Enhancing Social Competence in Adults with Autism

Statistical Analysis Plan (SAP)

NCT04349644

Document Date: April 26, 2019

Document Source: Statistical Design Plan from MH120149

4.4. Statistical Design and Power

Post-attrition Sample Size Used to Test Aims and Compute Statistical Power. We propose to recruit **approximately 46 participants to enroll 40 participants** over the first 2 years of the project with the last year being used to complete the treatment and assessment for the last wave of recruitment. Using a 10% attrition rate, which is the maximum attrition rate to date of past SENSE Theatre trials, the post-attrition sample size will be **40** (equally distributed between experimental and control groups). The analyses and power analyses are the based on post-attrition sample size and past findings involving the primary dependent variables.

Aim 1: To Assess Preliminary Effectiveness via Target Engagement of Social Salience.
Hyp 1.1. Adults with ASD in the EXP group will demonstrate significantly faster growth in IMF (primary DV), CMF (secondary DV) or WMS-F (secondary DV) than adults with ASD in the WLC group, verifying the intervention had predicted effect on the target.
Statistical Analysis: IMF is the primary DV because it uses the same procedure, stimuli, and metric that was used with our previous successful efficacy demonstrations of SENSE Theatre. The CMF and WMS-F are measures of the same constructs on which SENSE Theatre had effects in children and adolescents, but are measured with procedures appropriate for adults. Thus, the latter two variables are considered secondary. The effect on the primary and secondary dependent variables will be tested in separate analyses for posttest and follow-up measures. In each analysis, we will use multiple linear regression in which group (experimental = EXP or waitlist control = WLC) is analyzed as the dummy-coded independent variable, and the baseline (pre-intervention) variable is analyzed as a covariate (assuming homogeneity of slopes assumption is met). By including the baseline value as a covariate, the coefficient for group becomes the between-group (EXP vs. WLC) difference in the pre-test adjusted posttest (or follow-up) scores (i.e., a measure of change similar to residualized gain scores). This method is often superior to a method that uses raw change scores directly because it improves the statistical power of the significance test and improves the accuracy of effect size estimates relative to between-group differences in raw change scores [1]. If significant difference between groups are detected on other baseline variables that correlate with the outcome, those baseline variables will also be controlled in the multiple linear regression model, assuming appropriate assumptions are met. We will assess the homogeneity of slopes assumption by testing the interaction effect between baseline variable and treatment group assignment. If there is a significant interaction, then we will assess the treatment effect at various levels of the pre-intervention variable. If other regression assumptions (e.g., homoscedasticity, normality of residuals) are not met, robust standard error estimates will be used to test parameters in analogous models.

Power analyses for tests of hypothesis associated with Aim 1. Using a one-tailed test (based on the a priori predicted direction of differences), 80% power, 5% type 1 error, and a post-attrition sample size with equal groups of **40** (**20** in each group), the minimal detectable effect size is an R^2 of 0.14 (i.e., standardized mean difference of .80). The effect size of the between-group difference on the pretest-adjusted posttest ERP measure of memory for faces in our most recent RCT of SENSE Theatre was an R^2 of .18 (i.e., standardized mean difference of .94). Therefore, it is feasible that we will detect the predicted SENSE Theatre effect on our primary measure of memory for faces with the proposed sample size.

Hyp 1.2. Alternative hypothesis, adults in the EXP group will show a direct effect of better social communication with novel peers (CASS).

It is plausible that participation in the theatre program more broadly increases salience of social stimuli, leading to enhanced performance on face memory and social communication, which ultimately impact functional outcomes. In this way, successfully interacting and performing with supportive expert social models (peers) increases intrinsic social interest directly contributing to enhanced interactive communication and behavior [2]. Thus, we will test this expanded salience model by including social communication (CASS aggregated score) in the aforementioned model.

Aim 2: Test the Link Between the Degree of Target Engagement (Social Salience), Social Interaction, Anxiety/Depression (STAI, ABCL, CDI) and Functional Outcomes of Social Skills (SRS, ABAS).
Hyp 2.1: Adults in the EXP group will demonstrate significantly more growth on functional outcomes.
Hyp 2.2: There will be a positive association between (a) memory for faces (IMF, CMF or WMS-F) at post-test or follow-up and (b) social skills (CASS involvement and rapport) at posttest or follow-up.

Hyp 2.3: There will be a negative association between (a) memory for faces (IMF, CMF or WMS-F) at post-test or follow-up and social awareness symptoms (SRS-2 awareness T score), or anxiety/depression (STAI, ABCL, CDI) at posttest or follow-up.

Statistical Analysis: The aforementioned statistical analyses in Hyp 1.1 will be followed for Hyp 2.1 with the functional outcomes as the dependent variable. The SRS-2 is the instrument from which we will derive our primary DV because it is the same instrument we have used to show the association between our primary measure of memory for faces and social symptoms in our past work on children and adolescents with ASD. From the SRS-2, the Awareness subscale and its T score will be the subscale and score type used to test this aim because this is the variable that correlates with the primary measure of memory for faces (i.e., the ERP measure). The other measures are secondary DVs because, while they measure similar or identical constructs as those for which SENSE Theatre has had effects in children and adolescents, but these are measured with procedures appropriate for adults in the present proposal. In two separate analysis (one for posttest and one for follow-up), we will examine the association of memory for faces with the primary functional outcome measure using linear regression. We will also explore how changes in social awareness symptoms are related to changes in memory for faces by modeling the social awareness outcomes for all timepoints as a function of the change in memory for faces and a timepoint dummy variable using a generalized estimating equation. To assess the effect of change in memory for faces on change in social awareness symptoms we will test the change in memory for faces by timepoint interaction. Similar analyses will be conducted for secondary DVs.

Power analyses for tests of primary dependent variable in Aim 2: A one-tailed test of the association between memory for faces and social awareness symptoms is reasonable because past evidence indicates that association is negative and the theory predicting the association indicates that the association should be negative. Using a one-tailed test, 80% power, and 40 post-attrition sample size, the minimally-detectable effect size for the association between memory for faces and social awareness symptoms is an R square value of 0.16. In our past work, the association between ERP-measured memory for faces and SRS-measured social awareness symptom severity was $R^2 = 0.15$. Thus, it is feasible that we will detect the expected association between memory for faces and social awareness symptom severity with a post-attrition sample size of 40.

Interpretative framework for findings detected in the context of multiple significance testing without alpha adjustment. The R33 funding mechanism is designed to enable the acquisition of preliminary data to aid the design of more robust tests of the efficacy of treatments. Adjusting alpha for multiple significance testing at this early stage would increase the probability of Type II error. Therefore, the results will be considered pilot given the inflation of type 1 error resulting from using multiple significance tests to address the same aim. The pilot findings will provide the preliminary data needed to plan for a future, larger RCT.

The aforementioned design is rigorous and the results from the analyses will set the stage to inform Go/No-Go decisions as to whether further development or testing of the intervention in a larger sample of adult participants with ASD is warranted.

Future Directions:

Ultimately, we hope to test the hypothesis that SENSE Theatre will have a proximal effect on social cognition that will act as a mediator on social behavior/functioning. Creating composites should enhance power by increasing the reliability and thus validity of construct estimates and reduce the number of significance tests. If confirmed, then such suggests that the theory and data must be mature enough for us to expect replication and thus a probable contribution to cumulative knowledge building.

For the R33, the sample size will not allow us to test a mediation model. Instead we will test the A path (SENSE Theatre effect on social cognition) and B path (the association between Social Cognition and Social Behavior controlling for treatment effects on Social Behavior) and the C path (SENSE Theatre effect on Social Behavior). These data can help create the foundation for an R01 with adults by: A) indicating whether the association among the component variables really is > .4 for each construct and if not indicating which component variables are associated as expected with other constructs. B) If the target is engaged and linked to functional outcomes that have clinical significance for adults with ASD, then the data will serve as a base for future R01 sample size to

power tests of the mediated model.

As in our current R01, a larger RCT will allow us to use a multivariate multilevel growth model (involving a linear growth model for memory and another linear growth model for social behavior). We would be able to specify a mediation pathway at the participant-level from treatment (x) through memory for faces (m) to social behavior (y). The indirect effect could then be tested using the 95% confidence interval around the product of the unstandardized coefficients for the (x→m) and (m→y) paths of the multilevel mediation model following procedures of Lachowicz, Sterba, and Preacher [3]. The confidence interval will be estimated using the Monte-Carlo procedure described in Preacher and Selig, [4].

References for Statistics

1. Frison, L. & Pocock, S.J. Repeated measures in clinical trials: analysis using mean summary statistics and its implications for design. *Statistical Medicine* **11**, 1685-1704 (1992).
2. Chevallier, C., Kohls, G., Troiani, V., Brodkin, E.S. & Schultz, R.T. The social motivation theory of autism. *Trends Cogn Sci* **16**, 231-239 (2012).
3. Lachowicz, M.J., Sterba, S.K. & Preacher, K.J. Investigating multilevel mediation with fully or partially nested data. *Group Process Interg* **18**, 274-289 (2015).
4. Preacher, K.J. & Selig, J.P. Advantages of Monte Carlo confidence intervals for indirect effects. *Communication Methods and Measures* **6**, 77-98 (2012).