Attenuated psychophysiological reactivity following single-session group imagery rescripting versus verbal restructuring in social anxiety disorder: Results from a randomized controlled trial

Supplementary Material

Overview of Primary Self-Report and Diagnostic Measures

The following overviews the primary self-report outcome measures administered at pre- and post-treatment assessment points: The 14-item Social Interaction Phobia Scale (SIPS) [1] measures anxiety experienced in social situations and was administered to all participants before and after treatment (and at follow-up 2 weeks after treatment), with total scores serving as the primary self-report outcome measure for our study. The Cognitive Avoidance Questionnaire (CAQ) [2] assesses an individual's use of five cognitive strategies in response to threatening thoughts; we administered the Transformation of Images into Thoughts subscale of the CAQ in the current study, which assesses the use of verbal-linguistic activity to suppress imagery-related arousal, a hypothesized mechanism of maintenance of emotional disorders [3]; the Negative Self Portrayal Scale (NSPS) [4], assessing the extent to which specific self-attributes elicit concern about self-presentation in anxiety-provoking social situations; the Repetitive Thinking Questionnaire (RTQ) [5, 6] a 10-item measure of state repetitive negative thinking; and the SAD, major depressive disorder (MDD), and suicidality modules of the Mini-International Neuropsychiatric Interview (MINI) [7]. The MDD module allowed analysis of comorbid depressive symptomatology (in those with a principal SAD diagnosis) between groups in the current study (there were no differences between groups in the proportion of individuals with a diagnosis of MDD, as reported in the main text). The suicidality module was administered to judge suicide risk (those judged as high risk were not eligible to participate and were referred on to appropriate services). Additional outcome measures including memory modification following the intervention, self-rated anxiety (SUDS) and clinician-rated anxious behaviors during the speech task, and willingness to perform another speech task are overviewed below.

Physiological acquisition parameters

Physiological responses were recorded at a sampling rate of 1000 Hz with a BioPac MP150 System (running AcqKnowledge 3.9.1). EDA responding was acquired (with a BioPac GSR100C module) using two 8mm Ag/AgCl electrodes placed on the thenar and hypothenar eminences of the participants' non-dominant hand, with responses DC amplified at 5 μS/Volt. ECG was collected using three electrodes, one on the upper left, and one on the upper right chest (below the clavicles), and one medial to the anterior-superior iliac spine. The ECG was recorded with a BioPac ECG100C amplifier and bandpass filtered between 0.5 to 35 Hz (gain = 500).

Psychophysiology Pre-Processing

Time series data were obtained for both ECG and EDA for each subject and for each TSST condition (e.g., baseline, speech task, post-speech task) at pre- and post-treatment. Raw data were plotted to allow visual inspection, and segments with substantial artefact (e.g., poor contact of ECG/EDA electrodes; signal drift) were discarded before initial pre-processing (57/348 files; 16.38%) and subsequently treated as missing in analyses. ECG preprocessing and HRV estimation followed established guidelines [8]. The Pan-Tompkins algorithm [9] was implemented to extract the amplitude and temporal index of QRS complexes. R-R-intervals were calculated as the difference (in ms) between successive R waves from the QRS complexes. The *ARTiiFACT* toolbox [10] was used to detect artefacts from the distribution of R-R intervals and remove artefacts beyond the first and fourth quartile. The standard deviation of the interquartile range was estimated from the data (passing initial visual inspection), with artefact components estimated via interpolation of neighboring R-R-intervals. The accuracy of the preprocessing algorithm was judged via visual inspection of the corrected R-R interval data; 90% of the files (across all conditions) were deemed to be free of artefacts, and 37 files (~10%) required further cleaning (i.e., manual removal of multiple ectopic beats). 10% of files were independently crosschecked (author GAA) largely confirming the files as artefact free; however, 2 files/conditions from one participant in the waitlist group were judged as being inappropriate for inclusion (artefact in raw data files was too extensive to allow accurate preprocessing). This participant's data was hence treated as missing.

We estimated time- and frequency-domain HRV parameters on all valid data files for use in analyses, including the median of the R-R intervals (henceforth referred to as medianRR), mean heart rate (meanHR), standard deviation of the normal-to-normal intervals (i.e., intervals between successive QRS complexes that result from sinus node depolarizations; SDNN), root mean square of the sum of differences between successive NN intervals (RMSSD), as well as low frequency (0.04 – 0.15 Hz) and high frequency (0.15 – 0.4 Hz) power measured in absolute values (ms^2; LFabs and HFabs). We estimated two indices of skin conductance, i) mean skin conductance level (mSCL), and ii) number of skin conductance responses (SCRs) above a given threshold (.05 μS – detailed below), across each condition for each individual participant (see Table 1). Mean SCL was calculated in MATLAB. SCRs were estimated using the *Ledalab* toolbox [11]. *Ledalab* implements a model-based approach to estimating central sympathetic activity from observed skin conductance responses [12]. In brief, we used continuous decomposition analysis (CDA) to invert standard biophysical models of skin conductance (i.e., sudomotor nerve (SN) activity → SCRs), hence providing a model-based estimate of SN activity (i.e., SCRs →SN). EDA data were downsampled to 10Hz before estimation. Trough-to-peak (TTP) analysis was used to estimate the number of peaks in the inverted data using a minimum amplitude criterion of .05 μS.

The physiological correlates of HRV and skin conductance parameters have previously been overviewed in extensive detail [13-15]. In brief, the psychophysiological parameters used in this study reflect dynamic modulation of central and peripheral nervous systems in response to changing internal and environmental demands. The cardiovascular system in particular is designed to self-organize, via dynamic interactions with neural systems, in response to a range of internal and external stimuli, including metabolic imbalance, and changes in physical and mental activity [15]. Sophisticated analysis of heart rate variability signals provides a comprehensive index of adaptability of the autonomic nervous system. The parameters indexing beat-to-beat variation in the current study (medianRR, SDNN, RMSSD) and high frequency cyclical HRV (HFabs) are largely parasympathetically mediated whereas it is suggested that sympathetic activity is an important mediator of low frequency cyclical HRV (e.g., LFabs) [15], although the exact mechanisms that contribute to HRV are still a topic of debate [16]. Electrodermal activity (mean SCL and SCRs) reflects activity of the sympathetic branch of the autonomic nervous system [14]. Changes in heart rate (meanHR) arise from interactions between sympathetic

and parasympathetic responses. For example, a reduction in parasympathetic drive accompanied by increased sympathetic nerve activation results in increased heart rate and vice versa, whereas coactivation or coinhibition of both systems can result in alternate response patterns (i.e., increased, decreased or no change in heart rate) [17].

Analyses of physiological outcomes focused on differences between the groups on *reactivity* and *recovery* indices. Reactivity is the magnitude of psychophysiological responsiveness elicited by aversive, challenging, or engaging laboratory tasks [19]). Recovery refers to changes in stressor-induced psychophysiological responses following stressor termination [20]). In the current study, we defined reactivity as the mean difference in each physiological parameter before and during the TSST; recovery was defined as the mean difference in physiological parameters during and after the TSST. For example, at pre-treatment meanHR scores for the WC group were 79.10 before the speech task (TSST baseline) and 85.61 during the speech task; the average reactivity score was therefore 85.61 – 79.10 = 6.51. At post-treatment, the mean reactivity value was 5.76. The change in reactivity values for the WC group between pre and post-treatment was 6.51 – 5.76 = 0.75. In contrast, the mean change in reactivity for the IR group was 9.37, and the mean difference in reactivity change between the WC and IR groups was 0.75 – 9.37 = - 8.63, which was statistically significant (p = .004) and large in magnitude, corresponding to a standardized difference between means of d = -0.73.

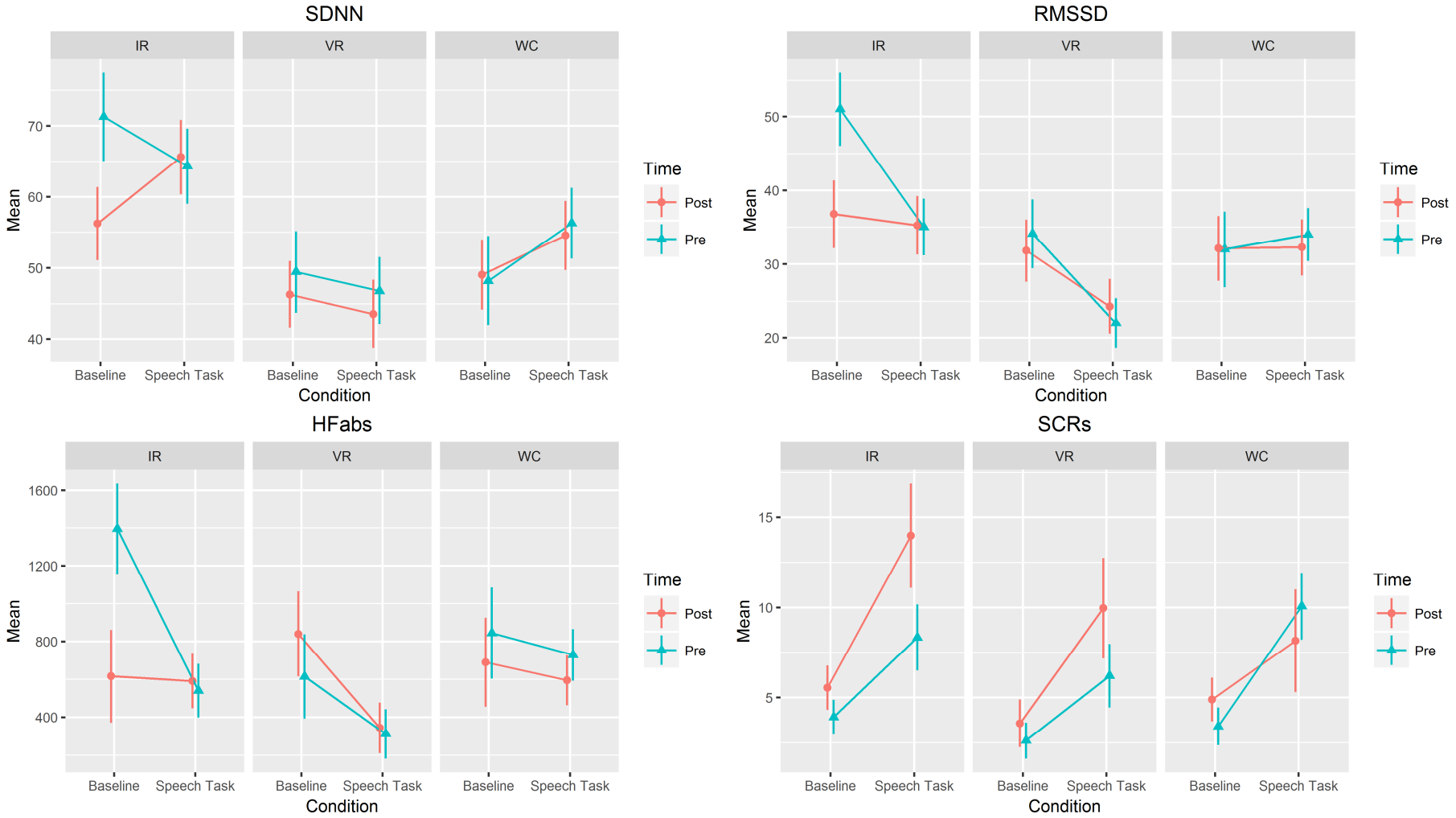
Single Session Group IR and VR Protocols

For both groups, sessions (each lasting 90 minutes) began with psychoeducation about social anxiety and a treatment rationale. Participants were then required to identify a negative socially evaluative situation as either images (IR) or thoughts (VR) that shaped how they currently feel in social situations, as well as their core beliefs about the situations. Participants in the IR group were guided to rescript images associated with this memory through visualization i) from their younger-self perspective (i.e., around the time of encoding the initial fear memory), ii) from the perspective of their "older self", whose role it was to intervene to assist the younger self (i.e., contributing adaptive information that may not have been accessible at encoding), and iii) from the perspective of their younger self observing the "older self" intervening. Participants in the VR group were guided to restructure thoughts associated with the fear memory using thought challenging (i.e., reality testing). This process involved considering evidence contrary to beliefs around the original event, generating

alternative thoughts and predictions, and constructing more realistic judgments about the probability and consequences around perceived evaluation. Throughout each intervention, participants were invited to reflect on their affect and bodily sensations, and to identify any shifts in meaning of the original fear memory. Upon completion, participants in both groups were asked to identify three take-home messages from the intervention.

Additional Details about Statistical Analyses

Mixed-model repeated measure (MMRM) statistical analyses were conducted using *R*. The analyses were performed using the 'gls' function of the R package nlme [21]. Restricted maximum likelihood was the method of estimation. A heterogeneous unstructured (co)variance structure was used to model the within-subject errors. We used the R package emmeans [22] to calculate estimated marginal means at pre- and post-treatment, and construct contrasts between the group means.



Supplementary Figure 1. Heart rate variability and skin conductance response parameter values for baseline and TSST speech task before and after treatment across groups for SDNN, RMSSD, HFabs and SCRs. Error bars represent standard errors. *HFabs: High frequency power (in absolute values – ms^2); SCRs: Count of skin conductance responses (> .05 µS); SDNN: Standard deviation of all normal R-R intervals; RMSSD: Root mean square of summed differences between successive normal R-R intervals.*

Assessment of Anxiety and Distress during TSST

Two additional self-report measures of distress and anxiety were administered to participants following the TSST at pre- and post-treatment time points. The *Subjective Unit of Distress Scale (SUDS)* is a self-report measure of anxiety, rated on a 0-10 point scale. A score of 0 represents no fear, while 10 represents the most fear the individual has ever felt. Participants were asked to rate their peak level of anxiety they experienced during the speech task ('peak SUDS') at pre- and post-treatment. In addition, participants were also asked to rate their willingness to perform another speech task (0 = not at all, 10 = completely willing) at pre- and post-treatment, providing a parsimonious measure of anxious avoidance. Trained assessors blind to group allocation rated overt anxious behaviors during the speech task (via videotape of TSST sessions) for participants at pre- and post-treatment. The following behaviors: 'soft speech', 'fast speech', 'hides face', and/or 'excuses about presentation quality', were coded as present or absent, and summed for analysis to provide an objective measure of anxious behavioral responding (as a count varying between 0 and 4, with higher values indicating more anxious behaviors). Group x time contrasts were estimated across these outcome measures as per the other self-report measures in the main manuscript. At the post-treatment assessment point, participants in the IR and VR groups were asked to rate what memory or perspective was most dominant in their mind (contrasting the original memory/perspective from the past situation they had *before* the intervention, and the new memory/perspective that they developed *after* the intervention), spanning 1 = old memory completely dominates, 4 = old and new memories are equally dominant, to 7 = new memory completely dominates. Higher scores on this scale thus reflect greater use of the new memory/perspective developed as a result of the intervention. Group contrasts (comparing IR and VR groups at post-treatment) were estimated using the 'gls' function in R for this latter scale. Supplementary Table 1 contains the means, standard errors, and contrasts testing whether mean changes in self-reported or observer-rated anxious responding and impact on memory/perspective modification between pre- and post-treatment differed between the groups. There were no statistically significant between group differences in the change of peak SUDS during speech task, willingness to perform another speech, or observer-rated anxious behaviors between pre- and post-treatment. There were no statistically significant differences between IR and VR groups in the type of memory most prominent at post-treatment with both groups, on average, appearing to use a combination of new and old memories (mean scores from 3.60 to 3.93 in IR and VR, respectively).

Supplementary Table 1. Estimated marginal means, standard errors and group x time interaction tests for speech task anxiety and memory modification.

	IR		VR		WC						
Measure	**Pre**	**Post**	**Pre**	**Post**	**Pre**	**Post**	**Contrast**	***M* Diff**	**S.E.**	**p**	**d**
Peak SUDS	5.35 (0.52)	4.04 (0.56)	5.23 (0.39)	4.02 (0.59)	5.68 (0.39)	5.53 (0.41)	IR vs. WC	1.16	0.77	0.136	0.62
							IR vs. VR	0.10	0.86	0.905	0.05
							VR vs. WC	1.06	0.75	0.161	0.57
Speech Willingness	6.94 (0.82)	6.29 (0.72)	5.95 (0.85)	5.76 (0.75)	5.37 (0.69)	5.26 (0.54)	IR vs. WC	0.54	0.98	0.582	0.16
							IR vs. VR	0.45	1.13	0.689	0.13
							VR vs. WC	0.09	1.02	0.929	0.03
Anxious Behaviors	0.82 (0.17)	0.59 (0.19)	0.64 (0.16)	0.41 (0.18)	0.79 (0.18)	0.53 (0.16)	IR vs. WC	-0.03	0.23	0.905	-0.04
							IR vs. VR	0.008	0.23	0.973	0.01
							VR vs. WC	-0.04	0.23	0.874	-0.05
Memory	-	3.60 (0.27)	-	3.93 (0.26)	-	-	IR vs. VR	-0.33	0.38	.384	-0.23

SUDS: Subjective Units of Distress Scale; *M* Diff: contrast mean difference.

Supplementary References

1. Carleton RN, Collimore KC, Asmundson GJ, McCabe RE, Rowa K, Antony MM: Refining and validating the Social Interaction Anxiety Scale and the Social Phobia Scale. Depress Anxiety 2009;26:E71-81.
2. Sexton KA, Dugas MJ: The Cognitive Avoidance Questionnaire: validation of the English translation. J Anxiety Disord 2008;22:355-370.
3. Borkovec TD, Alcaine O, Behar ES: Avoidance theory of worry and generalized anxiety disorder; in Heimberg R, Mennin D, Turk C (eds): Generalized anxiety disroder: Advances in research and practice. New York, Guildford, 2004, pp 77-108.
4. Moscovitch DA, Huyder V: The negative self-portrayal scale: development, validation, and application to social anxiety. Behav Ther 2011;42:183-196.
5. McEvoy PM, Mahoney AE, Moulds ML: Are worry, rumination, and post-event processing one and the same? Development of the repetitive thinking questionnaire. J Anxiety Disord 2010;24:509-519.
6. McEvoy PM, Tribodeau M, Asmundson G: Trait repetitive negative thinking: A brief transdiagnostic assessment. J Exp Psychopathol 2014;5:382-398.
7. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, Hergueta T, Baker R, Dunbar GC: The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatry 1998;59:22-33.
8. Quintana DS, Alvares GA, Heathers JA: Guidelines for Reporting Articles on Psychiatry and Heart rate variability (GRAPH): recommendations to advance research communication. Transl Psychiatry 2016;6:e803.
9. Pan J, Tompkins WJ: A real-time QRS detection algorithm. IEEE Trans Biomed Eng 1985;32:230-236.
10. Kaufmann T, Sutterlin S, Schulz SM, Vogele C: ARTiiFACT: a tool for heart rate artifact processing and heart rate variability analysis. Behav Res Methods 2011;43:1161-1170.
11. Bach DR, Friston KJ: Model-based analysis of skin conductance responses: Towards causal models in psychophysiology. Psychophysiology 2013;50:15-22.
12. Bach DR: A head-to-head comparison of SCRalyze and Ledalab, two model-based methods for skin conductance analysis. Biol Psychol 2014;103:63-68.

13. Heart rate variability. Standards of measurement, physiological interpretation, and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. Eur Heart J 1996;17:354-381.
14. Boucsein W: Electrodermal activity. Berlin, Springer, 1992.
15. Pumprla J, Howorka K, Groves D, Chester M, Nolan J: Functional assessment of heart rate variability: Physiological basis and practical applications. Int J Cardiol 2002;84:1-14.
16. Berntson GG, Quigley KS, Norman GJ, Lozano DL: Cardiovascular psychophysiology; in Berntson GG, Cacioppo JT, Tassinary LG (eds): Handbook of psychophysiology. Cambridge, Cambridge University Press, 2016, pp 183-216.
17. Berntson GG, Cacioppo JT, Quigley KS: Autonomic determinism: The modes of autonomic control, the doctrine of autonomic space, and the laws of autonomic constraint. Psychol Rev 1991;98:459-487.
18. Goldstein DS, Bentho O, Park MY, Sharabi Y: Low-frequency power of heart rate variability is not a measure of cardiac sympathetic tone but may be a measure of modulation of cardiac autonomic outflows by baroreflexes. Exp Physiol 2011;96:1255-1261.
19. Kamarck TW, Lovallo WR: Cardiovascular reactivity to psychological challenge: Conceptual and measurement considerations. Psychosom Med 2003;65:9-21.
20. Haynes SN, Gannon LR, Orimoto L, O'Brien WH, Brandt M: Psychophysiological assessment of poststress recovery. Psychol Assess 1991;3:356-365.
21. Pinheiro J, Bates D, DebRoy S, Sarkar D, R Core Team: nlme: Linear and Nonlinear Mixed Effects Models, 2017.
22. Lenth R, Love J, Herve M, R Core Team: emmeans, 2017.