

Distinguishing and Replicable Features of Youth Experiencing Psychotic-Like Experiences Via Multimodal Data Integration

Aristotle N Voineskos

aristotle.voineskos@camh.ca

Centre for Addiction and Mental Health

Julia Gallucci

Teodora Secara

Grace Jacobs

Prashanth Velayudhan

Anne L Wheeler

University of Toronto

Lauren Erdman

Stephanie Ameis

Colin Hawco

Centre for Addiction and Mental Health

Article

Keywords:

Posted Date: October 18th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10973424/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: There is **NO** Competing Interest.

Abstract

Although findings from population-level studies in human development have made important contributions to our understanding of normative brain, cognitive, and social development, few focus on children and youth with more significant impairment. Delineating those youth with risk profiles from others would be an important advance. Using data from the Adolescent Brain Cognitive Development Study, we examined 1,716 youth reporting distressing psychotic-like experiences (PLEs) (discovery $n = 858$; replication $n = 858$). Multimodal data, encompassing demographic, clinical, cognitive, neuroimaging, and socio-economic features were integrated using Similarity Network Fusion and clustered to identify data-driven subgroups. Cluster validity was evaluated via resampling, held-out measures, and longitudinal trajectories. Identified subgroups were also compared to a large 'normative sample' ($n = 5,225$). Five replicable subgroups emerged across discovery and replication samples, differing in cognitive performance, environmental adversity, and neural integrity. Subgroups were highly consistent, with an 89% overlap in top 15% contributing features, and demonstrated robust stability via resampling. Separation extended to out-of-model variables, and longitudinal analyses showed clear cognitive and distress trajectories. These replicated findings separate groups of youth with risk profiles from others, furthering a more precision-based understanding of the demographic, clinical, cognitive, neuroimaging, and socioeconomic factors underpinning such risk.

Introduction

Youth mental health has become a major public health concern ¹, yet large investments in interventions have yielded limited population-level gains ². One challenge is that not all children and adolescents are at equal risk, underscoring the need to identify those most likely to experience persistent adverse outcomes. Psychotic-like experiences (PLEs) represent an early risk indicator. PLEs are relatively common in youth, with an estimated prevalence of 17% in children and 8% in adolescents ³, and occur at substantially higher rates ($\sim 50\%$) among those with existing mental health disorders ⁴. Distressing PLEs, in particular, are associated with elevated risk for subsequent psychiatric illness ⁵ and suicidality ⁶. Converging evidence further links PLEs to lower cognitive ability, greater environmental adversity, and altered brain development, highlighting their relevance as early markers of vulnerability in adolescence ^{7–11}. Not all youth with PLEs experience adverse long-term outcomes, yet for those who do, consequences can be substantial. Developing a more precise understanding of neurodevelopmental, cognitive, and environmental features that collectively characterize youth who experience distressing and persisting PLEs may guide early intervention.

We used data from the Adolescent Brain Cognitive Development (ABCD) study, a large longitudinal investigation tracking over 11,000 children starting from ages 9–10 across the United States, from varying geographic regions and socioeconomic statuses ¹². ABCD's multimodal design encompasses neurocognitive assessment, brain imaging, and environmental measures, offering a robust framework for investigating various domains that shape mental health outcomes across development. Ongoing

refinement of features that may better characterize who among the general population of youth is at greatest risk of poor mental health outcomes is needed at a time of heightened concern in this population. In this study, we applied a data integration method, known as Similarity Network Fusion (SNF), which groups participants based on similarities across multiple data modalities¹³, to identify data-driven subgroups of youth with distressing PLEs. To rigorously assess reproducibility, we randomly split the sample into equal-sized discovery and replication subsets. We were interested in identifying the relative importance of cognitive, brain, and environmental measures in distinguishing groups in distress, and how that might compare to children not in distress. To evaluate out-of-model generalization, selected brain and behavioral variables were excluded from SNF and examined separately. We were also interested in how these groups might separate in relation to key outcome measures over time, such as persistence of PLEs or cognitive deficits, which are indicators of poorer mental health outcomes^{14–16}.

Methods and Materials

Participants

Data from the ABCD Study v5.1 release (doi:10.15154/z563-zd24) were analyzed (see Supplementary Figure S1 for CONSORT-style flow diagram). Of the 11,868 participants with baseline ABCD data available, 2,505 were excluded due to low-quality MRI data or excessive motion during resting-state fMRI (mean framewise displacement > 0.2 mm). Participants missing behavioral (n = 685) or environmental data (n = 713) were also excluded, along with 1,024 siblings, to minimize genetic confounds in clustering. To reduce zero inflation and emphasize relevant cognitive, environmental, and neural features, we further excluded 5,225 participants who did not endorse at least one distressing PLE, defined as present and at least moderately distressing (see ‘Clinical Measures’ section below for details). The primary sample consisted of 1,716 youth aged 9–10 years at the time of baseline measurement, endorsing distressing PLEs with complete, high-quality multimodal data. This sample was randomly divided into discovery (n = 858) and replication (n = 858) to enable independent validation of clustering results, while maintaining similar distributions of key demographic variables, including age, sex, and race/ethnicity. The NonPLE group (n = 5,225), who met all inclusion criteria but did not endorse any distressing PLEs, was retained as a reference sample. Demographic characteristics of the NonPLE group were comparable to the primary sample (Supplementary Figure S2). All participants provided informed consent, and data access was granted under the ABCD Study data use agreement.

Measures

Clinical Measures

PLEs were assessed using the Prodromal Questionnaire–Brief Child Version (PQ-BC), a 21-item self-report measure validated for school-age children^{17,18}. The PQ-BC evaluates subclinical psychotic symptoms experienced over the past month, including unusual thoughts and perceptual abnormalities. A

distress-weighted total PLE score was computed^{16,19}. Responses were coded as 0 = no, 1 = yes without distress, and 2–6 = yes with distress (rated from 2 “not very bothered” to 6 “extremely bothered”), yielding a total score range of 0–126. Distressing PLEs were defined as any item rated at least moderately bothersome (score > 3).

Psychopathology was assessed using the Child Behavior Checklist (CBCL), a 113-item parent-report measure of emotional and behavioral problems²⁰. Items were rated on a three-point Likert scale (0 = absent, 1 = occurs sometimes, 2 = occurs often). CBCL items compose eight syndrome scales: anxious/depressed, withdrawn/depressed, somatic complaints, social problems, thought problems, attention problems, rule-breaking problems, and aggressive behavior. These scales were used for out-of-model validation of group differences.

Cognitive Measures

Baseline cognitive performance was assessed using tasks from the NIH Toolbox: Picture Vocabulary (language), List Sorting (working memory), Flanker Inhibitory Control (attention), Pattern Comparison (processing speed), Picture Sequence (episodic memory), Dimensional Change Card Sort (executive function), and Oral Reading Recognition (language and academic skills)²¹. Participants also completed the Rey Auditory Verbal Learning Test (verbal learning and memory)²², Little Man Task (visuospatial processing and attention)²³, and WISC-V Matrix Reasoning task (fluid intelligence)²⁴. These ten measures provided a comprehensive assessment of cognitive function and were used as input features in the SNF analysis.

Environmental Measures

A wide array of environmental factors were assessed at baseline through caregiver- and youth-reported questionnaires, geocoded data, and census-derived metrics. Dimensional and summary scores were used. Variables with more than 10% missing data were excluded. For variables with $\leq 10\%$ missingness, missing values were imputed using k-nearest neighbors (kNN) imputation. Imputation was applied only to environmental variables to preserve as much information as possible across these diverse and externally sourced measures. In total, 41 measures were retained and used as environmental input features in the SNF analysis (see Supplementary Table S1 for details). These included indicators of socioeconomic status, social environment, developmental adversity, and proximal environmental exposures. Socioeconomic and proximal factors included parental education, household income, neighborhood safety, walkability, air pollution (NO₂ and PM), proximity to roads, lead exposure risk, and local crime and substance availability. Social environment indicators captured family dynamics and social support, including caregiver and parental support, family conflict, parental monitoring, peer relationships, experiences of discrimination, and school engagement. Developmental adversity

encompassed early life health and perinatal complications, such as birth weight, prematurity, parental age, breastfeeding history, developmental milestones, and prenatal exposure to substances.

Neuroimaging Measures

Baseline structural, functional, and diffusion MRI features were derived from preprocessed tabulated data. Structural scans were excluded based on both automated acquisition quality control and manual FreeSurfer quality control, following the ABCD protocol recommendations and guidelines²⁵. Structural metrics included cortical thickness and surface area across 68 cortical regions and subcortical volumes across 26 regions, based on the Desikan-Killiany atlas²⁶, with subcortical volumes normalized by intracranial volume. Resting-state functional MRI (rs-fMRI) data were acquired over four 5-minute runs (totaling 20 minutes), during which participants were instructed to keep their eyes open and fixate on a crosshair. Functional connectivity estimates were based on pairwise correlations between regions defined by the 13-network Gordon parcellation²⁷. Fisher Z-transformed correlation values were provided for all within- and between-network pairs ($n = 91$), as well as correlations between cortical networks and subcortical regions of interest ($n = 247$). These structural and functional measures served as input features for the SNF analysis. Diffusion MRI metrics were derived using AtlasTrack²⁸, a probabilistic atlas-based method for segmenting major white matter tracts. Fractional anisotropy (FA) and mean diffusivity (MD) across individual tracts, as well as summary measures averaged across all tracts, and separately for the left and right hemispheres ($n = 42$); diffusion measures were used for out-of-model validation of group differences. All neuroimaging features were harmonized across sites using ComBat with age and sex as covariates^{29,30}.

Data Integration and Clustering

SNF was applied separately to discovery and replication samples to integrate participant similarity across six data modalities: cognitive, environmental, structural imaging (cortical thickness, surface area, subcortical volume), and resting-state functional connectivity. For each modality, an individual-level similarity network was constructed based on Euclidean distances, using a nearest neighbors (k) value of 44 and a normalization parameter (α) of 0.7. These hyperparameters were selected through a grid search using the metaSNF framework³¹, using distress-weighted total PLE score as the optimization target (see Supplement for details). SNF then applied a nonlinear combination method to iteratively fuse the modality-specific networks into a single integrated similarity network, capturing shared structure across all data types. Normalized Mutual Information (NMI) was used to quantify the degree of overlap between similarity matrices derived from individual features and the fused similarity matrix incorporating all features (range from 0 to 1), with higher NMI scores indicating greater contribution to participant similarity. The final fused similarity network in each sample was used as input for spectral clustering to identify participant subgroups. See Fig. 1 for an overview of the analysis pipeline.

Cluster Stability

Internal cluster reproducibility was evaluated in discovery and replication samples using resampling across 1,000 iterations, each sampling 80% of participants. Stability was quantified by calculating the percentage of times each participant clustered with every other participant, along with the Adjusted Rand Index (ARI), which assesses consistency in cluster assignments across iterations^{32–34}.

Cluster Comparisons Across SNF Features

In both the discovery and replication samples, ANOVA was used to assess cluster differences across input features within each data modality, accompanied by standardized effect size estimates using eta squared. Where significant effects were detected, post hoc Tukey's HSD tests were performed to identify pairwise contrasts. False discovery rate (FDR) correction was applied to adjust for multiple comparisons. For imaging modalities, two-sample t-tests were conducted for each region of interest (structural) and each connectivity edge (functional), comparing each cluster to a reference group of youth with no distressing PLEs (NonPLE group). The resulting t-statistics for each region or edge were used to visualize the magnitude and direction of group differences.

Out-of-Model Features and Longitudinal Cluster Validation

To assess whether data-driven group differences extended beyond the features used in the SNF analysis, out-of-model white matter and psychopathology measures were analyzed using the same statistical approach as the measures included in SNF.

To evaluate the longitudinal validity of the identified clusters, we analyzed both cognitive performance and trajectories of distressing PLEs. Distressing PLE trajectories were examined using latent growth modeling. Specifically, the total PLE score weighted by distress (as measured by the PQ-BC) was modeled across three longitudinal time points: baseline, 1-year, and 2-year follow-up. These ABCD time points were selected to maximize the availability of complete data for longitudinal modeling (i.e., the greatest attrition, particularly in the most impaired subgroup, occurred after the 2-year follow-up, guiding the decision to limit analyses to this period). Upon verifying that discovery and replication samples demonstrated highly similar clustering solutions, the full primary sample was used to maximize statistical power and generalizability ($n = 1,545$). We applied both Latent Class Growth Analysis (LCGA; with fixed intercept and slope) and Growth Mixture Modeling (GMM; with random slope and/or intercept) to identify distinct longitudinal trajectories. Models with one to five classes were tested, and model selection was guided by multiple fit indices, including the Bayesian Information Criterion (BIC), and maximum log likelihood, as well as by class interpretability and adequate size consistent with prior recommendations³⁵. A chi-square test was used to examine the distribution of longitudinal PLE classes across clusters. Post hoc analysis of standardized residuals, reflecting deviations between observed and expected proportions for each latent class, was conducted to identify significant over- or under-representation within clusters, with FDR correction applied for multiple comparisons. Finally, we conducted multinomial logistic regression to examine whether cluster membership provided incremental

predictive value for longitudinal PLE class beyond baseline distress-weighted PLE severity. Model fit was compared using McFadden's pseudo- R^2 and likelihood ratio testing.

Cognitive performance was assessed at baseline and 2-year follow-up using a composite score derived from SNF cognitive measures (excluding the WISC-V Matrix Reasoning task, as it was collected only at baseline). These ABCD time points were selected to maximize the use of complete data for repeated measures modelling. For each measure at each time point, the mean and standard deviation in the NonPLE group ($n = 2,117$) were used as a reference to compute z-scores for all participants, which were then averaged within individuals to produce a single composite score for each time point. Similar to above, the full primary sample was used ($n = 1,565$). Longitudinal group differences in the composite score were evaluated using repeated-measures ANOVA, allowing us to test for both main effects of group and time as well as Group $\times$ Time interactions. Post hoc pairwise contrasts were conducted to characterize differences between clusters and relative to the NonPLE group at each time point, with FDR correction applied for multiple comparisons.

Code

The analysis code and a comprehensive list of input features for this study are available at https://github.com/juliagallucci/ABCD_SNF/

Results

Five-Cluster Solution with Consistent Feature Contributions

Following SNF and spectral clustering, a five-cluster solution was identified in both the discovery and replication samples (Supplementary Figure S3), with good to strong cluster cohesion and separation (average silhouette widths = 0.79 and 0.71, respectively; Supplementary Figure S4). Cluster sizes were as follows: in the Discovery sample, Cluster 1 $n = 102$, Cluster 2 $n = 179$, Cluster 3 $n = 159$, Cluster 4 $n = 247$, and Cluster 5 $n = 171$; in the Replication sample, Cluster 1 $n = 85$, Cluster 2 $n = 238$, Cluster 3 $n = 127$, Cluster 4 $n = 228$, and Cluster 5 $n = 180$. Participant demographics and behavioral scores stratified by cluster are presented in Supplementary Table S2, with Supplementary Figure S5 illustrating cluster-wise PLE symptom profiles. The top 15% of features contributing most strongly to clustering, based on NMI scores, showed an 89% overlap across discovery and replication samples, demonstrating strong consistency in key features (Fig. 2; **right**). Top features included measures of cognitive performance, environmental adversity, cortical thickness, and subcortical volume across both samples. Nine of ten cognitive tests ranked among the top features across samples, spanning domains of language, memory, attention, processing speed, executive function, visuospatial skills, crystallized, and fluid intelligence. Key environmental contributors included Area Deprivation Index, household income, and substance availability, reflecting socioeconomic and neighborhood-level influences. Neural contributors encompassed widespread areas previously implicated in clinical high-risk individuals, including cortical

regions in the frontal (e.g., pars triangularis, pars opercularis, precentral gyrus), temporal (e.g., superior temporal sulcus, fusiform gyrus), and parietal (e.g., precuneus, postcentral gyrus) cortices, as well as subcortical regions (e.g., hippocampus, amygdala)^{36–38}. Resting-state connectivity and surface area features contributed minimally, largely ranking in the bottom half of NMI scores.

Cluster Stability

In the discovery sample, participants clustered with their assigned group an average of 74.6% of the time, compared to 6.1% with other groups (ARI = 0.60; Fig. 2, **top left**), indicating moderate-to-high reproducibility of the clustering solution. The replication sample showed similar stability, with 65.0% within-group clustering compared to 8.4% across-groups (ARI = 0.47; Fig. 2, **bottom left**), demonstrating consistency in participant assignment across iterations³⁴.

Cluster Differences in Input Model Features

Significant group effects were observed across all input modalities ($pFDR < 0.05$), with the largest differences in cognitive performance, environmental adversity, cortical thickness, and subcortical volume. Cluster 1 showed the lowest cognitive performance and highest environmental adversity, whereas Cluster 5 showed the opposite (Fig. 3; Supplementary Tables S3-6). Differences may in part reflect demographic composition: Cluster 1 comprised younger participants (discovery: 9.7 ± 0.6 years; replication: 9.5 ± 0.5 years) and a higher proportion of Black youth (discovery: 40.6%; replication: 35.3%), while Cluster 5 included older participants (discovery: 10.2 ± 0.6 years; replication: 10.0 ± 0.6 years) and significantly more White youth (discovery: 59.6%; replication: 62.2%). Clusters 2–4 exhibited intermediate cognitive and environmental profiles but were distinguished by structural brain features relative to the NonPLE group: Cluster 2 showed increased subcortical volume, Cluster 3 showed reduced cortical thickness and subcortical volume, and Cluster 4 showed increased cortical thickness (Fig. 4; Supplementary Tables S7-10). Patterns were largely consistent across discovery and replication. Group differences in resting-state connectivity and surface area are reported in the Supplement (Supplementary Figures S6–S8; Supplementary Tables S11–S14).

Cluster Differences in Out-of-Model Features

Significant group differences were observed across out-of-model features in discovery and replication samples ($pFDR < 0.05$), indicating cluster separation extends beyond features used for subgroup identification. Cluster 5 showed higher FA and lower MD, indicating greater white matter integrity (Supplementary Table S15-16), while Cluster 1 exhibited elevated externalizing symptoms, including attention, social, and rule-breaking problems, relative to Clusters 4 and 5 (Supplementary Tables S17–S18).

Cluster Differences in Longitudinal PLE and Cognitive Trajectories

The most parsimonious solution for distressing PLEs trajectories was a three-class GMM with random effects for intercept and slope (maximum log-likelihood=-17521.22; BIC = 35,145.25; Fig. 5A). Supplementary Table S19 summarizes models derived from longitudinal LCGA and GMM, including model specifications, fit indices, and class proportions. The largest class, the *low trajectory* (n = 1,192; 77.2%), was characterized by low baseline symptom severity (intercept = 12.33) that remained relatively stable over time (slope=-4.72). The *remitting trajectory* (n = 149; 9.6%) began with elevated symptom severity (intercept = 30.50) and showed a steep decline over the follow-up period (slope=-12.87). The *persistently high trajectory* (n = 204; 13.2%) was characterized by high baseline symptom severity (intercept = 32.67), which remained elevated relative to other classes throughout the follow-up period (slope=-4.99). Trajectory distributions significantly differed across clusters ($\chi^2(8, n = 1545) = 33.29$, $p < .001$; Fig. 5B–C). The *persistently high trajectory* was significantly overrepresented in Cluster 1 (Std. Residual = 4.3, $pFDR < .001$) and underrepresented in Cluster 5 (Std. Residual=-2.7, $pFDR=.01$). The *low trajectory* was significantly underrepresented in Cluster 1 (Std. Residual=-3.2, $pFDR=.005$) and overrepresented in Cluster 4 (Std. Residual = 2.7, $pFDR=.01$). The *remitting trajectory* was significantly underrepresented in Cluster 4 (Std. Residual=-2.4, $pFDR=.02$) and overrepresented in Cluster 2 (Std. Residual = 2.2, $pFDR=.04$). Multinomial logistic regression found that while baseline PLE severity strongly predicts trajectories (pseudo- $R^2 = 0.23$), adding cluster membership significantly improves model fit ($\chi^2(8) = 68.32$, $p < .001$).

Cognitive performance showed a significant Group×Time interaction ($F(3,678) = 20.02$, $p < .001$). At baseline, Cluster 1 performed lowest, Clusters 2–4 scored in the intermediate range, and Cluster 5 scored highest; all clusters differed significantly from the NonPLE group (Figure S9; Supplemental Table S20). By year 2, this general ordering persisted, although Cluster 1 showed slight improvement and Cluster 5 a slight decline towards the NonPLE group, a pattern that may partly reflect regression to the mean.

Discussion

By integrating cognitive, environmental, and neuroimaging data, we identified distinct, replicable subgroups of youth with distressing PLEs. Subgroup distinctions were highly consistent across discovery and replication samples, with 89% overlap among the top 15% contributing features, and robust internal stability via resampling. Cluster separation extended to out-of-model features, including clinically (i.e., psychopathology) and biologically (i.e., white matter integrity) meaningful metrics that did not contribute to cluster identification. Cognitive differences evident at baseline largely persisted over time, and subgroups were associated with distinct longitudinal PLE trajectories. Together, findings suggest that multi-level data integration can reliably distinguish risk profiles for youth experiencing divergent outcomes.

While distressing PLE are associated with numerous adverse outcomes^{14–16}, they do not guarantee poor trajectories. Our analysis revealed marked heterogeneity in underlying features, suggesting that no

single neurobiological or developmental signature characterizes all affected youth³⁹. Two clusters differed in cognitive performance and environmental adversity, consistent with evidence suggesting cognitive ability mediates the influence of environmental (e.g., stress, trauma, socioeconomic disadvantage) and genetic risk factors on current and future PLEs, highlighting cognition as a potential pathway through which broader exposures shape psychopathological outcomes^{40,41}. Three clusters showed widespread structural brain differences, with global variations in cortical thickness and subcortical volume rather than region-specific changes, aligning with findings in population-based and high-risk samples linking PLEs to diffuse morphometric alterations^{16,42}.

Several large-scale cross-sectional studies have linked PLEs to cognitive, environmental, and developmental impairments⁷⁻⁹, as well as neural alterations^{10,11}; however, most focus on single data modalities. Our findings highlight the value of multi-domain integration⁴³, and approaches like SNF for identifying subgroups with shared risk and resilience profiles. For example, we demonstrated that some youth with distressing PLEs had intact cognition, associated with improved long-term outcomes. A minority of youth with distressing PLEs followed trajectories marked by either persistent or remitting symptoms, consistent with recent findings⁴¹. These patterns differed substantially across clusters. Cluster 1 (low cognitive performance, high environmental adversity) showed significant overrepresentation of the persistently high trajectory and underrepresentation of the low trajectory, whereas Cluster 5 (high cognitive performance, minimal environmental adversity) showed lower representation in the persistently high trajectory. Clusters 2 and 4, characterized by increased subcortical volume and cortical thickness, respectively, were disproportionately associated with remitting or low trajectories. Despite intermediate cognitive and environmental profiles, these global morphometric features may reflect a neurobiological marker of resilience. Cognitive strength, supportive environments, and healthy brain structure are associated with lower levels of persistent distress and greater symptom remission.

Neurocognitive development during childhood and adolescence is critical, as subtle deviations can elevate risk for later mental health problems⁴⁴. The Dunedin Multidisciplinary Health and Development Study showed early cognitive changes can precede psychotic symptoms, underscoring the value in detecting deficits in youth¹⁴. This work also highlighted the importance of normative frameworks for understanding typical cognitive development. Using a NonPLE reference group, we contextualized cognitive performance across clusters, identifying subgroups that diverged from typical trajectories. PLEs themselves also signal heightened risk for future challenges. The Avon Longitudinal Study of Parents and Children demonstrated that PLEs in childhood and young adulthood predict adverse outcomes, including educational, occupational, social, and substance-related problems¹⁵. Similarly, youth with persistent, distressing PLEs exhibit impairments across multiple domains, including cognition and psychopathology, emphasizing long-term challenges faced and the need for early intervention¹⁶. Our findings suggest that distressing PLEs combined with lower cognitive ability and greater environmental adversity may increase risk for persistent impairment and serve as early indicators of later difficulties.

We employed multiple approaches to evaluate stability and meaningfulness of our clusters, consistent with recent validation guidelines⁴⁵. The five-cluster solution replicated in a held-out sample, with high overlap among top contributing features, and showed strong participant-level stability via resampling. Importantly, out-of-model and longitudinal analyses showed that baseline-derived clusters captured meaningful differences beyond model inputs. Cognitive profiles remained stable over follow-up, and clusters predicted future PLE outcomes, underscoring their clinical relevance.

While large, population-based samples like ABCD enhance generalizability, many participants reported minimal or no distressing PLEs. To isolate relevant correlates, we focused on youth with distressing PLEs, given evidence that distress confers greater prognostic value⁴⁶. Although this reduced the sample, the subset remained well powered for discovery and replication and better aligned with what is clinically important. Additionally, dropping missing data or using single imputation may affect representativeness⁴⁷, and generalizability may be limited by the data reduction, analytic methods employed (i.e., SNF, and clustering). While our findings were replicated internally, future work should validate these results in an independent dataset. The lack of contribution from rs-fMRI and surface area should also be acknowledged. These null findings may reflect methodological limitations, as resting-state connectivity is sensitive to motion and noise, particularly in pediatric samples⁴⁸, which can reduce its influence on clustering relative to more robust features. Further, surface area differences may be subtle at this developmental stage compared to cortical thickness⁴⁹ and may contribute less to clustering solutions, which are driven more by features with stronger variation or signal-to-noise ratios. Finally, although most participants clustered consistently across resampling, some did not fit 'neatly' into a single subgroup, reflecting within-cluster heterogeneity. While continuum-based approaches better capture individual variability, discrete data-driven subgroups remain useful for risk stratification and targeted intervention⁵⁰.

Sufficient evidence has accumulated to suggest PLEs are a cause for concern during the developmental phase, although they may subside for many youth. However, the importance of knowing when they may persist, and the factors underlying their persistence, particularly when they cause distress is useful. Our findings reveal a group of youth with higher and distressing PLEs that are much more likely to persist compared to other groups, and show that environmental adversity is the most important associated factor.

Declarations

Disclosure

All authors have reported no biomedical financial or potential conflicts of interest.

Funding Declaration

No funding to disclose.

Author Contributions:

J.G. and M.T.S. contributed to study conception, data analysis, interpretation, and manuscript drafting. G.R.J., P.S.V., A.L.W., L.E., and S.H.A. contributed to data analysis and interpretation, as well as manuscript drafting. C.H. and A.N.V. contributed to study conception, data interpretation, manuscript drafting, and provided overall supervision and accountability.

Data Availability Declaration

The ABCD data are publicly available upon request at <https://docs.abcdstudy.org/>

References

1. McGorry, P., Gunasiri, H., Mei, C., Rice, S., Gao, C.X.: The youth mental health crisis: analysis and solutions. *Front. Psychiatry*. **15**, 1517533 (2024)
2. Montero-Marin, J., Allwood, M., Ball, S., et al.: School-based mindfulness training in early adolescence: what works, for whom and how in the MYRIAD trial? *Evid. Based Ment Health*. **25**(3), 117–124 (2022)
3. Kelleher, I., Connor, D., Clarke, M.C., Devlin, N., Harley, M., Cannon, M.: Prevalence of psychotic symptoms in childhood and adolescence: a systematic review and meta-analysis of population-based studies. *Psychol. Med*. **42**(9), 1857–1863 (2012)
4. Cleverley, K., Foussias, G., Ameis, S.H., et al.: The Toronto Adolescent and youth Cohort Study: Study design and early data related to psychosis spectrum symptoms, functioning, and suicidality. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging*. **9**(3), 253–264 (2024)
5. Lindgren, M., Numminen, L., Holm, M., Therman, S., Tuulio-Henriksson, A.: Psychotic-like experiences of young adults in the general population predict mental disorders. *Psychiatry Res*. **312**(114543), 114543 (2022)
6. DeVyllder, J.E., Lukens, E.P., Link, B.G., Lieberman, J.A.: Suicidal ideation and suicide attempts among adults with psychotic experiences: data from the Collaborative Psychiatric Epidemiology Surveys: Data from the collaborative psychiatric epidemiology surveys. *JAMA Psychiatry*. **72**(3), 219–225 (2015)
7. Cannon, M., Caspi, A., Moffitt, T.E., et al.: Evidence for early-childhood, pan-developmental impairment specific to schizophreniform disorder: results from a longitudinal birth cohort. *Arch. Gen. Psychiatry*. **59**(5), 449–456 (2002)
8. Karcher, N.R., Schiffman, J., Barch, D.M.: Environmental risk factors and psychotic-like experiences in children aged 9–10. *J. Am. Acad. Child. Adolesc. Psychiatry*. **60**(4), 490–500 (2021)

9. Calkins, M.E., Moore, T.M., Merikangas, K.R., et al.: The psychosis spectrum in a young U.S. community sample: findings from the Philadelphia Neurodevelopmental Cohort. *World Psychiatry*. **13**(3), 296–305 (2014)
10. Schoorl, J., Barbu, M.C., Shen, X., et al.: Grey and white matter associations of psychotic-like experiences in a general population sample (UK Biobank). *Transl Psychiatry*. **11**(1), 21 (2021)
11. Jacobs, G.R., Ameis, S.H., Ji, J.L., et al.: Developmentally divergent sexual dimorphism in the cortico-striatal-thalamic-cortical psychosis risk pathway. *Neuropsychopharmacology*. **44**(9), 1649–1658 (2019)
12. Garavan, H., Bartsch, H., Conway, K., et al.: Recruiting the ABCD sample: Design considerations and procedures. *Dev. Cogn. Neurosci.* **32**, 16–22 (2018)
13. Wang, B., Mezlini, A.M., Demir, F., et al.: Similarity network fusion for aggregating data types on a genomic scale. *Nat. Methods*. **11**(3), 333–337 (2014)
14. Poulton, R., Caspi, A., Moffitt, T.E., Cannon, M., Murray, R., Harrington, H.: Children’s self-reported psychotic symptoms and adult schizophreniform disorder. *Arch. Gen. Psychiatry*. **57**(11), 1053 (2000)
15. Davies, J., Sullivan, S., Zammit, S.: Adverse life outcomes associated with adolescent psychotic experiences and depressive symptoms. *Soc. Psychiatry Psychiatr Epidemiol.* **53**(5), 497–507 (2018)
16. Karcher, N.R., Loewy, R.L., Savill, M., et al.: Persistent and distressing psychotic-like experiences using adolescent brain cognitive developmentSM study data. *Mol. Psychiatry*. **27**(3), 1490–1501 (2022)
17. Karcher, N.R., Loewy, R.L., Savill, M., et al.: Replication of associations with psychotic-like experiences in middle childhood from the Adolescent Brain Cognitive Development (ABCD) study. *Schizophr Bull. Open*. **1**(1), sgaa009 (2020)
18. Karcher, N.R., Barch, D.M., Avenevoli, S., et al.: Assessment of the Prodromal Questionnaire-Brief Child Version for measurement of self-reported psychoticlike experiences in childhood. *JAMA Psychiatry*. **75**(8), 853–861 (2018)
19. Karcher, N.R., Modi, H., Kochunov, P., Gao, S., Barch, D.M.: Regional vulnerability indices in youth with persistent and distressing psychoticlike experiences. *JAMA Netw. Open*. **6**(11), e2343081 (2023)
20. Achenbach, T.M.: Achenbach system of empirically based assessment (ASEBA): Development, findings, theory, and applications. *University of Vermont, Research Center of Children, Youth & Families*. Published online 2009
21. Weintraub, S., Dikmen, S.S., Heaton, R.K., et al.: Cognition assessment using the NIH Toolbox. *Neurology*. **80**(11 Suppl 3), S54–S64 (2013)
22. Spreen, O., Strauss, E.: A compendium of neuropsychological tests: Administration, norms, and commentary. Published online 1998. <http://www.iapsych.com/articles/strauss2006.pdf>
23. Acker, W., Acker, C.: Bexley Maudsley automated processing screening and Bexley Maudsley category sorting test manual. *Windsor, England: NFER-Nelson*. Published online 1982

24. Wechsler, D., Kodama, H.: Wechsler intelligence scale for children. ;1. (1949).
<https://buros.org/sites/unl.edu.cehs.educational-psychology.buros/files/media/file/Ventresco%2C%20N.%20-%20Assignment%20Description.pdf>
25. Hagler, D.J. Jr., Hatton, S., Cornejo, M.D., et al.: Image processing and analysis methods for the Adolescent Brain Cognitive Development Study. *Neuroimage*. **202**(116091), 116091 (2019)
26. Desikan, R.S., Ségonne, F., Fischl, B., et al.: An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage*. **31**(3), 968–980 (2006)
27. Gordon, E.M., Laumann, T.O., Adeyemo, B., Huckins, J.F., Kelley, W.M., Petersen, S.E.: Generation and evaluation of a cortical area parcellation from resting-state correlations. *Cereb. Cortex*. **26**(1), 288–303 (2016)
28. Hagler, D.J. Jr., Ahmadi, M.E., Kuperman, J., et al.: Automated white-matter tractography using a probabilistic diffusion tensor atlas: Application to temporal lobe epilepsy. *Hum. Brain Mapp*. **30**(5), 1535–1547 (2009)
29. Johnson, W.E., Li, C., Rabinovic, A.: Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics*. **8**(1), 118–127 (2007)
30. Fortin, J.P., Cullen, N., Sheline, Y.I., et al.: Harmonization of cortical thickness measurements across scanners and sites. *Neuroimage*. **167**, 104–120 (2018)
31. Velayudhan, P.S., Xu, X., Kallurkar, P., et al.: Metasnf: Meta clustering with similarity network fusion in R. *arXiv [statCO]*. Published online October 23, (2024). <http://arxiv.org/abs/2410.17976>
32. Steinley, D., Brusco, M.J., Hubert, L.: The variance of the adjusted Rand index. *Psychol. Methods*. **21**(2), 261–272 (2016)
33. Hawco, C., Dickie, E.W., Jacobs, G., Daskalakis, Z.J., Voineskos, A.N.: Moving beyond the mean: Subgroups and dimensions of brain activity and cognitive performance across domains. *Neuroimage*. **231**, 117823 (2021)
34. Jacobs, G.R., Voineskos, A.N., Hawco, C., et al.: Integration of brain and behavior measures for identification of data-driven groups cutting across children with ASD, ADHD, or OCD. *Neuropsychopharmacology*. **46**(3), 643–653 (2021)
35. Weller, B.E., Bowen, N.K., Faubert, S.J.: Latent class analysis: A guide to best practice. *J. Black Psychol*. **46**(4), 287–311 (2020)
36. Del Re, E.C., Stone, W.S., Bouix, S., et al.: Baseline cortical thickness reductions in clinical high risk for psychosis: Brain regions associated with conversion to psychosis versus non-conversion as assessed at one-year follow-up in the shanghai-at-risk-for-psychosis (SHARP) study. *Schizophr Bull*. **47**(2), 562–574 (2021)
37. Vissink, C.E., Winter-van Rossum, I., Cannon, T.D., Fusar-Poli, P., Kahn, R.S., Bossong, M.G.: Structural brain volumes of individuals at clinical high risk for psychosis: A meta-analysis. *Biol. Psychiatry Glob Open. Sci*. **2**(2), 147–152 (2022)

38. Chan, R.C.K., Di, X., McAlonan, G.M., Gong, Q.Y.: Brain anatomical abnormalities in high-risk individuals, first-episode, and chronic schizophrenia: an activation likelihood estimation meta-analysis of illness progression. *Schizophr Bull.* **37**(1), 177–188 (2011)
39. Voineskos, A.N., Jacobs, G.R., Ameis, S.H.: Neuroimaging Heterogeneity in Psychosis: Neurobiological Underpinnings and Opportunities for Prognostic and Therapeutic Innovation. *Biol. Psychiatry.* **88**(1), 95–102 (2020)
40. Park, J., Lee, E., Cho, G., et al.: Gene-environment pathways to cognitive intelligence and psychotic-like experiences in children. *Elife.* **12**, RP88117 (2024)
41. Karcher, N.R., Dong, F., Johnson, E.C., et al.: Longitudinal trajectories of cognition and neural metrics as predictors of persistent distressing psychotic-like experiences across middle childhood and early adolescence. *bioRxivorg* Published online January. **22** (2025). 10.1101/2025.01.20.633817
42. Kirschner, M., Hodzic-Santor, B., Antoniades, M., et al.: Cortical and subcortical neuroanatomical signatures of schizotypy in 3004 individuals assessed in a worldwide ENIGMA study. *Mol. Psychiatry.* **27**(2), 1167–1176 (2022)
43. Hassan, A.A., Ali, A.A., Fouda, A.E., Hanafy, R.J., Fouda, M.E.: Leveraging embedding techniques in multimodal machine learning for mental illness assessment. *arXiv [eessAS]*. Published online April 2, 2025. <http://arxiv.org/abs/2504.01767>
44. Jacobson, S., Kelleher, I., Harley, M., et al.: Structural and functional brain correlates of subclinical psychotic symptoms in 11–13 year old schoolchildren. *Neuroimage.* **49**(2), 1875–1885 (2010)
45. Liu, T., Yu, H., Blair, R.H.: Stability estimation for unsupervised clustering: A review. *Wiley Interdiscip. Rev. Comput. Stat.* **14**(6), e1575 (2022)
46. Karcher, N.R.: Psychotic-like experiences in childhood and early adolescence: Clarifying the construct and future directions. *Schizophr Res.* **246**, 205–206 (2022)
47. Nakua, H., Hawco, C., Forde, N.J., et al.: Systematic comparisons of different quality control approaches applied to three large pediatric neuroimaging datasets. *Neuroimage.* **274**(120119), 120119 (2023)
48. Uddin, L.Q., Castellanos, F.X., Menon, V.: Resting state functional brain connectivity in child and adolescent psychiatry: where are we now? *Neuropsychopharmacology.* **50**(1), 196–200 (2024)
49. Wierenga, L.M., Langen, M., Oranje, B., Durston, S.: Unique developmental trajectories of cortical thickness and surface area. *Neuroimage.* **87**, 120–126 (2014)
50. Oliver, L.D., Hawco, C., Viviano, J.D., Voineskos, A.N.: From the Group to the Individual in Schizophrenia Spectrum Disorders: Biomarkers of Social Cognitive Impairments and Therapeutic Translation. *Biol Psychiatry.* Published online September. **20** (2021). 10.1016/j.biopsych.2021.09.007

Figures

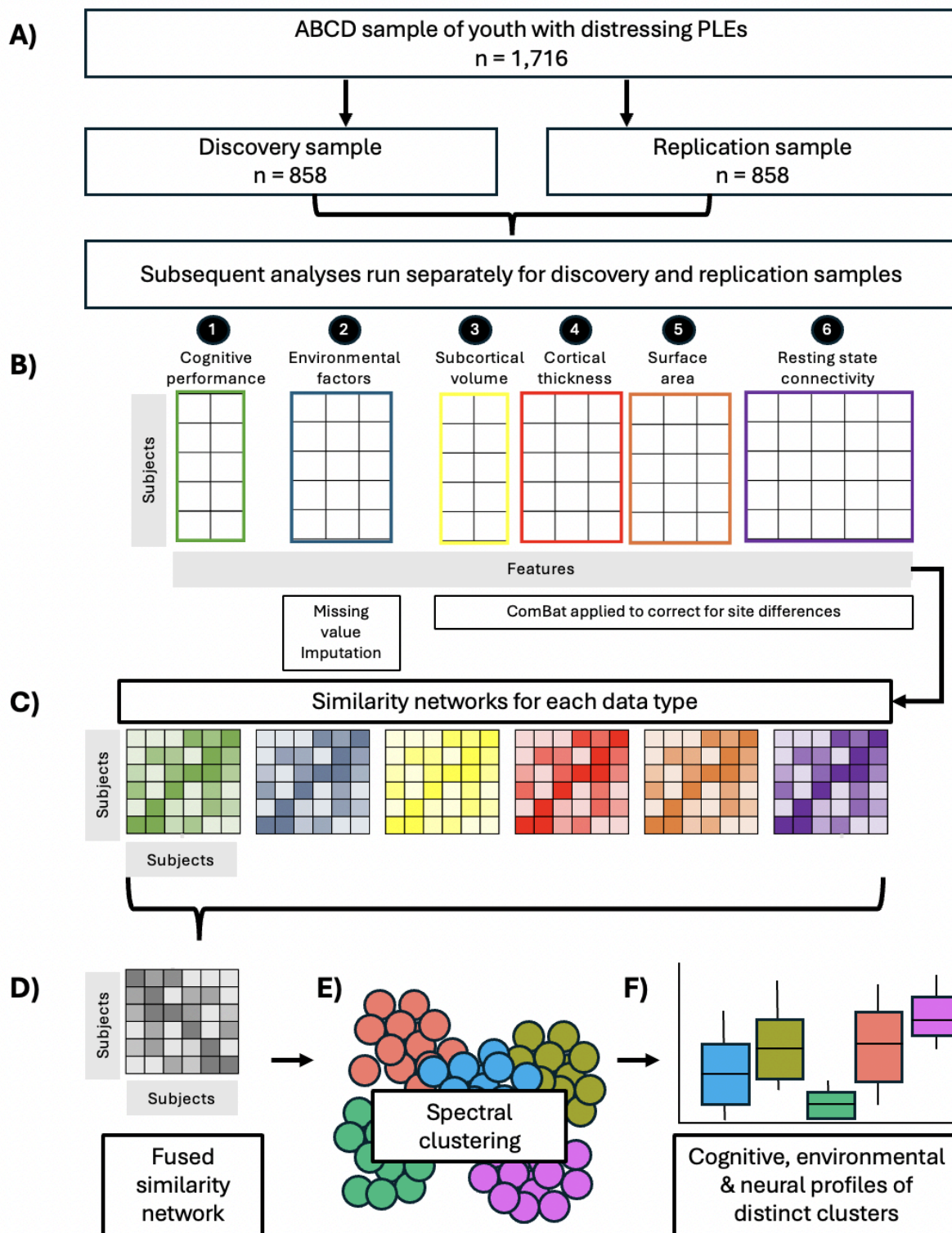


Figure 1

Schematic of multimodal data integration using Similarity Network Fusion (SNF) and clustering analysis. (A) Participants with distressing psychotic-like experiences (PLEs) were split into independent discovery and replication samples (n = 858 each). All subsequent steps were conducted separately in the discovery and replication samples. (B) Multimodal data were explored across six domains: cognitive performance, environmental factors, subcortical volume, cortical thickness, surface area, and resting-

state functional connectivity. Site effects were removed from neuroimaging features via ComBat, and missing values in environmental variables were imputed to ensure completeness. **(C)** For each modality, subject-by-subject similarity networks were constructed, reflecting pairwise similarity based on the respective feature set. **(D)** SNF was then applied to create a single fused similarity network that captures shared and complementary information across data types. **(E)** Spectral clustering was performed on the fused network to identify subgroups of participants with similar multimodal profiles. **(F)** The resulting clusters were characterized based on their distinct cognitive, environmental, and neural patterns.

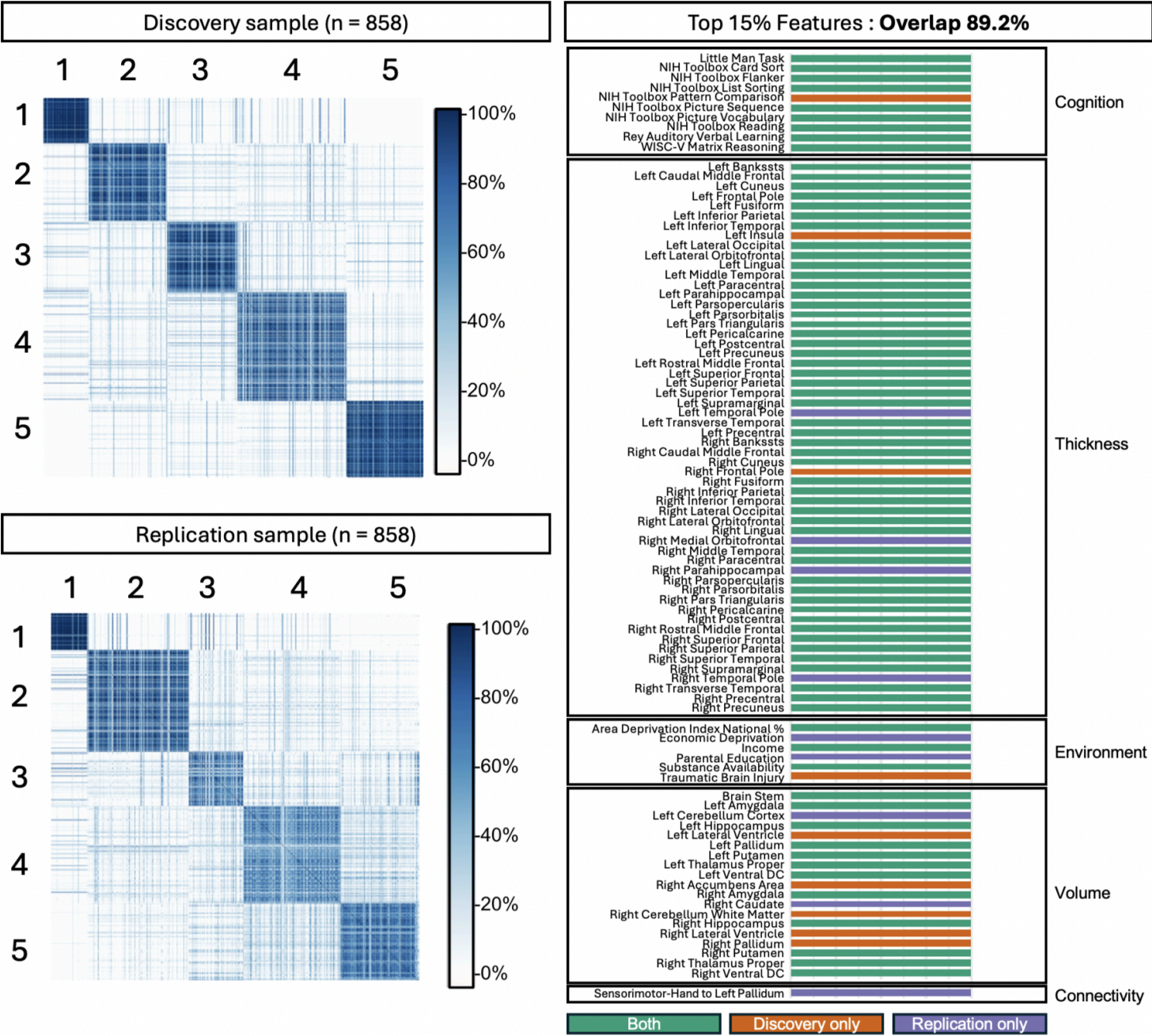


Figure 2

Participant agreement matrices and top contributing features across discovery and replication samples. Left panels show participant agreement matrices for the discovery (top) and replication

(bottom) samples (n=858 each), indicating how frequently each participant clustered with each other participant across 1000 iterations of 80% resampling. Darker shades of blue represent higher agreement. Right panel displays the top 15% of contributing features determined by normalized mutual information (NMI) score in each sample. The 15% threshold was chosen a priori to capture the strongest contributing features while minimizing inclusion of weaker, potentially noisy features. Features are color-coded by whether they appeared in both samples (green), in the discovery sample only (orange), or in the replication sample only (purple). A high overlap of 89.2% was observed between discovery and replication feature sets.

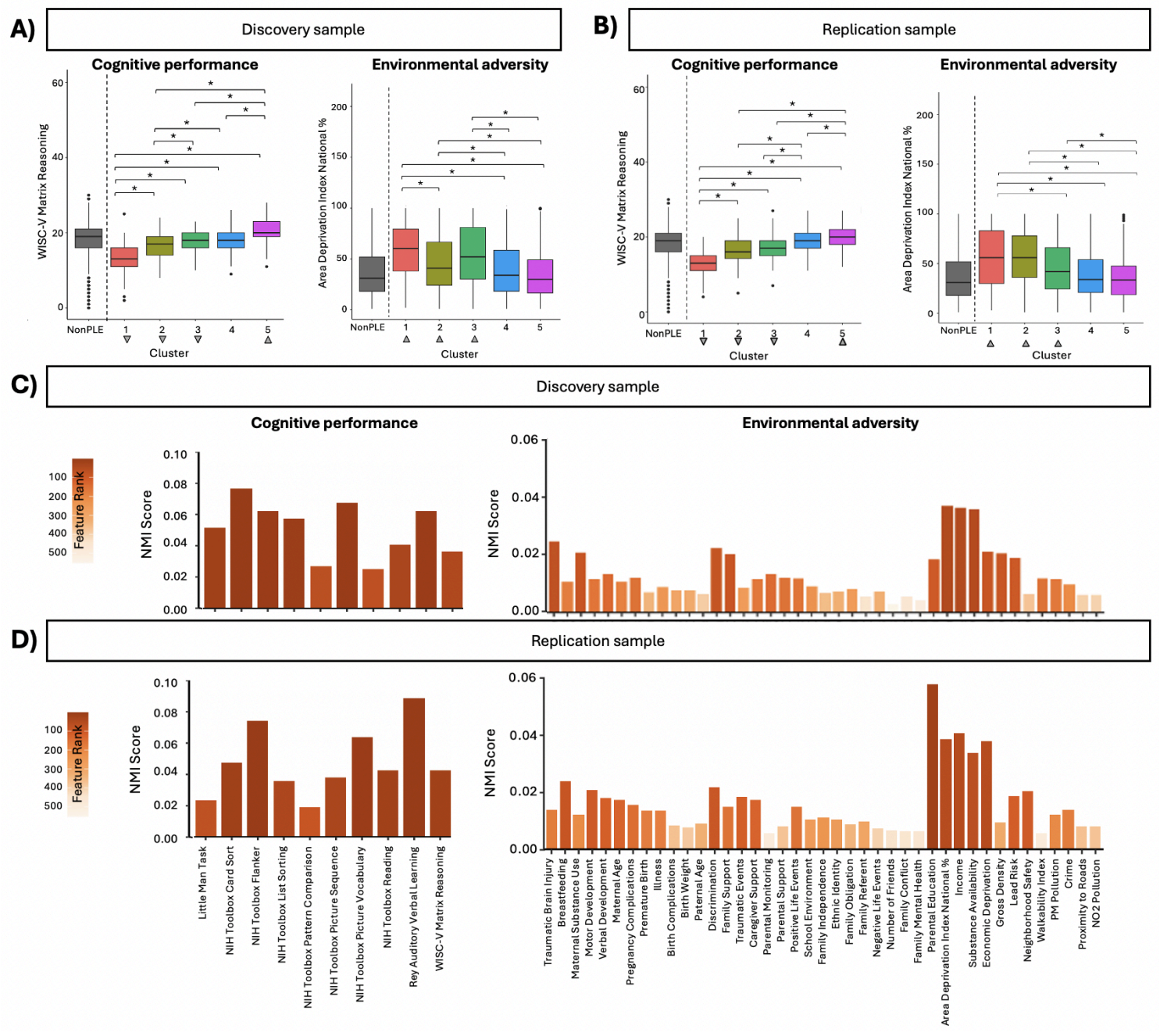


Figure 3

Cognitive performance and environmental adversity across data-driven clusters in discovery and replication samples. Boxplots (top) display one of the top contributing cognitive performance (WISC-V Matrix Reasoning) and environmental adversity (Area Deprivation Index; national percentile) measures. These were chosen to illustrate clear differentiation across clusters. A comparison group of youth with no distressing psychotic-like experiences (NonPLE) is included for visualization purposes only; pairwise comparisons were conducted among the five data-driven clusters. Asterisks denote significant pairwise differences ($p < .05$; *). Exploratory contrasts with the NonPLE group were also conducted; grey arrows underneath cluster numbers denote when means were significantly higher or lower than NonPLE. Results are shown for the discovery (A) and replication (B) samples. Barplots (bottom) display Normalized Mutual Information (NMI) scores for all cognitive features ($n=10$) and environmental measures ($n=41$) used in clustering, with higher scores indicating stronger feature importance. Bars are colored by rank, with darker colors representing higher-ranking features. Results are shown for the discovery (C) and replication (D) samples.

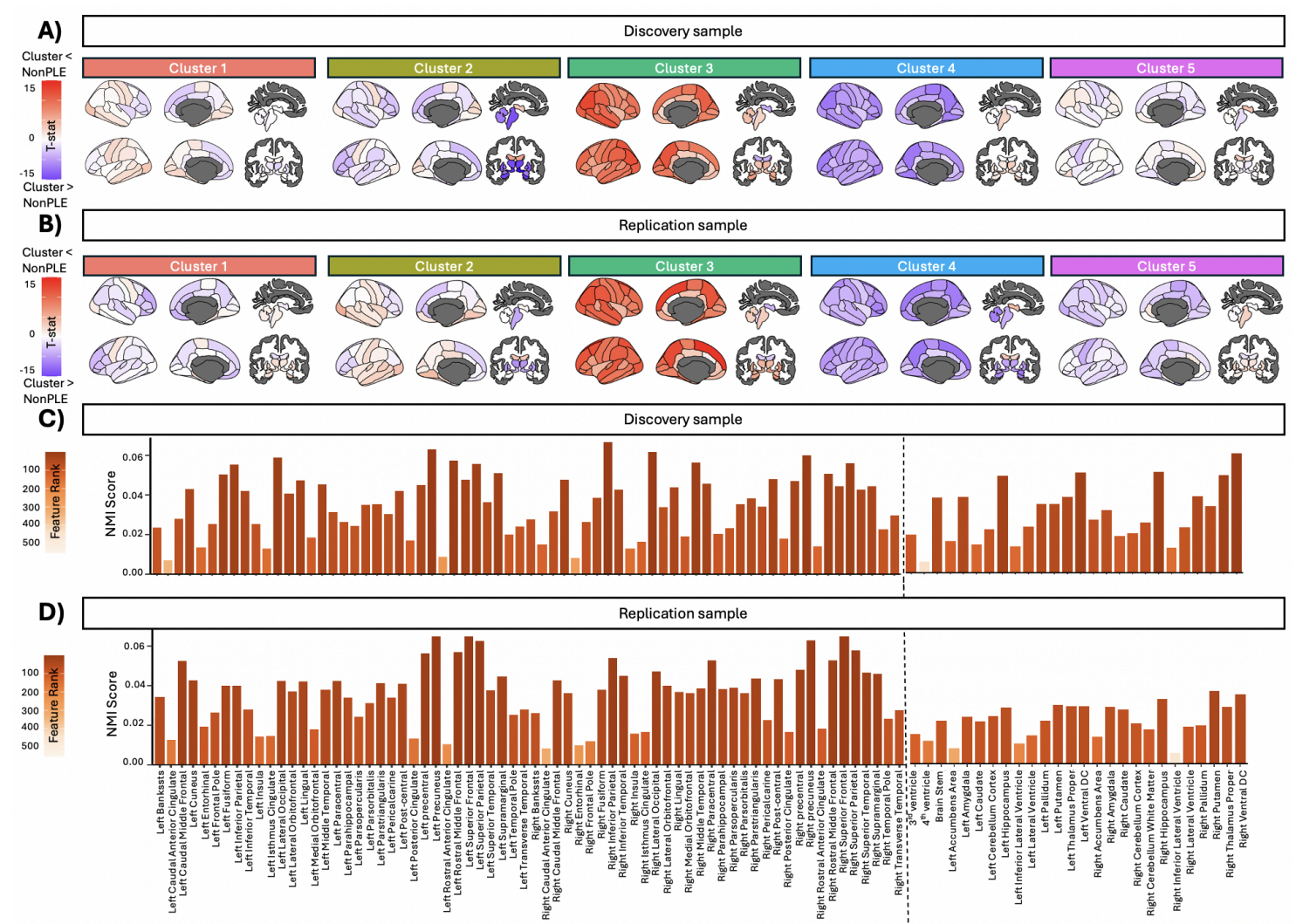


Figure 4

Cortical thickness and subcortical volume differences across data-driven clusters relative to the NonPLE group in discovery and replication samples. Brain maps illustrate cortical thickness and subcortical volume differences (t-statistics) between each cluster and a comparison group of youth with no distressing PLEs (NonPLE). T-statistics are shown for visualization purposes only and are not thresholded for statistical significance. Warmer colors indicate regions where clusters show lower thickness or volume than NonPLE (i.e., Cluster < NonPLE), and cooler colors indicate regions with greater thickness or volume (i.e., Cluster > NonPLE). Barplots display Normalized Mutual Information (NMI) scores for all cortical thickness (n=68) and subcortical volume (n=26) regions used in clustering, with higher scores indicating stronger feature importance. Bars are colored by rank, with darker colors representing higher-ranking features. Results are shown for the discovery (A, C) and replication sample (B, D).

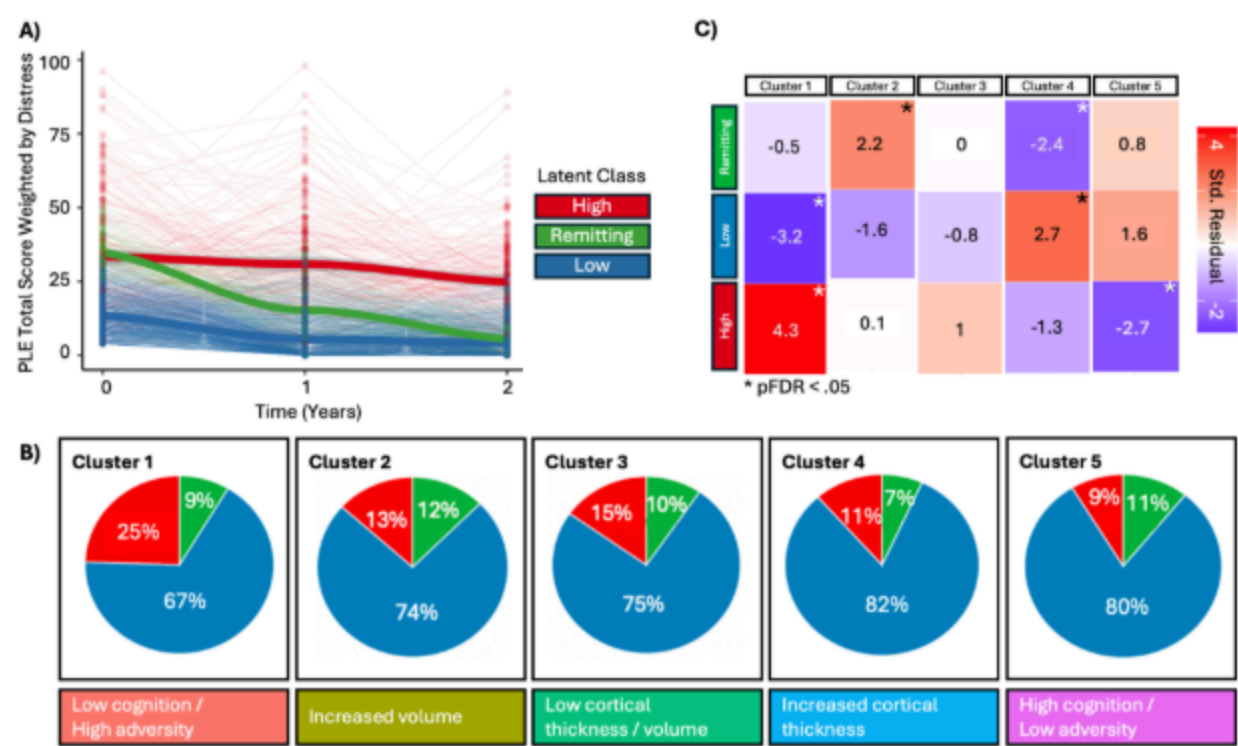


Figure 5

Longitudinal trajectories of psychotic-like experiences (PLEs) and their distribution across data-driven clusters. (A) Growth mixture modeling identified a three-class solution of PLE trajectories (PLE total score weighted by distress) across three time points: baseline, 1-year, and 2-year follow-up (n=1,545). The latent classes include Persistent PLEs (red; 13.2%), Remitting PLEs (green; 9.6%), and Low-level PLEs (blue; 77.2%). Thin lines represent individual trajectories, and bold lines indicate the average trajectory for each class. (B) Pie charts show the proportion of individuals from each latent class within the five data-driven clusters, with summaries below describing the key features of each cluster. (C) Standardized residuals from a chi-square test indicate deviations from expected distributions between clusters and latent classes. Red cells reflect an overrepresentation of a latent class within a cluster,

while blue cells reflect an underrepresentation. Asterisks denote statistically significant differences after FDR correction ($p\text{FDR} < .05$).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementABCDPLEs.docx](#)