PROTOCOL

We will implement the adaptive social VR game and determine the optimal dose. After we construct each subject’s affective model with the machine-learning algorithm, we will test the effects of the adaptive VR games in 40 medicated outpatients with schizophrenia (SZ). Demographically matched control (CO) subjects (n=16) will only participate in the affective modeling and social and cognitive assessments but not in VR training (see Human Subjects section for details). To determine the optimal dose of VR training, it is necessary to conduct a valid assessment for social functioning pre- and post-treatment. Details of assessments are outlined below. Optimal dose of physiology-based, adaptive VR intervention is unknown but Tsang & Man\textsuperscript{53} reported that just ten 30-minute sessions of conventional, non-adaptive VR training was sufficient to obtain an improvement in SZ. In our study, we will examine the effects of social VR training on social and cognitive measures, pre- and post-treatment. At baseline (t1), the PI’s will use stratified randomization to assign SZ to two dose conditions: high vs. low. Stratified randomization is used to ensure that potential confounding factors (e.g. age, gender) are evenly distributed between groups. If possible, we will try to match the two groups on IQ, and the Social Functioning Scale\textsuperscript{79}, a broad measure of social functioning, but if it cannot be done, we will statistically adjust for them when we test for treatment effects. Everybody will train twice a week but the low dose group will train for 30 mins and the high dose group for 2 x 30 mins per visit. VR task performance will be recorded at each visit. At t1, social, cognitive and symptom assessments will be conducted. The research staff who administer assessments are blind to the group assignment and SZ are unaware which group they are in. The research assistants who supervise VR sessions will NOT conduct assessments because they will be aware of the dosing. After t1, subjects play the VR game for 2.5 weeks, at which point (t2), we repeat the social, cognitive and symptoms assessments plus a satisfaction survey. Then they train for another 2.5 weeks. At 5 weeks (t3), we repeat the assessments and the satisfaction survey. The behavioral measures have been used in intervention trials to be reliable across time with minimal practice effects, or offer alternate forms\textsuperscript{85,100}. The satisfaction survey will solicit user feedback to improve our protocol in the future.

**VR Training:** SZ will play the “Day Out” game in the lab. Biosensors and eye tracking will be used. To boost adherence, we will call/text 2 hours before the visit. They will be paid after each visit and receive a bonus upon reaching the next level of difficulty in the game, and when they complete the training.

**ASSESSMENTS.**

**VR Task Measures:** The VR game has built-in performance measures, including but not limited to the total time spent on completing the specified task (e.g., find out avatar’s birthday at the bus stop), number of errors following the sequential steps, number of successful social interactions with avatars and deviation from specified scenarios. The number of difficulty levels mastered by the participant will determine the overall performance. We will track the VR scores per session to determine the rate of learning.

**Social Assessments:**

(1) **Social Functioning Scale (SFS)\textsuperscript{79}** has 79-items with 7 subscales: social engagement, interpersonal communication, daily living activities, recreation, social activities, independent living and employment. Raw scores of the subscales are
converted to scaled score (m=100, SD=15). (2) **Social Emotional Task from CogState**\(^2\) (SET-CS) is a facial emotion recognition task within CogState (see below). (3) **Bell Lysaker Emotion Recognition test (BLERT)**\(^{100}\). Subjects view 10s vignettes of an actor talking about life. After each vignette, subjects select one emotion that was portrayed. There are 3 different stories portrayed in 7 emotions (21 total). Correct responses are summed. (4) **Revised Eyes Test (EYES)**\(^{127}\) assesses the theory of mind\(^{87}\). It asks subjects to decide the mental state of a person from an image of his/her eyes. For each set of eyes, subjects select one word out of 4 choices that best describes what the person in the picture is thinking or feeling. There are 36 trials. These measures were selected for their validity, reliability, and coverage of important domains of social functioning\(^1,130\). Our VR game simulates and exercises social interactions, thus, targeting emotion recognition (BLERT, EYES, SET-CS), and theory-of-mind\(^{87}\) (EYES, BLERT). Practicing these skills will likely lead to broad, beneficial consequences for social functioning (SFS).

**Cognitive Assessments:** (1) **Wechsler Abbreviated Scale of Intelligence (WASI)**\(^{80}\) will estimate IQ at t1. It will not be repeated at t2 or t3. Baseline IQ may interact with the efficacy of social interventions\(^{20,81}\). (2) **CogState**\(^{82}\) is a computerized, standardized battery for assessing cognitive abilities (cogstate.com). It reliably detects cognitive changes over time, and shows minimal practice effects with repeated administrations\(^{83,84}\). It is widely used to measure cognition in treatment trials\(^{84-86}\). Results are automatically scored. It takes 30 minutes.

**Clinical Symptoms:** SAPS\(^74\) and SANS\(^75\) will be used to measure the severity of symptoms in SZ.

**Satisfaction Survey:** To obtain feedback about VR experience, we will give a 9-item survey that asks about difficulties, hurdles, enjoyment and motivational factors, and solicits suggestions.

**DATA ANALYSIS PLAN**

Does the VR intervention elicit significant pre-post change on the targeted social attention, measures of social functioning (SFS, SET-CS, BLERT, EYES) and clinical symptoms (SAPS and SANS)? What are the specific treatment parameters that provide the optimal benefit when cost-benefit and other considerations are taken into account? To address training effects on the social assessments, we will use a linear mixed effects model (LMM) analytic approach because it allows for maximally flexible modeling of effects and parameters and for a principled treatment of missing data\(^{128,141}\). To address this question, we will conduct repeated measures LMM analyses. Specific contrasts will directly assess: 1) The linearity of the slope of change within each group; 2) Pairwise differences among time points within each group; and, 3) Group differences in linearity and pairwise differences. Because our power calculations indicate that the most powerful tests of group differences will be provided by an ANCOVA framework (with pre-treatment values serving as the covariate)\(^{112,134}\), we will also use this analytic approach to compare the two dose conditions. Although it is important to minimize Type 2 errors in an initial study, we will also generate multiplicity-corrected \(p\) values across the 5 primary measures using simulation-based approaches\(^{142}\). A similar approach will be used to analyze the clinical symptom and other measures (e.g., CogState) although a generalized linear mixed effects (GLMM) model may be needed to accommodate non-normal distributional properties of the measures\(^{101}\). With a particular emphasis on the set of social measures,
we will also conduct analyses that assess effects on the set as a whole. Depending on the correlational structure of the measures, we will either: 1) Form 1 or 2 linear composites using prior findings on the dimensional structure of the measures and our own analyses; or 2) Conduct a multivariate analysis of variances using the flexible facilities of mixed-effects software to accommodate such models. We will assess clinically meaningful change by comparing the post-treatment SZ and normative distribution for each group. To provide a more fine-grained assessment of the proximal effects of training we will assess change over days on the VR task measures (e.g., # of difficulty levels mastered) to assess the rate of learning across, within, and between the dose conditions. Generalized linear and generalized nonlinear mixed effects models will be used to carefully assess the functional form of change over time using several approaches: a) Polynomial growth models; b) Penalized spline analyses; and, c) Non-linear mixed effects models directly estimating parameters particularly relevant to learning. It is important that analyses also address missing data and attrition because: (1) Minimization of drop-out is one of the goals of the VR intervention and will be an important determinant of the dose and duration to use in the R33 phase; and (2) To minimize biases, it is important to use a principled approach to the treatment of missing data on the other outcome measures. We will: (1) Conduct logistic regression and, if feasible, discrete survival analyses predicting drop-out based on weekly data; (2) Use pattern-mixture models and related approaches to estimate treatment effects in the presence of drop-out and assess how different patterns of attrition might moderate effects; and, (3) Try to include in models auxiliary variables that predict drop-out. Supplementary analyses will also test whether performance changes during the training phase mediate the effects of the intervention conditions on social functioning and symptom measures.

BIBLIOGRAPHY


23. VISN 5 Mental Illness Research, Education, and Clinical Center (MIRECC) - VA Social Skills Training Program. [http://mirecc.va.gov/visn5/training/social_skills.asp](http://mirecc.va.gov/visn5/training/social_skills.asp)


