Official title: Leveraging Computational Social Sciences and Natural Language Processing to Optimize Engagement and Response to Low-intensity CBT for Depression and Anxiety

NCT number: NCT04870099

Date of the document: October 17, 2022

Statistical analysis plan

A more thorough analysis plan is reported in the Open Science Foundation page for this study https://osf.io/j32uw/ and in the main outcome paper (Lorenzo-Luaces et al., 2023).

Lorenzo-Luaces, L., Howard, J., De Jesús-Romero, R., Peipert, A., Buss, J. F., Lind, C., ... & Starvaggi, I. (2023). Acceptability and outcomes of transdiagnostic guided self-help bibliotherapy for internalizing disorder symptoms in adults: A fully remote nationwide open trial. *Cognitive Therapy and Research*, *47*(2), 195-208.

Sample size justification and power. We initially powered the trial to be able to detect a statistically significant difference in engagement relative to the 65.1% engagement rate from the meta-analysis by Van Ballegooijen et al. (2014). We used an online (https://sample-size.net/sample-size-conf-interval-proportion/) calculator (Kohn & Senyak, 2021) to estimate the required sample size to test for a statistically significant difference between a sample proportion and the expected value (i.e., 65.1%) at a p value < .05. This analysis suggested we needed to recruit at least 95 individuals to have an adequately powered acceptability trial. We ultimately recruited more participants than was necessary (n=141) because there was uncertainty about our ability to retain all individuals and because, during the initial recruitment period, the PI (LL-L) obtained funding for an extension assessing the effects of GSH-CBT on natural language metrics of social media data. That subcomponent of the study (i.e., whether GSH-CBT produces effects detectable via social media) is not being used in the current analysis. We used G*Power 3.1 to estimate what magnitude of within-person changes we were powered to detect with 141 participants. The results of that analysis suggested that we could detect small-medium within-

person changes (d = 0.25, at p < .05 and power of 80%) with the achieved sample size (i.e., n = 141).

Analytic strategy. All analyses were conducted in R version 4.1.2 (see R Core Team, 2013) using the R Studio Graphic User Interface. All code and deidentified data are available on the Open Science Foundation (OSF) site https://osf.io/j32uw/. First, we report the percentage of participants who progressed from beginning the survey to the end of the study.

We followed an intent-to-treat (ITT) approach, analyzing data from all individuals who reached the onboarding call to confirm trial participation. There are at least two other definitions of who entered the trial that are possible in these types of online trials. One is less conservative, considering as study subjects only individuals who completed at least one post-baseline assessment after the baseline eligibility survey. Another is more conservative and considers as participants all individuals who qualified for the study regardless of whether or not we could reach them to confirm participation. We also used two different definitions of completers: completing at least 50% of GSH-CBT sessions and reaching participants to be study completers if they completed at least 50% of their GSH-CBT sessions. Given variability in how study entry and completion could be defined, we conducted a "multiverse" analysis to assess the acceptability of the intervention (see Steegen et al., 2016). In a multiverse analysis, a researcher conducts and presents all possible ways of analyzing data (i.e., here, the engagement rate). Multiverse analysis has been recommended as a way of increasing transparency in psychological sciences by presenting readers multiple versions of the data, as opposed to simply presenting the analyses that give the result with the most favorable effects. We chose a non-inferiority margin of +/- 10% to determine if observed engagement rates were equivalent to the rates from the meta-analysis by Van Ballegooijen et al. (2014).

We report baseline demographic and clinical characteristics by presenting means and standard deviations for continuous variables and percentages for categorical variables (see Table 1). To assess the preliminary efficacy of the intervention, we conducted mixed regression models using the lme4 and lmerTest packages in R to regress internalizing symptoms (K6), well-being (WHO-5), ERQ-reappraisal, and ERQ-suppression on time in GSH-CBT. We coded the time variable by dividing each week by six, the total number of weeks in the study.