**Supplementary Material 1**

**Supplementary Methods. Statistical analyses**

***Assessment of evidences for reliability and validity of PAUSS scores***

In this section, we followed the international standards of the American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (2014)1. Overall, evidence of reliability refers to the accuracy in measurements should the measurements be repeated. Evidence of validity refers to the scale capability to measure the intended concept/construct.

Evidence for reliability of PAUSS item scores in the ASD and the SSD sample

Separately for the ASD and the SSD sample, we first obtained descriptive data for each of the eight PAUSS item scores. Second, we investigated the evidence for score internal consistency with item-item correlations estimated with Pearson’s product–moment correlation coefficient. We considered item-item correlations in the range of 0.15–0.50 acceptable2. Furthermore, we obtained discrimination indices for the PAUSS items. There are various ways of operationalizing item score discrimination. For ordinal/continuous scales, the corrected item total–correlation index is particularly valuable. This index represents the degree to which differences among participants’ item scores are consistent with differences in their total test scores. The higher the corrected item-total correlation index, the better the item discriminates among high and low overall “scorers” on the test, based on the scores on that item. As a rule, item-total correlation values above .30 are considered acceptable2. Finally, we performed analyses to assess the evidence for item score internal consistency through Cronbach’s alpha and the statistic “alpha if the item is removed”. An alpha between 0.7 and 0.9 was considered acceptable3.

Evidence for validity of the PAUSS to assess autism symptom severity in the ASD and the SSD sample

One of the elements that assess the validity of the structure of an instrument is the convergent validity, which implies the extent to which scores on that instrument are similar to scores of other instruments that are intended to measure the same construct. To explore the evidence of convergent validity of the PAUSS, we carried out two-tailed bivariate correlation analyses between the PAUSS score/subscores and the gold-standard to measure autism symptom severity (ADOS-2 CSS score/subscores), as well as between the PAUSS score/subscores and two proxy scores that measure autistic symptom severity: (i)the ADI-R CBA score/subscores, and (ii) the SRS total score.

In the ASD sample, we first confirmed that PAUSS/CSS/ADIR-CBA/SRS scores were not significantly associated with any potentially confounding demographic/clinical variable (e.g. age, IQ). Since CSS score/subscores were not evenly distributed in ASD, we used Spearman’s Rho tests to assess PAUSS-CSS associations, and Pearson’s r tests to assess PAUSS-ADIR score and PAUSS-SRS score associations.

In the SSD sample, we also explored score for all scores the associations with potentially confounding demographic/clinical variables. We found associations between the IQ and (a) PAUSS total score, (b) Overall CSS, and (c) SRS total score (all rho<-0.500; p.05). We also found an association between parental SES and PAUSS total score (Mann Whitney’s U=.30.500; p<.05). Therefore, using linear regression, we corrected PAUSS scores/subscores by socioeconomic status (SES) and intelligence quotient (IQ), and CSS and SRS scores by IQ, and unstandardized residuals for these scores were saved. Since all scores were not evenly distributed in the sample, we used non-parametric (Spearman’s rho) tests to assess the “corrected-PAUSS - CSS/ADIR-CBA/SRS” associations.

*ASD, “autistic-SSD” and “non-autistic SSD” group comparisons*

Given the sample size (n=26) and the PAUSS total score range in our SSD sample [8-31], and as we considered the PAUSS score a continuous measure of autism symptom severity, we used the median PAUSS score in the SSD group to classify individuals into an “autistic-SSD” (PAUSS>17, *n*=13) and a “non-autistic SSD” (PAUSS≤17, *n*=13) phenotype group. We assessed differences in demographic and clinical variables between the ASD, the “autistic-SSD” and the “non-autistic SSD” groups. Given the sample size, we used non-parametric tests, i.e. Kruskal-Wallis tests with post-hoc Dunn-Bonferroni correction (for continuous variables) and Fisher’s exact tests with post-hoc Bonferroni correction (for categorical variables). We used the same statistical approach to conduct an additional exploratory analysis to compare ASD, early-onset (age at first psychotic episode<18 years) and adult-onset SSD cases (≥18 years) in terms of their PAUSS scores and clinical/demographic variables.

*Association between autistic symptom load (as defined by the PAUSS total score) and global functioning (as defined by the CGAS/GAF score as a quantitative variable) in SSD*

To ascertain the predictive value of the PAUSS total score on individual’s functioning (CGAS/GAF) we ran exploratory multivariate linear regression models. We conducted preliminary linear regressions to ensure no violation of the assumptions of linearity, independence, homoscedasticity, normality, and multicollinearity. Alongside the PAUSS total score, we included in the models the following covariates: (i) variables that were bivariately associated with either the PAUSS total score (IQ and parental SES) or the CGAS/GAF score (PANSS positive score and developmental deviance as defined by the ADI-R diagnostic algorithm), and (ii) other clinically relevant variables that might act on either the PAUSS or the GAF/CGAS score (i.e. age, sex, age at psychosis onset). We did not include the PANSS negative score, PANSS total score and c-PAS score in the models since they were all highly collinear with the PAUSS total score (all r>0.650; p<.01) and with the CGAS/GAF score (all r>0.700; p<.001).

We first ran simple linear regression models for each predictor separately and the CGAS/GAF score as the dependent variable. Then, we ran sequential stepwise multivariate regression models including the PAUSS score plus variable-related blocks of confounders as follows: model 1: PAUSS total score + demographics (age, sex, parental SES); model 2: PAUSS total score + clinical variables (age at psychosis onset, IQ, PANSS positive score); model 3: PAUSS total score + developmental deviance. Finally, to examine whether the PAUSS total score remained as significant predictor of CGAS/GAF, we built a fully adjusted model (model 4), in which the PAUSS total score and all potential confounders were included in a stepwise multivariate regression. We computed the adjusted R2 for all models and all significant predictors within these models.

Given that the PAUSS construct is mainly derived from the PANSS negative subscale (including six out of the seven negative symptom items, plus two general symptom items) we also built a fully adjusted model with the PANSS negative score as the independent predictor and the same potential confounders using stepwise multivariate regression, to compare the variance explained by this predictor relative to the PAUSS score.

**References to Supplementary Material 1**

1. AREA, APA, NCME Standards for Educational and Psychological Testing. Washington: DC: American Educational Research Association; 2014.

2. Nunnally JC, Bernstein IH. The Assessment of Reliability. Psychometric Theory 1994; **3**:248-292.

3. Wu ZJ, Huang Y, Fu YC, Zhao XJ, Zhu C, Zhang Y, et al. Characterization of a Chinese KCNQ1 mutation (R259H) that shortens repolarization and causes short QT syndrome 2. J Geriatr Cardiol 2015; **12**:394-401. https://doi.org/10.11909/j.issn.1671-5411.2015.04.002.