) has 16 years of experience facilitating CBT groups (along with individual CBT) and supervising other practitioners. All lead clinicians completed their post-graduate qualifications and were supervised weekly by the lead researcher. Clinicians were also required to co-facilitate a group with PM or LS before leading subsequent groups and training other clinicians.

*Intervention content*

Our previous studies evaluating IE-CBT (McEvoy, Erceg-Hurn, Saulsman, & Thibodeau. 2015; McEvoy and Saulsman, 2014) compared the imagery-enhanced protocol to historical controls who completed the protocol developed by Rapee and colleagues(2009). Compared to the historical verbally-based protocol, the imagery-enhanced protocol introduced imagery-based techniques throughout all sessions and rearranged some session content based on clinical experience. The larger effects of the verbally-based protocol in this trial, compared to historical controls completing the previous verbally-based protocol, may be due to the rearranged content. In eTable 2 we outline the main session content of the two protocols evaluated in this trial, but also the historical VB-CBT protocol, to inform readers of the key similarities and differences so that the methods and findings from this trial can be contextualized with the historical studies. The main difference between the current and historical VB-CBT protocols was that the first three sessions of content in the historical protocol was covered in the first two sessions of the current VB-CBT protocol, so that thought monitoring and challenging began earlier. As a consequence, subsequent content (behavioral experiments, safety behaviors experiments, video-feedback to challenge self-images) was also introduced earlier. In vivo behavioral experiments are also included in three sessions rather than two, although the cumulative number of completed experiments is similar. These differences may account for why the past benchmarking studies revealed superiority of IE-CBT compared to the historical VB-CBT protocol, and why the VB-CBT protocol in the current trial yielded larger effect sizes than the historical VB-CBT protocol. The historical VB-CBT was modified for the current trial because, to answer the key question of whether the imagery-enhancements improved outcomes to a verbally-based approach, it was critical that the session content aligned across the two conditions with only the mode of intervention varying.

**eTable 2. Main components of each session for the imagery-enhanced and verbally-based CBT conditions in the current trial, but also for historical verbally-based protocol evaluated in previous research**

|  |  |  |  |
| --- | --- | --- | --- |
| **Session** | **Imagery-enhanced CBT a** | **Verbally-based CBT a** | **Verbally-based CBT (historical) b** |
| 1 | **Socialize to the model**: multi-sensory social images as key maintaining factor, imagery monitoring | **Socialize to the model**: verbal thoughts as key maintaining factor, imagery monitoring | **Socialize to the model**: verbal thoughts as key maintaining factor |
| 2 | **Identify/challenge** thoughts and imagery, balanced image | **Identify/challenge** thoughts, balanced thought | **Identify** thoughts |
| 3 | **Behavioral experiments introduction**: use of imagery to guide predictions, use imagery to update predictions, coping imagery | **Behavioral experiments introduction**: use of thoughts to guide predictions | **Challenge** thoughts |
| 4 | **Behavioral experiments 1**: in vivo | **Behavioral experiments 1**: in vivo | **Behavioral experiments introduction**: use of thoughts to guide predictions |
| 5 | **Safety behaviors**: visualize using and not using safety behaviors before behaviorally testing the effects of safety behaviors | **Safety behaviors**: behaviorally test effects of using and not using safety behaviors | **Behavioral experiments 1**: in vivo |
| 6 | **Negative self-image**: video feedback session | **Negative self-image**: video feedback session | **Safety behaviors**: behaviorally test effects of using and not using safety behaviors |
| 7 | **Behavioral experiments 2**: in vivo | **Behavioral experiments 2**: in vivo | **Negative self-image**: video feedback session |
| 8 | **Attention training**: visualize interactions with task- vs self-focused attention before testing the effects behaviorally | **Attention training**: test the effects of task- vs self-focused attention behaviorally | **Attention training**: test the effects of task- vs self-focused attention behaviorally |
| 9 | **Core beliefs 1**: imagery rescripting | **Core beliefs 1:** downward arrowing, monitoring inconsistent evidence | **Behavioral experiments 2 & 3**: in vivo |
| 10 | **Behavioral experiments 3**: in vivo | **Behavioral experiments 3**: in vivo | **Core beliefs 1:** downward arrowing, monitoring inconsistent evidence |
| 11 | **Core beliefs 2**: new core beliefs and action plans via imagery | **Core beliefs 2**: New action plans via verbal discussion | **Core beliefs 2**: New action plans via verbal discussion |
| 12 | **Review**: relapse prevention via imagery, future imagery | **Review**: relapse prevention and future planning without imagery | **Review**: relapse prevention and future planning without imagery |

a Treatments evaluated in the randomized controlled trial reported in this paper, b Verbally-based CBT protocol previously compared to imagery-enhanced CBT by McEvoy and Saulsman (2014) and McEvoy et al (2015). See McEvoy, Saulsman, & Rapee (2018) for more detailed information about each strategy.

***Assessments (Extended)***

*Primary outcome measures.* The three pre-registered primary outcomes captured self-report symptoms, diagnoses, and clinician-rated severity. All self-report measures were highly reliable, as indexed by Omega (*ω*). The 20-item Social Interaction Anxiety Scale (SIAS, *ω* = .91; Mattick and Clarke, 1998) was used to assess self-reported social interaction anxiety, and was administered at baseline, sessions 4, 8, 12, and 1- and 6-month follow-up. The 5-point response scale is Not at all (0), Slightly (1), Moderately (2), Very (3), or Extremely (4) characteristic of me, with total scores ranging from 0 to 80. The 12-week test-retest reliability is .92 (Mattick and Clarke, 1998). The SCID-5 was administered by blind assessors to all participants at baseline and again at 1- and 6-month follow-up. Principal and comorbid diagnoses at baseline were re-assessed at each follow-up. To ascertain current diagnostic status and avoid overlaps with previous assessment time points, participants were asked to reflect on how they had been feeling “in the past few weeks” at 1- and 6-month follow-up. Severity of diagnoses was rated by clinicians on an 8-point scale at baseline and 1- and 6-month follow-up, with a score of 4 (at least occasional symptoms that definitely interfere with functioning) designated as the minimum threshold for a diagnosis.

*Secondary outcome measures.* The following measures were administered at baseline, each session, and at follow-ups and all were highly reliable: the 8-item PROMIS depression (*ω* = .99) and anxiety (*ω* = .94) scales (Pilkonis *et al.*, 2011) measure symptom severity over the previous 7-days on a 5-point rating scale from Never (1), Rarely (2), Sometimes (3), Often (4), to Always (5); the Brief Fear of Negative Evaluation scale – straightforwardly worded (*ω* = .92, BFNE, Rodebaugh *et al.*, 2004) is an 8-item self-report measure of fear or concern about being negatively evaluated on a 5-point scale from Not at all (1), Slightly (2), Moderately (3), Very (4), to Extremely (5) ; Fear of Positive Evaluation scale (FPE,*ω* = .81, Weeks *et al.*, 2008) is a 10-item measure (two reverse scored items are excluded from total scores) assessing fear and distress related to positive evaluation from others on a 10-point rating scale from Not at all true (0) to Very true (9). The Social Phobia Scale (SPS, *ω* = .97; Mattick and Clarke, 1998), a 20-item measure of performance anxiety, was administered in sessions 1, 4, 8, 12, and follow-ups and is assessed on the same scale as the SIAS. Additional secondary measures (McEvoy *et al.*, 2017) will be reported elsewhere.

**References**

Mattick, R. P. & Clarke, J. C. (1998). Development and validation of measures of social phobia scrutiny fear and social interaction anxiety. *Behav Res Ther* 36, 455-70.

McEvoy PM, Erceg-Hurn DM, Saulsman LM, Thibodeau MA. Imagery enhancements increase the effectiveness of cognitive behavioural group therapy for social anxiety disorder: a benchmarking study. *Behav Res Ther.* 2015;65:42-51.

McEvoy, P. M., Moulds, M. L., Grisham, J. R., Holmes, E. A., Moscovitch, D. A., Hendrie, D., Saulsman, L. M., Lipp, O. V., Kane, R. T., Rapee, R. M., Hyett, M. P. & Erceg-Hurn, D. M. (2017). Assessing the efficacy of imagery-enhanced cognitive behavioral group therapy for social anxiety disorder: Study protocol for a randomized controlled trial. *Contemp Clin Trials* 60, 34-41.

McEvoy PM, Saulsman LM. Imagery-enhanced cognitive behavioural group therapy for social anxiety disorder: a pilot study. *Behav Res Ther.* 2014;55:1-6.

McEvoy, P. M., Saulsman, L. M. & Rapee, R. M. (2018). *Imagery-Enhanced CBT for Social Anxiety Disorder*. Guildford Press: New York.

McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med (Zagreb).* 2012;22(3):276-282.

Pilkonis, P. A., Choi, S. W., Reise, S. P., Stover, A. M., Riley, W. T., Cella, D. & Group, P. C. (2011). Item banks for measuring emotional distress from the Patient-Reported Outcomes Measurement Information System (PROMIS(R)): depression, anxiety, and anger. *Assessment* 18, 263-83.

Rapee RM, Gaston JE, Abbott MJ. Testing the efficacy of theoretically derived improvements in the treatment of social phobia. *J Consult Clin Psychol.* 2009;77(2):317-327.

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. *Psychol Assess* 16, 169-81.

Weeks, J. W., Heimberg, R. G. & Rodebaugh, T. L. (2008). The Fear of Positive Evaluation Scale: assessing a proposed cognitive component of social anxiety. *J Anxiety Disord* 22, 44-55.