A norm-anchored structural deviation framework for characterizing cognitive heterogeneity in schizophrenia

Chen Chen*
Department of Anatomy, Faculty of Medical Science, Naresuan University, Phitsanulok, 65000, Thailand

*Corresponding author email address: chench61@alumni.nu.ac.th

Abstract

Cognitive heterogeneity is a core feature of schizophrenia (SCZ). Conventional approaches examine this heterogeneity using domain-specific scores, which may not fully reflect the underlying cognitive structure. In this study, a norm-anchored cognitive structural deviation (NCSD) framework was developed to examine such heterogeneity from a structure-informed perspective. The HC-derived latent cognitive structure (N-LCS) captured performance across the assessed tasks and remained stable under external validation in an independent cohort. Patients with SCZ showed greater global deviation from the N-LCS, along with altered loading directions of Wisconsin Card Sorting Test (WCST)-derived executive indicators which were consistent across robustness analyses, and altered correlation patterns among cognitive measures relative to HC. These features were quantified using three NCSD-derived indices: the cognitive normative deviation index (CNDI), loading pattern divergence (LPD), and correlation structure discrepancy (CSD). CNDI discriminated SCZ from HC with stable performance under cross-validation. LPD and CSD were associated with anxiety ratings, with LPD also showing a trend-level association with positive symptoms. Exploratory clustering identified a three-cluster solution with clear separation and moderate stability. Together, these findings show that cognitive heterogeneity in SCZ involves both global deviation from the N-LCS and structural alteration. NCSD provides a refined framework to characterize such heterogeneity and may inform precision psychiatry and functional recovery.

Introduction

Cognitive impairment is a core feature of schizophrenia (SCZ) and a determinant of functional outcome [1]. Its clinical presentation is heterogeneous, with patients differing in both the overall severity of impairment and cognitive profiles [2]. Current clinical and research practice routinely assesses cognition using domain-specific scores. However, accumulating evidence suggests that human cognition is better understood as a structured network, with interdependent domains organized within a broader latent cognitive structure (LCS) [3]. This raises a fundamental question: does the cognitive heterogeneity observed in SCZ reflect quantitative variation relative to the LCS, or a structural reorganization of it?

Two approaches are often discussed. The normative modelling approach establishes a normative reference and interprets SCZ as deviations from it [4]. In contrast, clustering-based approaches seek to identify patient subgroups with distinct cognitive phenotypes [5]. However, these subgroups often reflect differences in severity rather than qualitatively distinct phenotypes [6, 7].

Empirical evidence to date appears to favor the view that cognitive heterogeneity primarily reflects quantitative variation in impairment severity. Yet neuroimaging studies revealed that SCZ involves alterations in the topological organization of brain networks, including disrupted hierarchical structure and abnormal connectivity among hub regions [8, 9]. These findings demonstrate that cognitive heterogeneity in SCZ can be attributed not only to differences in deviation magnitude but also to structural shifts in the LCS relative to a normative reference. Because existing approaches tend to examine these aspects separately, an analytical framework is required to assess them jointly.

In the present study, we developed a norm-anchored cognitive structural deviation (NCSD) framework to assess individual departures from the normative LCS derived from HC (N-LCS). Moving beyond conventional score-based measures of cognition, NCSD enables a refined characterization of cognitive heterogeneity in SCZ and may support individualized clinical assessment within precision psychiatry.

2. Methods

2.1 Data

The primary dataset comprised patients with SCZ (n = 32), their unaffected siblings (UC, n = 32), and HC (n = 33), originally reported by Xu and Xian (2023) and publicly available [10].

For external validation of the N-LCS constructed from the primary dataset, we used the Center of Biomedical Research Excellence (COBRE) schizophrenia dataset [11], accessed via the Collaborative Informatics and Neuroimaging Suite Data Exchange tool (COINS; http://coins.mrn.org/dx). The COBRE dataset included 90 HC and 78 SCZ participants, but only the HC sample was used for validation.

Both datasets contained neurocognitive measures across multiple domains, as well as demographic and clinical variables. All data were fully anonymized, and the present analyses did not require additional ethical approval.

2.2 Neurocognitive measures

To ensure comparability, analyses focused on four cognitive domains assessed in both datasets: processing speed, attention, working memory, and reasoning and problem solving.

In the primary dataset, these domains were assessed using the Wisconsin Card Sorting Test (WCST), Digit Span Forward (DSF), Digit Span Backward (DSB), Verbal Fluency Test (VFT), Stroop Word, Stroop Color, Stroop Interference, Trail Making Test A (TMTA), and Trail Making Test B (TMTB). Three WCST subscores were included: correct responses, perseverative errors (PE), and trials to complete the first category (1st category).

To reduce dimensionality and address multicollinearity, Stroop performance was summarized as a composite score (Stroop-comp). For the Trail Making Test, the difference between TMTB and TMTA (TMT-diff) was calculated to reduce shared variance attributable to basic processing speed. Scores indicating poorer performance were reverse-coded so that higher values consistently reflected better cognitive ability.

In the COBRE dataset, cognitive measures were derived from the MATRICS Consensus Cognitive Battery (MCCB) [12, 13], which provides standardized T-scores for the corresponding domains.

2.3 Data preprocessing

Data were screened for missing values and outliers prior to analysis. Cognitive variables were examined for skewness within the HC group and transformed when necessary, using order-normalization estimated in HC data and applied to all participants.

In the primary dataset, cognitive variables were standardized using HC-based z-scores, so that scores in the UC and SCZ groups were expressed relative to the HC normative reference.

2.4 Normative cognitive modeling and deviation indices

Parallel analysis in the HC group of the primary dataset suggested a single-factor solution [14]. Confirmatory factor analysis indicated that this solution was not rejected ($p = 0.137$), although overall fit was modest (RMSEA = 0.10, 90% CI [0.00, 0.19]) [15]. To establish the N-LCS, a single-factor exploratory factor analysis was conducted. HC-derived loadings were used to compute factor scores for all participants, situating individuals on the normative cognitive dimension.

To examine the stability of the N-LCS in an independent cohort, external validation was conducted in the HC group of the COBRE dataset. The number of factors was evaluated using the Kaiser criterion [16], parallel analysis, and the minimum average partial (MAP) procedure [17], and factor loadings were examined for consistency with those observed in the primary dataset.

Deviation from the HC-derived normative model was quantified using Mahalanobis distance based on the HC covariance matrix [18]. This index was defined as the cognitive normative deviation index (CNDI).

Similarity of loading pattern between HC and SCZ was evaluated using Tucker's congruence coefficient [19], and loading pattern divergence (LPD) was defined as between-group differences in factor loadings. Robustness of these loading differences was assessed using Spearman correlation, bootstrap resampling, and re-estimation following multivariate outlier control.

Correlation structure discrepancy (CSD) was computed by comparing the Pearson correlation matrices between HC and SCZ, identifying variable pairs with the largest absolute differences [20].

2.5 Clinical discrimination, association, and exploratory subgroup analysis of deviation indices

Logistic regression was used to test whether the CNDI could distinguish between SCZ and HC. Model performance was evaluated with 10-fold cross-validation.

Within the SCZ group, linear regression models were fitted to examine associations of CNDI, LPD, CSD with clinical measures. Gaussian mixture modeling (GMM) was applied to the deviation indices to explore whether subgroup could be detected. The number of mixture components (G) was set to 1–3 to account for sample size limitations [21]. Bootstrap resampling was used to assess the stability of the solution.

2.6 Statistical analysis

Group comparisons were conducted using t-tests or one-way analysis of variance (ANOVA), as appropriate. Post hoc pairwise comparisons were performed using Tukey's honestly significant difference (HSD) test. Statistical significance was set at $p < 0.05$ (two-tailed). All analyses were conducted in R (version 4.5.0).

3. Results

3.1 Demographic characteristics

In the primary dataset, the three groups (HC, UC, SCZ) did not differ significantly in age, sex, or years of education (all $p > 0.05$), consistent with the original report [10]. Detailed demographic, cognitive, and clinical information is reported in the original study.

In the COBRE dataset, the HC group had a mean age of 37.5 years (SD = 11.8), 23 were female (25.6%), and the mean years of education was 13.8 (SD = 1.70).

3.2 External validation of N-LCS in COBRE dataset

Validation in the COBRE HC sample confirmed the stability of the N-LCS. The Kaiser criterion indicated a one-factor solution, with the first eigenvalue (2.02) above 1 and subsequent values below [16]. Parallel analysis also supported this solution, and the MAP identified one factor as optimal (Figure 1).

The loading pattern was consistent with that observed in the primary dataset, with stronger loadings for working memory and attention and a weaker loading for reasoning.

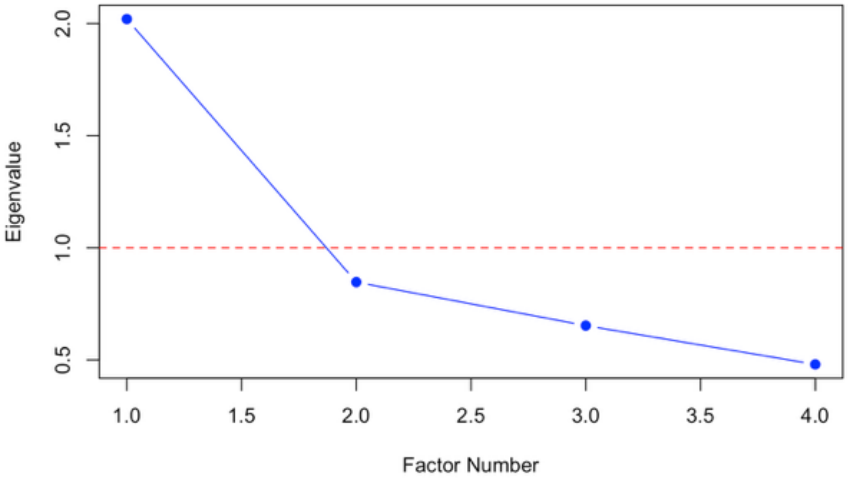

Figure 1. Scree plot of eigenvalues from parallel analysis in the COBRE HC sample

3.3 Group differences in CNDI

CNDI showed significant differences across HC, UC, and SCZ ($p = 2.93 \times 10^{-6}$; see Figure 2 and Table 1). Post hoc Tukey's HSD tests confirmed significant pairwise differences (all $p < 0.05$), with SCZ exhibiting the greatest deviation. A significant linear trend was also observed ($p=1.39 \times 10^{-6}$), with increasing deviation across groups.

Table 1. CNDI across three groups

Group	CNDI	p-value
HC	6.85 ± 2.43	
UC	25.3 ± 20.7	2.93×10^{-6}**
SCZ	44.2 ± 46.6	

Data are presented as mean ± SD and compared using one-way ANOVA.
** $p < 0.01$

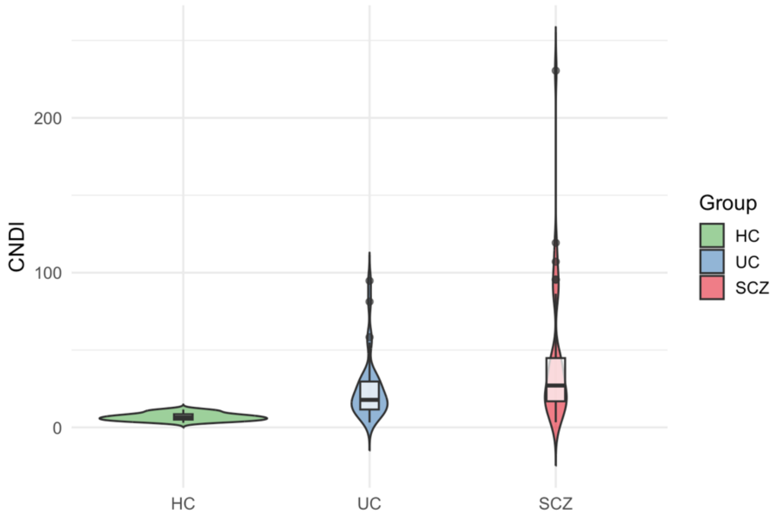

Figure 2. Differences in cognitive normative deviation between groups

3.4 Divergence in factor loading and correlation structure

Tucker's congruence coefficient was 0.73, indicating that the SCZ loading pattern was not structurally equivalent to that of HC [19].

All three WCST subscores showed reversals in loading direction, negative in HC but positive in SCZ (see Figure 3). This reversal was robust across Spearman correlation, bootstrap resampling, and re-estimation after controlling for multivariate outliers. In contrast, the remaining cognitive measures retained consistent loading directions across groups.

CSD analysis revealed that the largest absolute differences were observed for the correlations between 1st category and VFT, and between 1st category and STROOP-comp, both exceeding 0.50, with stronger correlations in SCZ than in HC.

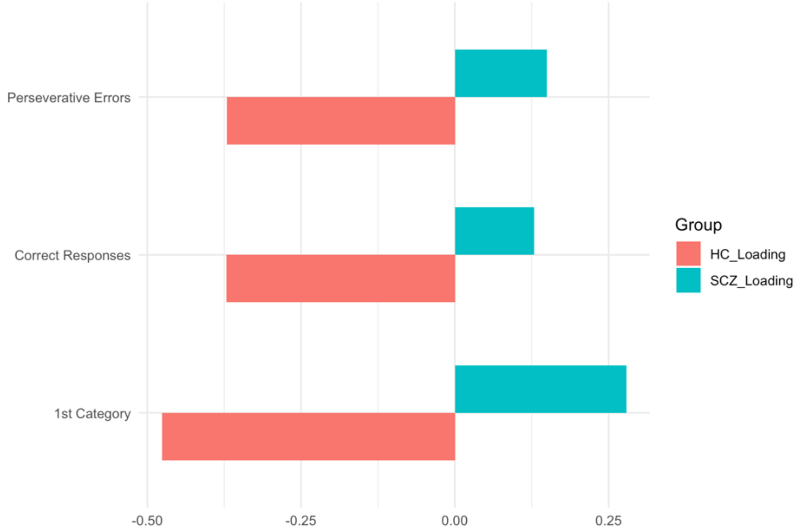

Figure 3. Factor loadings of WCST subscores in HC and SCZ

3.5 Clinical discrimination, association, and exploratory subgroup analysis of deviation indices

3.5.1 Clinical discrimination

CNDI showed good discrimination between SCZ and HC. Logistic regression achieved a high accuracy of 93.9%, with a sensitivity of 0.88 and a specificity of 1.00. The area under the receiver operating characteristic curve (AUC) was 0.92. Ten-fold cross-validation yielded a mean AUC of 0.90, indicating stable model performance (see Figure 4).

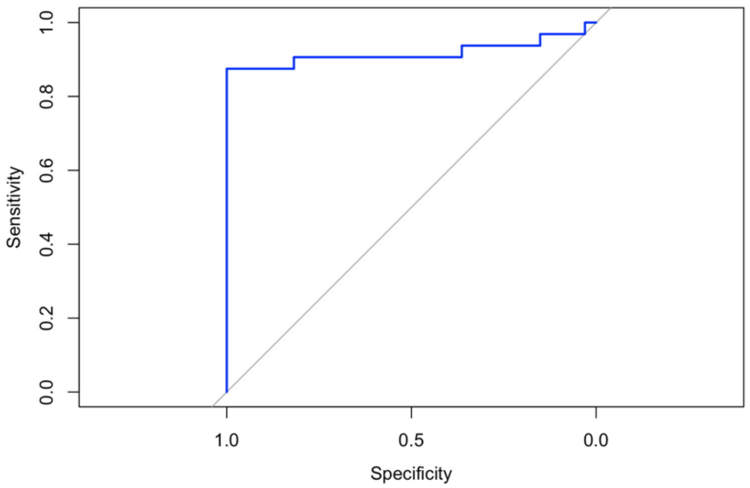

Figure 4. Receiver operating characteristic curve based on ten-fold cross-validation

3.5.2 Symptom associations within SCZ

Within the SCZ group, linear regression models were fitted using CNDI, LPD, and CSD as predictors to examine symptom associations. The model including LPD and CSD was significant for anxiety (HAMA; $R^2 = 0.31$, $p = 0.005$). LPD ($\beta = 1.90$, $p = 0.0045$) and CSD ($\beta = -0.167$, $p = 0.0033$) were independently associated with anxiety. For positive symptoms, LPD showed a trend-level association ($\beta = 2.10$, $p = 0.059$), while other indices were not significant.

3.5.3 Exploratory subgroup structure within schizophrenia

GMM selected a three-cluster solution based on the Bayesian Information Criterion (BIC), with cluster sizes of 18, 8, and 6. The mean posterior membership probability was 0.95, demonstrating clear cluster assignment. Bootstrap resampling yielded a mean adjusted Rand index (ARI) of 0.60, suggesting moderate stability of the solution [22].

Cluster 2 was characterized by high CNDI and low LPD, whereas clusters 1 and 3 both showed moderate LPD but differed in CNDI level, with lower values in cluster 1 and higher values in cluster 3 (see Figure 5).

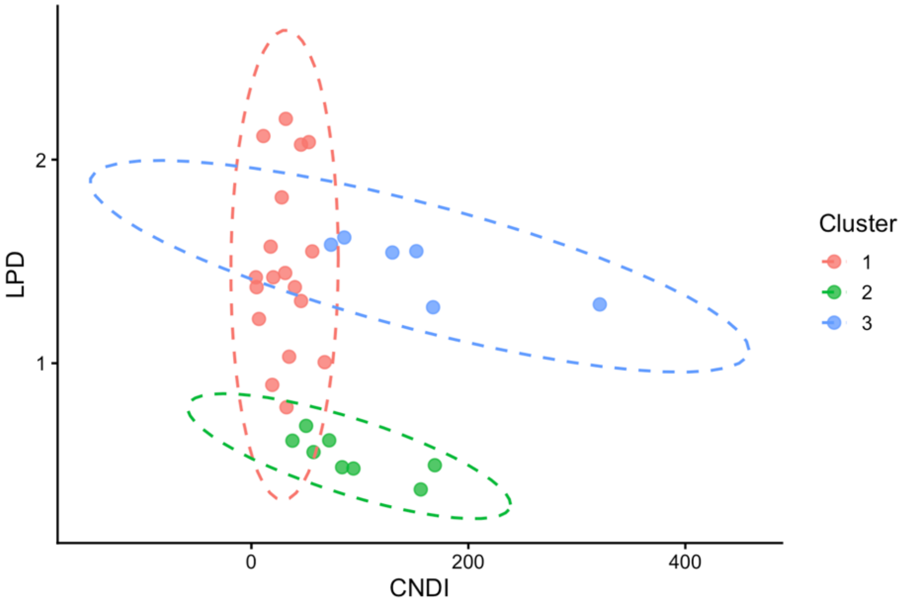

Figure 5. Exploratory three-cluster structure within the schizophrenia group identified using Gaussian mixture modeling

Discussion

Using the NCSD framework, the present study jointly examined deviation severity and structural alterations in cognition among patients with SCZ. The findings indicate a pattern of graded deviations from the N-LCS together with localized shifts in factor loadings.

A stable N-LCS captured performance across the assessed tasks, consistent with evidence that cognitive abilities are explained by a shared underlying dimension, described as a general cognitive factor [23-25]. Work in network neuroscience further suggests that cognition arises from interactions among distributed brain systems, with interdependent domains organized within a broader cognitive architecture that supports efficient information exchange and system-wide flexibility [3, 26]. Individual differences relative to the N-LCS may reflect varying susceptibility to environmental or developmental stressors, which may contribute to the onset of psychiatric disorders under adverse conditions [27]. In line with this, the present study observed a progressive increase in CNDI across HC, UC, and SCZ. Within the SCZ group, altered loading patterns were also found in WCST-derived executive indicators. Such alterations align with meta-analytic neuroimaging evidence showing that, although HC and SCZ recruit broadly similar executive control networks during task performance, core regions within the prefrontal-cingulate control system show altered activity in patients [28]. Together, these findings provide a conceptual basis for applying NCSD to examine cognitive heterogeneity in SCZ. However, it remains unclear whether deviation severity precedes structural alteration or arises as a consequence of it.

External validation using the COBRE dataset confirms the stability of the N-LCS [29]. Originally established in a modest sample, the N-LCS was reproduced in a larger independent cohort, supporting its generalizability. In the primary dataset, the study groups did not differ significantly in age, sex, or education, suggesting that the N-LCS is unlikely to be driven by demographic factors. This interpretation is further supported by replication in COBRE, despite demographic differences between the two cohorts. Moreover, the N-LCS was observed in cohorts that differed in ethnic background and cognitive test batteries, showing that it represents a general organization of cognitive domains rather than a measurement-specific artifact. These properties make the N-LCS a stable structural reference for implementing NCSD.

The deviation indices derived from NCSD showed empirical utility. The CNDI discriminated SCZ from HC with stable performance under cross-validation. This could be explained by cognitive structure capturing group differences more effectively than domain-specific scores [28]. Such scores reduce complex cognitive abilities to a single value [30], limiting their ability to reflect the structured organization of cognition [3]. Additionally, LPD and CSD were associated with HAMA scores, and LPD also showed a trend-level relationship with positive symptoms, echoing previous studies linking anxiety to executive deficits [31] and WCST performance to positive symptoms [32].

Exploratory clustering provided evidence that cognitive impairment in SCZ may not be homogeneous, with a three-cluster solution showing clear separation. One cluster was characterized by high CNDI with minimal LPD, whereas the other two showed similar levels of LPD but differed in deviation magnitude. This underscores that cognitive heterogeneity in SCZ is shaped by both deviation severity and structural alteration, rather than representing only a quantitative continuum or a completely distinct structural reorganization. This interpretation is consistent with findings from a large-scale cognitive profiling study [33]. Moderate bootstrap stability implies that the clustering solution captures meaningful heterogeneity but remains open to refinement in future studies. Results from these indices further demonstrate the empirical utility of NCSD.

Several limitations should be acknowledged. First, the present analysis included four cognitive domains, which may not fully capture the architecture of human cognition. Future studies including additional cognitive domains and alternative cognitive measures are needed to evaluate the robustness and generalizability of NCSD and its derived deviation indices. Second, the cross-sectional design limits our understanding of the temporal dynamics of NCSD and its potential prognostic value. Third, the absence of biological or neuroimaging measures limits the ability to explore the neural correlates and mechanistic basis of NCSD.

In conclusion, this study developed the NCSD framework for examining cognitive heterogeneity in SCZ. Conceptually, NCSD is grounded in prior work on cognitive structure. Empirically, the N-LCS demonstrated stability as a structural reference for implementing NCSD, and the derived deviation indices showed potential for capturing individual cognitive variation. By moving beyond conventional domain-specific approaches, NCSD provides a structure-informed perspective revealing that cognitive heterogeneity in SCZ arises from both deviation severity and structural alteration, offering a practical framework to inform precision psychiatry and functional recovery.

Acknowledgments

The primary dataset used in this study was obtained from Xu and Xian (2023). The validation dataset was downloaded from the Collaborative Informatics and Neuroimaging Suite Data Exchange tool (COINS; http://coins.mrn.org/dx). Data collection for the COBRE dataset was performed at the Mind Research Network and supported by a Center of Biomedical Research Excellence (COBRE) grant (5P20RR021938/P20GM103472) from the National Institutes of Health to Dr. Vince Calhoun.

The investigators associated with these datasets contributed to the design of the original research and data collection but did not participate in the analyses reported in this manuscript or in the preparation of this manuscript. The author gratefully thanks them for making these datasets publicly available.

References

1. Green MF. Cognitive impairment and functional outcome in schizophrenia and bipolar disorder. J Clin Psychiatry. 2006;67 Suppl 9:3-8; discussion 36-42.
2. McCutcheon RA, Keefe RSE, McGuire PK. Cognitive impairment in schizophrenia: aetiology, pathophysiology, and treatment. Mol Psychiatry. 2023;28(5):1902-18.
3. Petersen Steven E, Sporns O. Brain Networks and Cognitive Architectures. Neuron. 2015;88(1):207-19.
4. Marquand AF, Kia SM, Zabihi M, Wolfers T, Buitelaar JK, Beckmann CF. Conceptualizing mental disorders as deviations from normative functioning. Molecular Psychiatry. 2019;24(10):1415-24.
5. Lim K, Smucny J, Barch DM, Lam M, Keefe RSE, Lee J. Cognitive Subtyping in Schizophrenia: A Latent Profile Analysis. Schizophrenia Bulletin. 2020;47(3):712-21.
6. Haining K, Gajwani R, Gross J, Gumley AI, Ince RAA, Lawrie SM, et al. Characterising cognitive heterogeneity in individuals at clinical high-risk for psychosis: a cluster analysis with clinical and functional outcome prediction. European Archives of Psychiatry and Clinical Neuroscience. 2022;272(3):437-48.
7. Vaskinn A, Haatveit B, Melle I, Andreassen OA, Ueland T, Sundet K. Cognitive Heterogeneity across Schizophrenia and Bipolar Disorder: A Cluster Analysis of Intellectual Trajectories. Journal of the International Neuropsychological Society. 2020;26(9):860-72.
8. Bassett DS, Bullmore E, Verchinski BA, Mattay VS, Weinberger DR, Meyer-Lindenberg A. Hierarchical organization of human cortical networks in health and schizophrenia. J Neurosci. 2008;28(37):9239-48.
9. van den Heuvel MP, Sporns O, Collin G, Scheewe T, Mandl RC, Cahn W, et al. Abnormal rich club organization and functional brain dynamics in schizophrenia. JAMA Psychiatry. 2013;70(8):783-92.
10. Xu F, Xian Z. Study investigating executive function in schizophrenia patients and their unaffected siblings. PLoS One. 2023;18(4):e0285034.
11. Çetin MS, Christensen F, Abbott CC, Stephen JM, Mayer AR, Cañive JM, et al. Thalamus and posterior temporal lobe show greater inter-network connectivity at rest and across sensory paradigms in schizophrenia. Neuroimage. 2014;97:117-26.
12. Nuechterlein KH, Green MF, Kern RS, Baade LE, Barch DM, Cohen JD, et al. The MATRICS Consensus Cognitive Battery, part 1: test selection, reliability, and validity. Am J Psychiatry. 2008;165(2):203-13.
13. Kern RS, Nuechterlein KH, Green MF, Baade LE, Fenton WS, Gold JM, et al. The MATRICS Consensus Cognitive Battery, part 2: co-norming and standardization. Am J Psychiatry. 2008;165(2):214-20.
14. Lim S, Jahng S. Determining the number of factors using parallel analysis and its recent variants. Psychol Methods. 2019;24(4):452-67.
15. MacCallum RC, Browne MW, Sugawara HM. Power analysis and determination of sample size for covariance structure modeling [doi:10.1037/1082-989X.1.2.130]. US: American Psychological Association; 1996.
16. Kaiser HF. The application of electronic computers to factor analysis. Educational and Psychological Measurement. 1960;20:141-51.

17. Velicer WF. Determining the number of components from the matrix of partial correlations. Psychometrika. 1976;41(3):321-7.
18. Reprint of: Mahalanobis, P.C. (1936) "On the Generalised Distance in Statistics.". Sankhya A. 2018;80(1):1-7.
19. Lorenzo-Seva U, ten Berge JMF. Tucker's Congruence Coefficient as a Meaningful Index of Factor Similarity. Methodology. 2006;2(2):57-64.
20. Steiger JH. Tests for comparing elements of a correlation matrix. Psychological Bulletin. 1980;87(2):245-51.
21. Fraley C, Raftery AE. Model-Based Clustering, Discriminant Analysis, and Density Estimation. Journal of the American Statistical Association. 2002;97(458):611-31.
22. Warrens MJ, van der Hoef H. Understanding the Adjusted Rand Index and Other Partition Comparison Indices Based on Counting Object Pairs. Journal of Classification. 2022;39(3):487-509.
23. Spearman C. 'General intelligence,' objectively determined and measured. The American Journal of Psychology. 1904;15(2):201-93.
24. Jensen AR. The g factor: The science of mental ability. Westport, CT, US: Praeger Publishers/Greenwood Publishing Group; 1998. xiv, 648-xiv, p.
25. Zhang L, Feng J, Liu C, Hu H, Zhou Y, Yang G, et al. Improved estimation of general cognitive ability and its neural correlates with a large battery of cognitive tasks. Cerebral Cortex. 2024;34(2).
26. Barbey AK. Network Neuroscience Theory of Human Intelligence. Trends in Cognitive Sciences. 2018;22(1):8-20.
27. McEwen BS, Gianaros PJ. Stress- and allostasis-induced brain plasticity. Annu Rev Med. 2011;62:431-45.
28. Minzenberg MJ, Laird AR, Thelen S, Carter CS, Glahn DC. Meta-analysis of 41 Functional Neuroimaging Studies of Executive Function in Schizophrenia. Archives of General Psychiatry. 2009;66(8):811-22.
29. Salazar de Pablo G, Studerus E, Vaquerizo-Serrano J, Irving J, Catalan A, Oliver D, et al. Implementing Precision Psychiatry: A Systematic Review of Individualized Prediction Models for Clinical Practice. Schizophr Bull. 2021;47(2):284-97.
30. Van Der Maas HLJ, Dolan CV, Grasman RPPP, Wicherts JM, Huizenga HM, Raijmakers MEJ. A dynamical model of general intelligence: The positive manifold of intelligence by mutualism. Psychological Review. 2006;113(4):842-61.
31. Korczak A, Styła R. Anxiety and executive functions relationships in schizophrenia: A meta-analysis. Personality and Individual Differences. 2021;177:110643.
32. Stratta P, Daneluzzo E, Bustini M, Prosperini P, Rossi A. Processing of context information in schizophrenia: relation to clinical symptoms and WCST performance. Schizophrenia Research. 2000;44(1):57-67.
33. Huang Y, Wang W, Hei G, Shao T, Li L, Yang Y, et al. Subgroups of cognitive impairments in schizophrenia characterized by executive function and their morphological features: a latent profile analysis study. BMC Medicine. 2025;23(1):13.