HHS Public Access
Author manuscript
Schizophr Res. Author manuscript; available in PMC 2025 April 01.

Published in final edited form as:
Schizophr Res. 2024 April ; 266: 92–99. doi:10.1016/j.schres.2024.02.027.

Trajectories and Predictors of Response to Social Cognition Training in People with Schizophrenia: A Proof-of-Concept Machine Learning Study

Kathleen Miley[1,2], **Michael V. Bronstein**[2], **Sisi Ma**[3], **Hyunkyu Lee**[4], **Michael F. Green**[5,6], **Joseph Ventura**[6], **Christine I. Hooker**[7], **Mor Nahum**[8], **Sophia Vinogradov**[2]

[1]HealthPartners Institute, Minneapolis, Minnesota, USA

[2]Department of Psychiatry and Behavioral Sciences, University of Minnesota Medical School, Minnesota, USA

[3]Institute for Health Informatics, University of Minnesota, Minnesota, USA

[4]Department of Research and Development, Posit Science Inc., San Francisco, CA

[5]VA Greater Los Angeles, Los Angeles, CA

[6]Department of Psychiatry and Biobehavioral Sciences, University of California, Los Angeles, CA

[7]Department of Psychiatry and Behavioral Sciences, Rush University Medical Center, Chicago, IL

[8]School of Occupational Therapy, Hebrew University of Jerusalem

Abstract

Corresponding author: Kathleen Miley, PhD, MSN, mile0087@umn.edu.
Contributors:
Kathleen Miley,[1,2] Michael V. Bronstein,[2] Sisi Ma,[3] Hyunkyu Lee,[4] Michael F. Green,[5,6] Joseph Ventura,[6] Christine I. Hooker,[7] Mor Nahum,[8] & Sophia Vinogradov[2]
[1]HealthPartners Institute, Minneapolis, Minnesota, USA
[2]Department of Psychiatry and Behavioral Sciences, University of Minnesota, Minnesota, USA
[3]Institute for Health Informatics, University of Minnesota, Minnesota, USA
[4]Department of Research and Development, Posit Science Inc., San Francisco, CA
[5]VA Greater Los Angeles, Los Angeles, CA
[6]Department of Psychiatry and Biobehavioral Sciences, University of California, Los Angeles, CA
[7]Department of Psychiatry and Behavioral Sciences, Rush University Medical Center, Chicago, IL
[8]School of Occupational Therapy, Hebrew University of Jerusalem
Corresponding author:
Kathleen Miley, PhD, MSN
mile0087@umn.edu
Author contributions:
KM: Conceptualization, Methodology, Writing – original draft, Writing – review and editing; MB: Conceptualization, Data curation, Formal analysis, Methodology, Writing – original draft, Writing – review & editing; SM: Supervision, Writing – review & editing; HL: Investigation, Writing – review & editing; MG Investigation, Writing – review & editing; JV: Investigation, Writing – review & editing; Investigation; Writing – review & editing; MN: Conceptualization, Funding acquisition, Investigation, Methodology, Supervision, Writing – review & editing; SV: Conceptualization, Funding acquisition, Investigation, Methodology, Supervision, Writing – review & editing

Author disclosures: Hyunkyu Lee, PhD is a paid employee of Posit Science, the company which holds rights for the computerized SocialVille training, which is described in the manuscript.

Background.—Social cognition training (SCT) can improve social cognition deficits in schizophrenia. However, little is known about patterns of response to SCT or individual characteristics that predict response.

Methods: 76 adults with schizophrenia randomized to receive 8-12 weeks of remotely-delivered SCT were included in this analysis. Social cognition was measured with a composite of six assessments. Latent class growth analyses identified trajectories of social cognitive response to SCT. Random forest and logistic regression models were trained to predict membership in the trajectory group that showed improvement from baseline measures including symptoms, functioning, motivation, and cognition.

Results: Five trajectory groups were identified: Group 1 (29%) began with slightly above average social cognition, and this ability significantly improved with SCT. Group 2 (9%) had baseline social cognition approximately one standard deviation above the sample mean and did not improve with training. Groups 3 (18%) and 4 (36%) began with average to slightly below-average social cognition and showed non-significant trends toward improvement. Group 5 (8%) began with social cognition approximately one standard deviation below the sample mean, and experienced significant deterioration in social cognition. The random forest model had the best performance, predicting Group 1 membership with an area under the curve of 0.73 (SD 0.24; 95% CI [0.51-0.87]).

Conclusions: Findings suggest that there are distinct patterns of response to SCT in schizophrenia and that those with slightly above average social cognition at baseline may be most likely to experience gains. Results may inform future research seeking to individualize SCT treatment for schizophrenia.

Keywords

machine learning; precision medicine; psychosis; treatment response; cognitive remediation

1. Introduction

Individuals with schizophrenia experience deficits in all major domains of social cognition, including emotion and social perception, emotion processing, theory of mind, and attributional style (Savla et al., 2013). Although these deficits strongly predict social and occupational functioning (Fett et al., 2011), few interventions on them have robust empirical support. Social cognition training (SCT) is one such intervention. Meta-analyses suggest that SCT has moderate to large effects on numerous domains of social cognition and functioning (Kurtz and Richardson, 2012; Nijman et al., 2020). SCT may also have other benefits for individuals at clinical high risk for psychosis and those with schizophrenia, including alleviating positive symptoms (Hooker et al., 2014; Kurtz and Richardson, 2012) and reducing amotivation (Nahum et al., 2021). Although SCT is a promising intervention, it is also time and effort intensive, and yields heterogeneous social cognitive and functional outcomes – though this heterogeneity has not yet been rigorously characterized. Parsing this heterogeneity and identifying predictors of favorable response to SCT interventions is therefore a key research priority (Horan and Green, 2019).

Previous studies examining associations between patient characteristics and gains following SCT have produced mixed results. One meta-analysis of these studies (Kurtz and Richardson, 2012) found that longer illness duration predicted better social cognitive outcomes, while younger age and higher education predicted better generalization to functional outcomes. In a more recent meta-analysis (Nijman et al., 2020), none of the examined patient characteristics predicted social cognitive outcomes, while only male gender predicted generalization of benefits to social functioning. In addition to yielding mixed results, previous studies have not adequately established whether variables likely to influence SCT outcomes – such as individual differences in baseline performance of specific social cognitive skills or motivation – might predict response to SCT. There is therefore a need to better characterize the characteristics associated with individual benefits from SCT.

The Present Study

The present study sought to identify patterns and predictors of change in social cognition in response to SocialVille – a neuroplasticity-informed, adaptive, computerized and remotely-delivered SCT intervention. A recent study found that 12 weeks of SocialVille training elicited improvements across several outcomes, including a composite measure of social cognition, facial affect perception, social perception, motivation, social functioning and functional capacity, in a randomized controlled trial in patients with schizophrenia (Nahum et al., 2021). To this end, we applied latent class growth analysis (LCGA) to identify groups of participants who followed similar social cognition trajectories in the Treatment of Social cognition in Schizophrenia Trial (TRuSST) of SocialVille (Nahum et al., 2021; Rose et al., 2015). We then trained a machine-learning model to predict trajectory-group membership from baseline social cognitive, motivation, functional, symptom, and demographic information.

2. Method

2.1 Participants

Participants were 76 clinically stable adults with schizophrenia who were randomly assigned to complete a course of SocialVille SCT as part of TRuSST (Nahum et al., 2021). Participants were recruited from multiple sites in the United States and were required to have an estimated IQ ≥ 70 on the Wechsler Test of Adult Reading, and be clinically stable as indicated by a mean severity rating of no more than moderate (mean score ≤ 4) on the positive symptoms subscale and total symptoms score of the Positive and Negative Syndrome Scale (PANSS), and not have active suicide ideation. Recruitment protocols and measures are fully described elsewhere (Nahum et al., 2021; Rose et al., 2015).

2.2 Intervention

SocialVille is a neuroplasticity-informed computerized SCT program developed by Posit Science and is fully described elsewhere (Nahum et al., 2021; Rose et al., 2015). Briefly, it includes 27 exercises collectively targeting vocal and visual affect perception, social cue perception, theory of mind, self-referential style, and empathy. Exercises are designed to improve the speed and accuracy of brain systems involved in social information processing (Nahum et al., 2013). To promote neuroplasticity, exercises are intensive, adaptive, and reinforcing (Vinogradov et al., 2012). Participants completed training on loaned laptop computers, with the goal of 3-5 training sessions per week for a total of 40 training sessions over 8-12 weeks. On average, participants completed 27 training sessions (Nahum et al., 2021).

2.3 Measures

Social Cognition Composite.—We used a composite measure of social cognition, which was the primary outcome measure of TRuSST, to investigate trajectories of social cognition in individuals undergoing SCT. The composite had six equally-weighted indicators: facial emotion recognition (The Penn Emotional Recognition Test, ER40) (Gur et al., 2001), emotional prosody identification (The Prosody Identification Test, PROID) (Russ, 2008), immediate and delayed memory for faces (The Penn Faces Memory Test, PFMT) (Gur et al., 2001), the Mayer-Salovey-Caruso Emotional Intelligence Test's (MSCEIT) (Mayer et al., 2003) managing emotions subscale, and empathic accuracy (The Empathic Accuracy [EA] Task) (Lee et al., 2011).

The remaining primary and secondary outcome measures from the TRuSST trial, as captured during the trials' baseline (pre-treatment) session, were used as features to predict latent trajectory group membership. These measures captured:

Demographic information.—Demographic information included participant age, race, and sex.

Additional social cognition measures.—We included three additional measures of social cognition: the Awareness of Social Interaction Test (TASIT) (McDonald et al., 2006), which indexes social perception and theory of mind, the Faux Pas Test (Stone et al., 1998), which measures theory of mind, and the Ambiguous Intentions Hostility Questionnaire (AIHQ) (Combs et al., 2007b), which examines attributional style.

Symptoms.—The Positive and Negative Syndrome Scale (PANSS) (Kay et al., 1987) was used to measure psychosis symptom severity.

Functioning.—Functional capacity was measured with the UCSD Performance-based skills assessment (UPSA-2) (Patterson et al., 2001). Additional measures of functioning included the Global Functioning Scale: Social and Role (GFS) (Auther et al., 2006; Niendam et al., 2006), the Quality of Life Scale (QLS) (Heinrichs et al., 1984), the Specific Levels of Functioning Scale (SLOF) (Schneider and Struening, 1983), and the Social Functioning Scale (SFS) (Birchwood et al., 1990).

Motivation.—The behavioral inhibition and activation scale (BIS/BAS) (Carver and White, 1994) was used to measure sensitivity to anticipated rewards and punishments. The Temporal Experience of Pleasure Scale (TEPS) (Gard et al., 2006) was used to measure anticipatory and consummatory pleasure. Clinician ratings of motivation were derived from the abbreviated Quality of Life Scale (Heinrichs et al., 1984).

Intelligence.—Intelligence was estimated using the Wechsler Test of Adult Reading (WTAR) (Holdnack, 2001).

2.4 Analyses

2.4.1 Data Preprocessing: Outliers and Missing Data—For each participant, assessment waves with missing data were deleted list-wise. Outliers were detected using the method of Hubert and Van Der Veeken (Hubert and Van der Veeken, 2008) and then winsorized. Variables were standardized (z-scored) prior to analyses.

2.4.2 Latent Class Growth Analysis—The *Flexmix* package, version 2.3.18, was used to identify individuals whose social cognition followed similar trajectories during the TRuSST trial. In characterizing these trajectories, three assessment timepoints were considered: baseline, mid-point (after M=20 training sessions), and final (after M=40 training sessions). Time was defined as the number of completed training sessions, log transformed and incremented by one, allowing us to better capture participants' trajectories of social cognition with a linear growth model. The number of sessions, rather than other metrics of time was used because this metric was expected to be most closely associated with changes in social cognition. Fifty random initializations of model parameters were used to better avoid convergence to local maxima.

Final model selection was informed by multiple fit metrics. Models with lower AIC and BIC, and higher log likelihood, were considered a superior fit to the data. Each model was also compared to ones with a larger number of groups using bootstrapped likelihood ratio tests, in keeping with simulation studies suggesting these tests' utility for choosing group numbers (Nylund et al., 2007). The best-fitting model was evaluated using the average maximum posterior probability of group assignments (preferably >.70), the relative entropy of these posteriors (preferably >.80), and the odds of correct classification of participants (preferably > 5). Sufficient group size (preferably > 5% of the sample) was also evaluated.

2.4.3 Predicting Class Membership from Baseline Data—Only one group in our best-fitting latent trajectory model experienced a statistically significant improvement in social cognition with training. We trained logistic regression and random forest learners to predict membership in this group [coded: 1] vs. all others [coded: 0], with the dual aims of better understanding this group's characteristics and enhancing our ability to predict which individuals with schizophrenia would benefit from social cognitive training.

Model Performance Assessment.: A 10-fold cross-validation procedure, with a nested 6-fold cross-validation loop for tuning model hyperparameters, was used to estimate the mean and standard deviation of performance metrics.

Because classes were imbalanced (22 cases were in the positive class vs. 54 in the negative class), the main outcome of interest for assessing model performance was the area under the curve describing the tradeoff between precision and recall (AUR-PR) across thresholds for converting model-derived class membership probabilities into hard class labels. A point estimate of the AUC-PR was generated using the PRROC package in R (version: 1.3.1). A

95% confidence interval around this point was generated via the logit estimation method (Boyd et al., 2013).

The Generalized Threshold Shifting [GHOST] protocol (Esposito et al., 2021) was used to select an optimal threshold at which to convert model-derived class membership probabilities into hard class labels, as this protocol outperforms other selection methods (e.g., random under-sampling) for imbalanced classification problems. Briefly, the model selected in each cross-validation iteration was used to generate membership probabilities for the training data. Random subsamples of 20% of the training data [and associated membership probabilities] (n=100) were taken, and the median Cohen's kappa for each potential probability threshold across these subsamples was calculated. The threshold with the highest median kappa was selected and applied to the test data in each cross-validation iteration. The most commonly chosen threshold across cross-validation iterations was recommended for use with our final model.

Algorithm Selection.: We considered random forest and logistic regression models, both of which are learners that: (1) are appropriate for classification problems and (2) are interpretable, allowing for knowledge extraction. The model with the highest AUC-PR and F1 score was considered the best-performing.

Feature Selection.: Due to the number of variables included in our analysis, we performed a feature selection step to determine the optimal number of variables and reduce the feature space for each model. Recursive feature elimination was used for feature selection in the logistic regression model, whereas the Boruta algorithm (Kursa and Rudnicki, 2010) was used to select features in the random forest model. In both models, class imbalance was addressed via inverse frequency weighting. These methods are designed to optimize signal-to-noise ratios by eliminating variables unlikely to be related to the outcome of interest, thus minimizing concerns related to a large set of variables.

Feature Importance.: Shapley values (estimated using the *fastshap* package [version 0.0.7]), the absolute marginal contribution of features to the deviation of predictions from the sample mean, averaged across all possible combinations of features, were used to quantify feature importance (Boehmke and Greenwell, 2019). This approach, which conceptualizes features as players in a game where they collaborate to earn a payoff (the difference between the predicted and average outcome), has several advantages, including that contributions to the final prediction are fairly distributed across features.

3. Results

3.1 Differential Drop-Out and Missing Data

Study attrition and demographic information is fully described elsewhere (Nahum et al., 2021). Briefly, our final sample includes 52 male and 23 female participants, with an average age of 42.86 (SD=12.78). 28% of participants (N = 21 of 76) assigned to SCT with SocialVille dropped out of the study. Individuals who dropped out were included in the latent class growth analysis. Individuals who dropped out performed moderately worse at baseline on the SC composite measure (χ^2=5.02, p=.025, Cohen's d approximation=.47).

Thus, data were not missing completely at random (MCAR). However, drop-out was unrelated to any other study variable, including age, gender, and functional capacity. Further, among participants who returned for post-training assessments, only 3% of data on SC composite scores was missing and drop-out did not differ significantly across trajectory groups. Thus, data were assumed missing at random (MAR), consistent with the assumption of our Latent Class Growth Analysis.

3.2 Latent Trajectory Groups

3.2.1 Model Selection and Characteristics—A latent class trajectory growth model with five groups was the best fit to the data (Table 1). This model was superior to alternatives across all examined model fit indices (AIC, BIC, log-likelihood). Selection of the five-group model was also supported by bootstrapped likelihood ratio tests, which suggested that the true number of groups exceeded four, but not five. The five-group model featured high average posterior probabilities of group membership (≥.89) and high odds of correct classification (≥15:1), suggesting that individuals could be assigned to trajectory groups with high confidence and that misclassification was rare. This model also had an acceptable relative entropy (0.86), implying sufficient separation of the latent class centroids.

The five latent trajectory groups can be described as follows (Figure 1; Table 2): Group 1 (29% of participants) began with slightly above average (of the sample mean) social cognition, as indexed by our composite measure, and this ability improved with SCT. Group 2 (9% of participants) was initially just over one standard deviation above the sample mean social cognitive ability but did not improve with training. Groups 3 and 4 (18% and 36% of participants, respectively) began with slightly below-average social cognition and showed non-significant trends toward improving with training. Finally, Group 5 (8% of participants) had an initial social cognitive ability approximately one standard deviation below the sample mean, and experienced statistically significant deterioration in this ability, despite SCT.

3.2.2 Univariate Comparison of Clinical Variables Across Groups—Statistical comparisons with *ANOVAs* and Fisher's Exact Tests, uncorrected for multiple comparisons, suggested that the groups differed on only a few of the measured variables (Table 3; Supplementary Figure S1). Group 1 had better performance on social cognitive measures – including the SC composite, the TASIT, the Faux Paus test, and the AHIQ – and a lower level of positive and general symptoms than all other groups except Group 2.

3.2.3 Predictors of Group 1 Membership—Only one group (Group 1) in our best-fitting latent trajectory model experienced a statistically significant improvement in social cognition with training (Table 2). We trained several learners to predict membership in this group [coded: 1] vs. all others [coded: 0], with the dual aims of enhancing our ability to predict which individuals with schizophrenia would benefit from social cognitive training and better understanding this group's characteristics.

Model Performance.: Performance metrics for each model are listed in Table 4. At the threshold suggested by the GHOST protocol (.85), the average area under the random forest model's precision-recall curve was 0.73 (*SD*=0.24; 95% CI=[0.51 0.87]) across cross-validation folds. This model had an average F1 score of 0.67 (*SD*=0.13). While the average AUC-PR for the logistic regression model was similar (0.72; *SD*=0.31), this model also had a lower average F1 score (0.61, *SD*=0.24; 95% CI=[0.53 0.85]). Thus, the random forest model was considered superior. This model had the following hyperparameters, chosen via our cross-validation procedure: the forest had 500 trees with at least one case per leaf; splits were based on one randomly-selected variable and the "extratrees" split rule.

Feature Importance.: Average Shapley values (Figure 2) suggested that the most important features for predicting Group 1 membership in the random forest model were, in descending order: ability to infer the speaker's beliefs on the TASIT (average [absolute] value=0.10), habitual tendency to manage emotions (MSCEIT D scale; average value=0.09), emotion recognition on the ER40 (average value=0.08), prosody detection ability (PROID; average value=0.06), functional capacity (UPSA; average value=0.03), and ability to infer what the speaker intends on the TASIT (average value=0.03). Inspection of dependence plots (Supplementary Figure S2) suggested that higher scores on these features were, in general, associated with an increased likelihood of being in Group 1, and benefiting from social cognitive training.

In the logistic regression model, the sole retained predictor was ability to infer the speaker's beliefs on the TASIT.

4. Discussion

The present study took a proof-of-concept, data-driven approach to parsing the heterogeneity in response to a computerized SCT, SocialVille, which was previously not well characterized. Results suggested that trajectories of social cognition during SCT can be described by placing individuals into five groups. Only one of these (consisting of approximately 30% of participants) showed statistically significant improvement in social cognition, two more (together containing another 54% of participants) had non-significant trends in this direction, and the remaining groups either did not improve or became worse on measures of social cognition despite training. These results reinforce the need for models that can identify people who will likely benefit from SCT, and for further research identifying how SCT can best be modified to promote robust response for other patients. Notably then, when the model we built in this study predicted that an individual was a member of the trajectory group that benefited from SocialVille SCT, it was correct 61% of the time. It was also able to identify 75% of the members of this group. Although these results should be considered preliminary given our small sample size, they provide valuable foundations for future work seeking to personalize and tailor SCT to individual patients.

The distinct trajectory groups that we observed could have implications for tailoring SCT based on individual patient needs. Approximately 30% of participants fell into the group that responded to SCT (Group 1). In our random forest model, better social perception and theory of mind (i.e., inferring beliefs and intentions), emotion management, emotion recognition, prosody identification, and functional capacity were the most important predictors of membership in this group. Better theory of mind abilities, as measured by the

TASIT, may be an especially salient prognostic indicator – these were the most important predictor in the random forest model and the sole predictor identified in our logistic regression model. Patients with these characteristics may be especially good candidates for SocialVille SCT, which intensively and adaptively targets multiple social cognitive abilities (i.e., lower-level affect perception to higher-level theory of mind), and is delivered fully remotely without the requirement of highly trained therapists or frequent clinic visits (Nahum et al., 2021).

Two other groups experienced a non-significant trend towards response. Together these groups comprised another 54% of the sample and had slightly below to below average social cognition at baseline. The non-significant trend could be attributed to the small sample size or could indicate that due to their lower baseline social cognitive abilities, this group requires more intensive or longer duration SCT to elicit substantial benefits than might be required for Group 1 patients. However, as treatment duration and intensity have not been identified as treatment moderators in meta-analytic investigations of heterogeneous social cognition interventions (Nijman et al., 2020; Yeo et al., 2022), further intervention research is needed to determine whether tailoring SocialVille SCT in these ways results in more favorable response for a subset of patients. Once this evidence base is established, it may be possible to develop predictive tools to guide tailoring the intensity and duration of training needed to elicit a robust response for specific patients.

The two remaining groups did not show evidence of benefitting from SCT. One group had above average social cognition at baseline and did not show further improvements. The other had marked impairments in social cognition at baseline which further deteriorated throughout the course of the intervention. While preliminary, these results could indicate that individuals who have significantly above average social cognition abilities at baseline may not garner additional benefits from training and may instead benefit from investing time in other treatment services. In contrast, those with significantly impaired social cognition at baseline may require higher intensity, longer duration, or a different type of social cognitive intervention, such as SCT delivered in person with a 1:1 coach or a social skills focused intervention (Horan et al., 2018; Roberts and Penn, 2009). This group also had the highest severity of symptoms and may first require better symptom stabilization, as has been found for neurocognitive remediation (Biagianti et al., 2021; Wykes et al., 2011). In line with these hypotheses, some social cognitive interventions, such as Social Cognitive and Interaction Training, have evidence for efficacy in more impaired schizophrenia samples (Combs et al., 2007a). Additionally, previous work in neurocognitive remediation has suggested that lower baseline cognition predicts greater response specifically to compensatory and problem-solving focused cognitive interventions (Rodewald et al., 2014; Twamley et al., 2011). Whether elements underpinning these programs, such as strategy coaching, are necessary in SCT interventions for those who have greater baseline social cognition impairments requires further investigation.

Few previous studies have examined predictors of response to SCT interventions and results have been mixed. In a proof-of-concept study, baseline neurocognition did not predict response to Understanding Social Situations, an intervention targeting higher order social cognition, delivered individually by a trained therapist with drill and practice exercises

(Fiszdon et al., 2017). A study of a group meta-cognitive training intervention with a focus on social cognition found that poorer social cognition at baseline was a predictor of better treatment gains (Alvarez-Astorga et al., 2019). Due to the significant heterogeneity in intervention components and targets of these interventions, it is difficult to compare results in the context of the current study. The neurocognitive training literature on predictors of response has also been mixed with respect to baseline cognitive performance, with the influence of baseline cognitive ability on training response differing based on specific intervention targets, components, and theoretical frameworks (e.g., neuropsychology vs. neuroplasticity based cognitive remediation) (Biagianti et al., 2021; Reser et al., 2019). Finally, in line with our findings, one previous study in healthy adults examined trajectories of response to a working memory training, finding that better baseline cognition was associated with positive response trajectories (Guye et al., 2017).

Identification of distinct treatment response profiles allows for the development of prediction models that could guide treatment selection and tailoring for individual patients, which is an important goal for time intensive interventions like SCT (Horan and Green, 2019). Due to the heterogeneity of response observed in cognitive training, prediction models utilizing machine learning may especially have a role in improving prediction accuracy (Cearns et al., 2019; Schnack, 2019). Our preliminary, proof of concept models to predict membership in the latent trajectory group that improved with training demonstrated adequate precision and recall, with overall similar performance for the random forest and less complex logistic regression model. The logistic regression model, which relied only on inferring speakers' beliefs as measured by the TASIT, could be explored further as a practical tool that can more easily be deployed in clinical settings (Dwyer et al., 2018) to predict SCT response as it relied on only one measure of social cognition. To our knowledge, this is the first report of machine learning prediction of response to a social cognitive intervention. Future research is needed to understand whether these prediction models might generalize to other types of social cognitive interventions and could be used to help guide treatment planning.

Several limitations should be kept in mind when interpreting our results, in addition to the aforementioned small sample size, which could lead to overfitting of the machine learning models. Cognitive training is a time and effort intensive intervention. As such, cognitive training trials, like ours, often contend with high drop-out rates. The resulting missing data can influence LCGA results. A second limitation of LCGA is that groups can vary with study design and modeling choices, such as length of participant follow-up or the definition of time employed. Thus, future research examining the replicability of our results is warranted. The analysis of differences in clinical variables between latent trajectory groups were not corrected for multiple comparisons although several measures were assessed, increasing risk for Type-I error. Neurocognitive ability may be a potential predictor of response to cognitive training interventions (Biagianti et al., 2021), and we did not have neurocognitive data available to use in our models. While not a limitation per se, the groups identified here are best conceptualized as approximations of continuous, but unknown, population heterogeneity in trajectories of social cognition with use of SocialVille (Van De Schoot et al., 2017), rather than distinct groups that are "natural kinds." Moreover, there is substantial within-group heterogeneity in each groups' trajectories (see Figure 1), making caution warranted when drawing inference about likely outcome of treatment with SocialVille from likely group membership.

5.0 Conclusions

We identified five profiles of response to the SocialVille SCT intervention. Additionally, we identified preliminary baseline characteristics predicting favorable social cognitive improvement following the intervention, including less impaired social cognition and better functional capacity and symptom control. The different response trajectories identified highlight the need for personalized social cognitive interventions to promote robust treatment response for more patients to this highly scalable intervention, and accurate prediction tools that can be used to better guide treatment planning.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Funding statement:

Research reported in this publication was supported by the National Institute of Mental Health (NIMH) Award R44MH091793 and by the National Center for Advancing Translational Sciences of the National Institutes of Health (Award Number UL1-TR002494). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

Alvarez-Astorga A, Sotelo E, Lubeiro A, de Luis R, Gomez-Pilar J, Becoechea B, Molina V, 2019. Social cognition in psychosis: Predictors and effects of META-cognitive training. Prog Neuropsychopharmacol Biol Psychiatry 94, 109672. [PubMed: 31228639]

Auther AM, Smith CW, Cornblatt BA, 2006. Global Functioning: Role Scale (GF: Role). University of California, Los Angeles, Los Angeles, CA.

Biagianti B, Castellaro GA, Brambilla P, 2021. Predictors of response to cognitive remediation in patients with major psychotic disorders: A narrative review. J Affect Disord 281, 264–270. [PubMed: 33341008]

Birchwood M, Smith J, Cochrane R, Wetton S, Copestake S, 1990. The Social Functioning Scale. The development and validation of a new scale of social adjustment for use in family intervention programmes with schizophrenic patients. Br J Psychiatry 157, 853–859. [PubMed: 2289094]

Boehmke B, Greenwell BM, 2019. Hands-on machine learning with R. CRC press.

Boyd K, Eng KH, Page CD, 2013. Area under the Precision-Recall Curve: Point Estimates and Confidence Intervals. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 451–466.

Carver CS, White TL, 1994. Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: the BIS/BAS scales. Journal of personality and social psychology 67(2), 319.

Cearns M, Hahn T, Baune BT, 2019. Recommendations and future directions for supervised machine learning in psychiatry. Transl Psychiatry 9(1), 271. [PubMed: 31641106]

Combs DR, Adams SD, Penn DL, Roberts D, Tiegreen J, Stem P, 2007a. Social Cognition and Interaction Training (SCIT) for inpatients with schizophrenia spectrum disorders: preliminary findings. Schizophr Res 91(1-3), 112–116. [PubMed: 17293083]

Combs DR, Penn DL, Wicher M, Waldheter E, 2007b. The Ambiguous Intentions Hostility Questionnaire (AIHQ): a new measure for evaluating hostile social-cognitive biases in paranoia. Cogn Neuropsychiatry 12(2), 128–143. [PubMed: 17453895]

Dwyer DB, Falkai P, Koutsouleris N, 2018. Machine Learning Approaches for Clinical Psychology and Psychiatry. Annu Rev Clin Psychol 14, 91–118. [PubMed: 29401044]

Esposito C, Landrum GA, Schneider N, Stiefl N, Riniker S, 2021. GHOST: Adjusting the Decision Threshold to Handle Imbalanced Data in Machine Learning. J Chem Inf Model 61(6), 2623–2640. [PubMed: 34100609]

Fett AK, Viechtbauer W, Dominguez MD, Penn DL, van Os J, Krabbendam L, 2011. The relationship between neurocognition and social cognition with functional outcomes in schizophrenia: a meta-analysis. Neurosci Biobehav Rev 35(3), 573–588. [PubMed: 20620163]

Fiszdon JM, Roberts DL, Penn DL, Choi KH, Tek C, Choi J, Bell MD, 2017. Understanding Social Situations (USS): A proof-of-concept social-cognitive intervention targeting theory of mind and attributional bias in individuals with psychosis. Psychiatr Rehabil J 40(1), 12–20. [PubMed: 27281040]

Gard DE, Gard MG, Kring AM, John OP, 2006. Anticipatory and consummatory components of the experience of pleasure: a scale development study. Journal of research in personality 40(6), 1086–1102.

Gur RC, Ragland JD, Moberg PJ, Turner TH, Bilker WB, Kohler C, Siegel SJ, Gur RE, 2001. Computerized neurocognitive scanning: I. Methodology and validation in healthy people. Neuropsychopharmacology 25(5), 766–776. [PubMed: 11682260]

Guye S, De Smoni C, von Bastian CC, 2017. Do individual differences predict change in cognitive training performance? A latent growth curve modeling approach. Journal of Cognitive Enhancement 1, 374–393.

Heinrichs DW, Hanlon TE, Carpenter WT Jr., 1984. The Quality of Life Scale: an instrument for rating the schizophrenic deficit syndrome. Schizophr Bull 10(3), 388–398. [PubMed: 6474101]

Holdnack J, 2001. WTAR: Wechsler Test of Adult Reading manual. Psychol. Corp, San Antonio, TX.

Hooker CI, Carol EE, Eisenstein TJ, Yin H, Lincoln SH, Tully LM, Dodell-Feder D, Nahum M, Keshavan MS, Seidman LJ, 2014. A pilot study of cognitive training in clinical high risk for psychosis: initial evidence of cognitive benefit. Schizophr Res 157(1-3), 314–316. [PubMed: 24954429]

Horan WP, Dolinsky M, Lee J, Kern RS, Hellemann G, Sugar CA, Glynn SM, Green MF, 2018. Social Cognitive Skills Training for Psychosis With Community-Based Training Exercises: A Randomized Controlled Trial. Schizophr Bull 44(6), 1254–1266. [PubMed: 29300973]

Horan WP, Green MF, 2019. Treatment of social cognition in schizophrenia: Current status and future directions. Schizophr Res 203, 3–11. [PubMed: 28712968]

Hubert M, Van der Veeken S, 2008. Outlier detection for skewed data. Journal of Chemometrics: A Journal of the Chemometrics Society 22(3-4), 235–246.

Kay SR, Fiszbein A, Opler LA, 1987. The positive and negative syndrome scale (PANSS) for schizophrenia. Schizophr Bull 13(2), 261–276. [PubMed: 3616518]

Kursa MB, Rudnicki WR, 2010. Feature selection with the Boruta package. Journal of statistical software 36, 1–13.

Kurtz MM, Richardson CL, 2012. Social cognitive training for schizophrenia: a meta-analytic investigation of controlled research. Schizophr Bull 38(5), 1092–1104. [PubMed: 21525166]

Lee J, Zaki J, Harvey PO, Ochsner K, Green MF, 2011. Schizophrenia patients are impaired in empathic accuracy. Psychol Med 41(11), 2297–2304. [PubMed: 21524334]

Mayer JD, Salovey P, Caruso DR, Sitarenios G, 2003. Measuring emotional intelligence with the MSCEIT V2.0. Emotion 3(1), 97–105. [PubMed: 12899321]

McDonald S, Bornhofen C, Shum D, Long E, Saunders C, Neulinger K, 2006. Reliability and validity of The Awareness of Social Inference Test (TASIT): a clinical test of social perception. Disabil Rehabil 28(24), 1529–1542. [PubMed: 17178616]

Nahum M, Lee H, Fisher M, Green MF, Hooker CI, Ventura J, Jordan JT, Rose A, Kim SJ, Haut KM, Merzenich MM, Vinogradov S, 2021. Online Social Cognition Training in Schizophrenia: A Double-Blind, Randomized, Controlled Multi-Site Clinical Trial. Schizophr Bull 47(1), 108–117. [PubMed: 32614046]

Nahum M, Lee H, Merzenich MM, 2013. Principles of neuroplasticity-based rehabilitation. Prog Brain Res 207, 141–171. [PubMed: 24309254]

Niendam TA, Bearden CE, J. JK, Cannon TD, 2006. Global Functioning: Role Scale (GF: Role). University of California, Los Angeles, Los Angeles, CA.

Nijman SA, Veling W, van der Stouwe ECD, Pijnenborg GHM, 2020. Social Cognition Training for People With a Psychotic Disorder: A Network Meta-analysis. Schizophr Bull 46(5), 1086–1103. [PubMed: 32162658]

Nylund KL, Asparouhov T, Muthén BO, 2007. Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. Structural equation modeling: A multidisciplinary Journal 14(4), 535–569.

Patterson TL, Goldman S, McKibbin CL, Hughs T, Jeste DV, 2001. UCSD Performance-Based Skills Assessment: development of a new measure of everyday functioning for severely mentally ill adults. Schizophr Bull 27(2), 235–245. [PubMed: 11354591]

Reser MP, Slikboer R, Rossell SL, 2019. A systematic review of factors that influence the efficacy of cognitive remediation therapy in schizophrenia. Aust N Z J Psychiatry 53(7), 624–641. [PubMed: 31177813]

Roberts DL, Penn DL, 2009. Social cognition and interaction training (SCIT) for outpatients with schizophrenia: a preliminary study. Psychiatry Res 166(2-3), 141–147. [PubMed: 19272654]

Rodewald K, Holt DV, Rentrop M, Roesch-Ely D, Liebrenz M, Funke J, Weisbrod M, Kaiser S, 2014. Predictors for improvement of problem-solving during cognitive remediation for patients with Schizophrenia. J Int Neuropsychol Soc 20(4), 455–460. [PubMed: 24589198]

Rose A, Vinogradov S, Fisher M, Green MF, Ventura J, Hooker C, Merzenich M, Nahum M, 2015. Randomized controlled trial of computer-based treatment of social cognition in schizophrenia: the TRuSST trial protocol. BMC Psychiatry 15, 142. [PubMed: 26138715]

Russ J., 2008. Development of PROID, a computerized emotional prosody identification task. PennScience Journal 6, 14–20.

Savla GN, Vella L, Armstrong CC, Penn DL, Twamley EW, 2013. Deficits in domains of social cognition in schizophrenia: a meta-analysis of the empirical evidence. Schizophr Bull 39(5), 979–992. [PubMed: 22949733]

Schnack HG, 2019. Improving individual predictions: Machine learning approaches for detecting and attacking heterogeneity in schizophrenia (and other psychiatric diseases). Schizophr Res 214, 34–42. [PubMed: 29074332]

Schneider LC, Struening EL, 1983. SLOF: a behavioral rating scale for assessing the mentally ill. Soc Work Res Abstr 19(3), 9–21. [PubMed: 10264257]

Stone VE, Baron-Cohen S, Knight RT, 1998. Frontal lobe contributions to theory of mind. J Cogn Neurosci 10(5), 640–656. [PubMed: 9802997]

Twamley EW, Burton CZ, Vella L, 2011. Compensatory cognitive training for psychosis: who benefits? Who stays in treatment? Schizophr Bull 37 Suppl 2(Suppl 2), S55–62. [PubMed: 21860048]

Van De Schoot R, Sijbrandij M, Winter SD, Depaoli S, Vermunt JK, 2017. The GRoLTS-checklist: guidelines for reporting on latent trajectory studies. Structural Equation Modeling: A Multidisciplinary Journal 24(3), 451–467.

Vinogradov S, Fisher M, de Villers-Sidani E, 2012. Cognitive training for impaired neural systems in neuropsychiatric illness. Neuropsychopharmacology 37(1), 43–76. [PubMed: 22048465]

Wykes T, Huddy V, Cellard C, McGurk SR, Czobor P, 2011. A meta-analysis of cognitive remediation for schizophrenia: methodology and effect sizes. Am J Psychiatry 168(5), 472–485. [PubMed: 21406461]

Yeo H, Yoon S, Lee J, Kurtz MM, Choi K, 2022. A meta-analysis of the effects of social-cognitive training in schizophrenia: The role of treatment characteristics and study quality. Br J Clin Psychol 61(1), 37–57. [PubMed: 34291465]

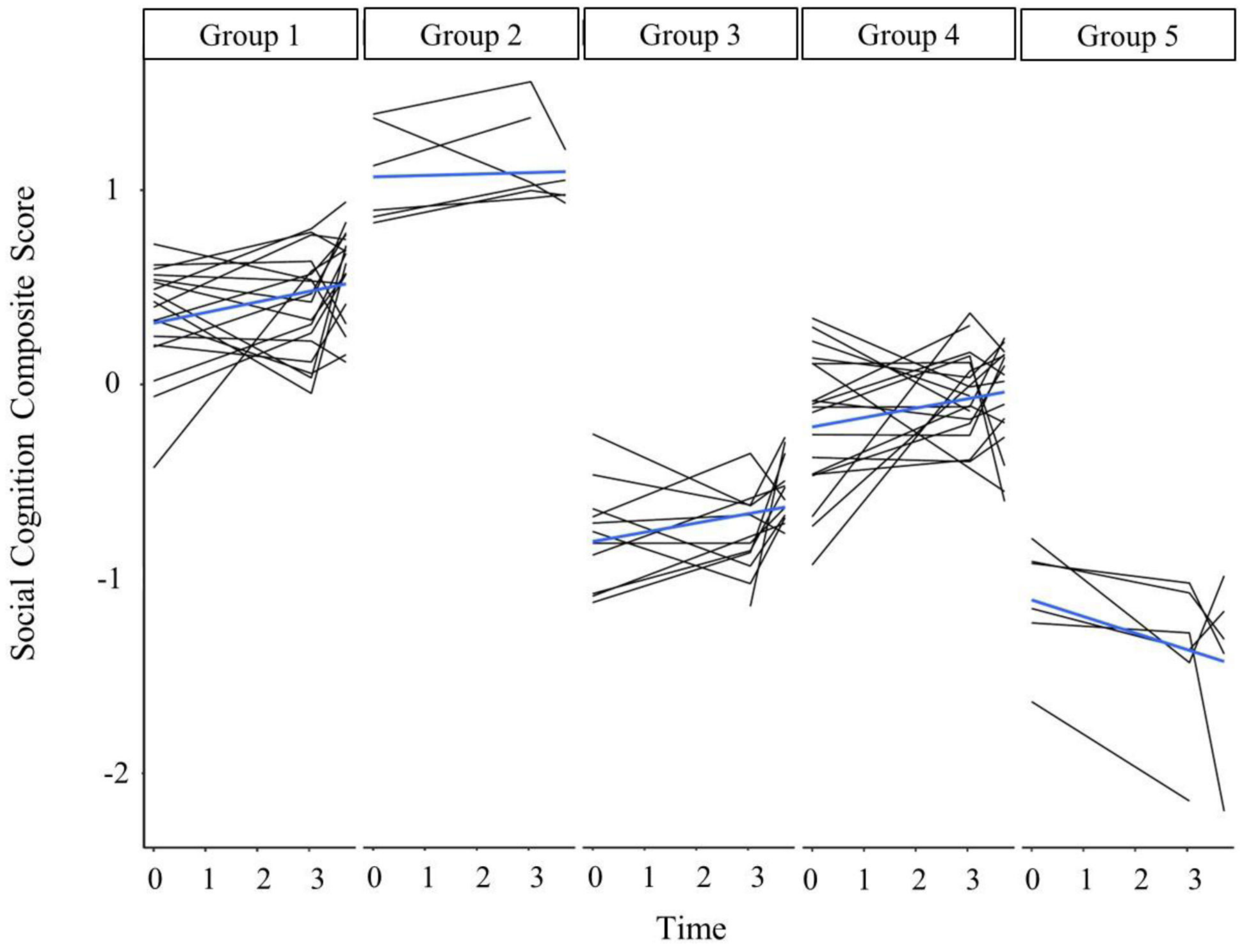

Figure 1.

Trajectories of each group. Criterion=composite measure of social cognition scores (standardized). Time=log(number of training sessions+1). Black lines are individual trajectories, blue lines are their group average.

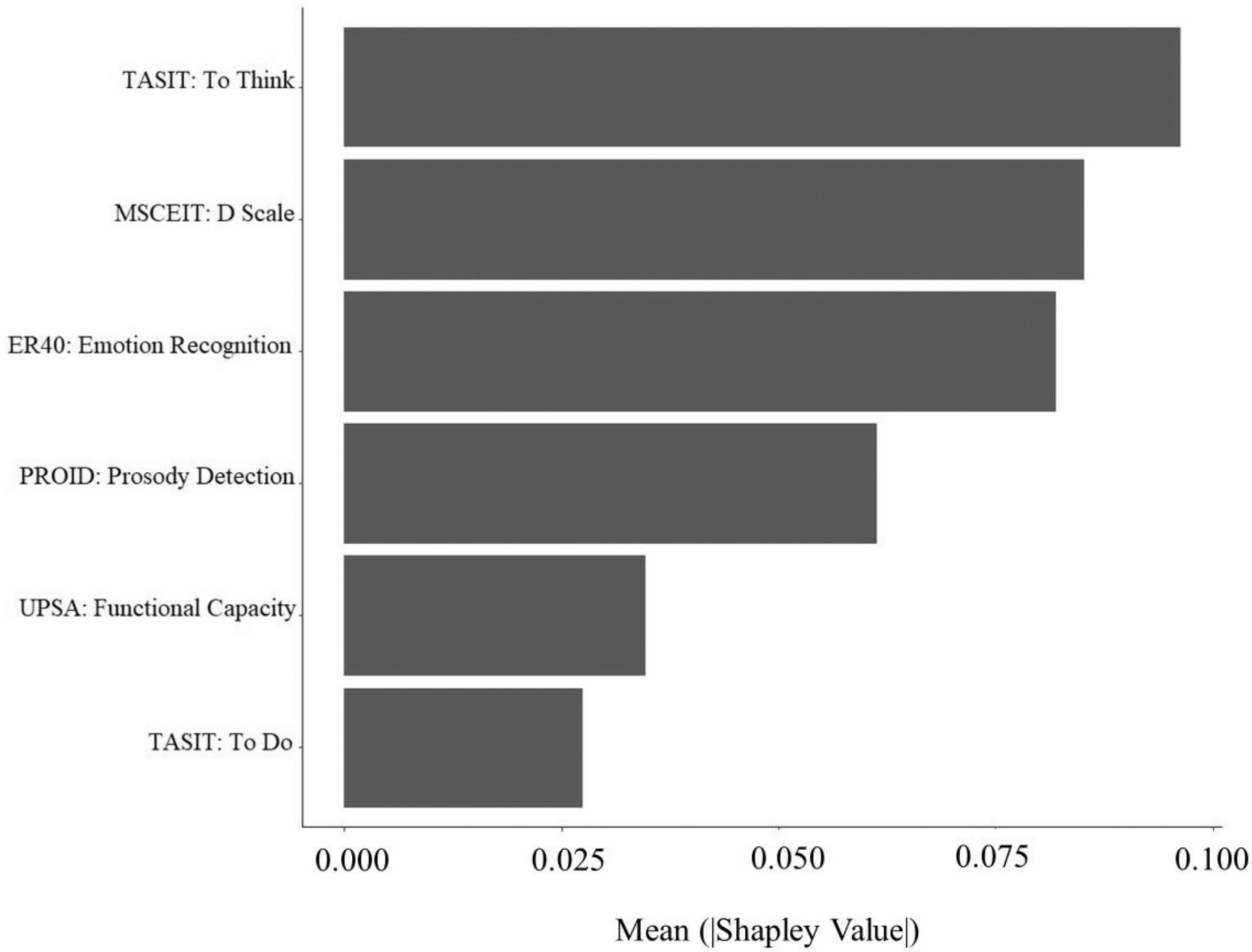

Figure 2. Shapley values of each predictor in the final random forest model.

Note. ER40 = Penn Emotion Recognition Test; MSCEIT D Scale: Mayer-Salovey-Caruso Emotional Intelligence Test's managing emotions subscale; TASIT = The Awareness of Social Interaction Test; UPSA: UCLA Performance Based Skills Assessment.

Table 1.

Group-based Trajectory Model Selection

#(Groups)	AIC	BIC	Log Likelihood	Average Posterior Probability	E_k	Odds of Correct Classification	Percent Assigned to Group	BLRT: #(groups) > selected
1	405.04	414.75	−199.52	..	..	..	100	p < .001
2	328.13	347.55	−158.07	.94 .91	.78	9 16	62 39	p=.001
3	290.95	320.08	−136.48	.93 .93 .96	.85	39 12 71	25 51 24	p=.003
4	278.37	317.21	−127.19	.96 .96 .90 .88	.84	89 252 19 13	22 9 32 37	p=.010
5	**262.58**	**311.13**	**−116.29**	**.90 .98 .90 .89 .93**	**.86**	**23 414 39 15 156**	**29 9 18 36 8**	**p=.132**
6	264.79	323.04	−114.39	.94 .79 .83 .92 .96 .98	.81	55 66 9 50 278 462	24 5 36 18 8 9	p=.433
7	265.03	333.00	−111.52	.96 .97 .81 .92 .79 .90 .93	.84	290 372 14 200 17 39 72	8 9 24 18 20 16	p=.319

Note. Bold=best model. Posterior probabilities represent the chance, on average, that an individual assigned to a given trajectory group is, in fact, a member of that group (values >.8 are desirable). Odds of correct classification are the ratio of the average posterior probability to the probability of assignment to each group based on its overall frequency. It thus quantifies how informative the model's group assignments are (vs. random chance). Values > 5 are desirable. E_k=relative entropy. Relative entropy captures the degree to which a given individual has unequal posterior probabilities for each group – such that, for example, the model strongly suggests assignment to one group over others. Values >.8 are desirable. BLRT=Bootstrapped Likelihood Ratio Test, with 1000 resamples. This bootstrapping procedure creates a distribution for the difference in fit between a model with k and k-1 groups, allowing inference about the probability that the number of groups should be greater than that selected.

Table 2.

Group Trajectories for Selected Model

Group	Parameter	Estimate	SE	z	*p*
1	**Intercept**	**0.30**	**0.07**	**3.99**	**<.001**
	Slope	**0.06**	**0.03**	**2.25**	**.024**
2	**Intercept**	**1.05**	**0.12**	**9.05**	**<.001**
	Slope	0.01	0.04	<1	.830
3	**Intercept**	**−0.80**	**0.11**	**7.45**	**<.001**
	Slope	0.04	0.03	1.13	.260
4	**Intercept**	**−0.21**	**0.08**	**2.52**	**.012**
	Slope	0.04	0.02	1.78	.076
5	**Intercept**	**−1.10**	**0.12**	**9.09**	**<.001**
	Slope	**−0.09**	**0.04**	**2.10**	**.036**

Note. Bold=statistically significant. This table describes the estimated slope and intercept of composite social cognition over the course of training. The final column is the *p*-value for a z-test of the parameter against zero. Because the slope of Group 1 alone is significant and positive – indicating improvement of social cognition with training, that group is the focus of our predictive analyses. For an exploratory comparison of group slope/intercept parameters, see Supplementary Table S1.

Table 3.

Comparison of Differences Across Groups

Group	1	2	3	4	5	*Test Statistic*	*p*
Drop-Out	Y:18, N:4	Y:5, N:2	Y:10, N:4	Y:14, N:13	Y:5, N:1	--	.225
Gender	**M:15, F:7**	**M:4, F:3**	**M:6, F:8**	**M:22, F:5**	**M:6, F:0**	**--**	**.046**
Age	40.68 (14.46)	28.00 (11.52)	48.57 (14.24)	45.33 (11.62)	39.00 (13.40)	1.68	.199
Race						--	.207
White	*13*	*5*	*4*	*11*	*3*	--	--
African American	*8*	*2*	*9*	*9*	*3*	--	--
Other	*1*	*0*	*1*	*7*	*0*	--	--
SC Composite	**0.34 (0.26)**	**1.06 (0.24)**	**−0.79 (0.25)**	**−0.22 (0.31)**	**−1.10 (0.30)**	**47.70**	**<.001**
WTAR	100.14 (10.60)	106.00 (8.66)	91.21 (10.92)	97.59 (10.99)	92.50 (14.02)	2.94	.090
AIHQ: Item	14.18 (3.98)	12.43 (3.74)	12.29 (4.60)	12.89 (4.14)	16.17 (5.27)	<1	.868
AIHQ: Hostility	9.95 (3.85)	8.71 (2.75)	7.93 (2.79)	9.74 (3.94)	12.17 (4.07)	<1	.581
AIHQ: Aggression	**7.36 (1.73)**	**6.86 (1.68)**	**8.00 (2.21)**	**7.56 (2.38)**	**11.17 (1.60)**	**5.08**	**.027**
Faux Pas	**0.77 (0.11)**	**0.86 (0.11)**	**0.62 (0.17)**	**0.73 (0.17)**	**0.50 (0.19)**	**8.22**	**.005**
TASIT: To Think	**13.81 (1.37)**	**12.43 (1.90)**	**9.57 (2.14)**	**11.33 (2.83)**	**9.83 (1.62)**	**21.14**	**<.001**
TASIT: To Feel	**11.63 (2.44)**	**12.00 (2.83)**	**9.35 (1.27)**	**10.66 (1.82)**	**8.50 (1.97)**	**9.14**	**.003**
TASIT: To Say	**11.09 (2.90)**	**11.29 (3.59)**	**8.57 (1.79)**	**10.30 (1.75)**	**7.83 (2.32)**	**6.05**	**.016**
TASIT: To Do	**11.55 (2.24)**	**11.28 (1.80)**	**9.21 (1.81)**	**9.74 (2.05)**	**8.00 (1.41)**	**18.91**	**<.001**
BIS/BAS: BIS	21.22 (5.17)	19.57 (4.69)	21.29 (3.93)	21.44 (3.43)	21.50 (2.56)	<1	.068
BIS/BAS: BAS	12.83 (1.99)	14.09 (1.52)	14.00 (1.58)	13.49 (1.94)	14.22 (1.59)	2.34	.130
TEPS Total	40.30 (5.96)	39.71 (5.42)	41.07 (5.89)	42.38 (7.95)	42.00 (5.67)	1.38	.244
QLS: Motivation	3.09 (0.90)	3.71 (0.49)	2.71 (0.82)	3.07 (0.87)	2.67 (1.03)	<1	.326
SLOF: Interpersonal	24.64 (3.97)	28.00 (4.20)	25.14 (7.42)	27.03 (5.02)	25.83 (6.05)	<1	.261
SLOF: Acceptability	27.27 (2.98)	28.71 (1.98)	26.64 (2.17)	27.59 (2.59)	26.00 (3.52)	<1	.615
SLOF: Activities	48.22 (9.41)	54.14 (1.21)	43.35 (12.21)	46.19 (12.21)	50.33 (8.16)	<1	.501
GFS: Social	5.91 (1.60)	6.57 (1.27)	6.14 (1.66)	5.96 (1.72)	5.50 (1.64)	<1	.720
GFS: Role	4.18 (2.15)	6.00 (2.38)	3.93 (1.64)	4.22 (1.97)	4.50 (2.26)	<1	.812
PANSS Positive	**13.77 (3.89)**	**12.43 (4.12)**	**15.71 (6.31)**	**16.07 (4.15)**	**17.83 (6.43)**	**5.91**	**.018**
PANSS Negative	15.72 (6.18)	14.00 (3.21)	17.93 (7.51)	16.89 (5.42)	23.00 (7.38)	3.87	.053
PANSS General	**27.64 (5.90)**	**25.14 (7.10)**	**33.21 (7.84)**	**30.78 (7.89)**	**34.00 (10.39)**	**5.22**	**.025**
UPSA	**37.36 (4.25)**	**38.57 (5.22)**	**31.85 (5.29)**	**33.25 (6.88)**	**27.00 (4.52)**	**16.61**	**<0.001**

Note. Bold=statistically significant. Test statistics are *F*-values from ANOVA, and omitted when Fisher's Exact test was used. For continuous values, values are reported as: *M* (*SD*). The significant result for gender was mainly driven by the presence of fewer males and more females than expected by chance in Group 3. Group differences for clinical variables in this table are visualized in the heatmap in Supplementary Figure S1. Abbreviations: AIHQ = Ambiguous Intentions Hostility Questionnaire; BIS/BAS = Behavioral Inhibition and Activation Scale; GFS: Global Functioning Scale; PANSS = Positive and Negative Syndrome Scale; QLS = Quality of Life Scale; SC = Social Cognition; SLOF = Specific Levels of Functioning Scale; TASIT = The Awareness of Social Interaction Test; TEPS = Temporal Associations of Pleasure Scale; WTAR = Wechsler Test of Adult Reading.

Author Manuscript

Author Manuscript

Author Manuscript

Table 4.

Performance of Models Predicting Trajectory Group 1 Membership

	LR Model: *M* (*SD*)	**RF Model: *M* (*SD*)**
Precision	0.62 (0.28)	0.61 (0.17)
Recall	0.67 (0.29)	0.75 (0.36)
F1	0.61 (0.24)	0.67 (0.13)
AUC-PR	0.72 (0.31)	0.73 (0.24)

Note. No Information Rate [AUC-PR]=0.29 [always predict that case is a member of group 1]. No Information Rate [F1]=.45 [again, always predicting a case is a member of group 1]. Decision thresholds: .6 (LR model) and .85 (RF model). AUC-PR = Area Under the Curve-Precision Recall; LR=logistic regression. RF=random forest.