Clustering cognitive phenotypes in affective and non-affective psychosis

Katharina M Bracher[1], Afra Wohlschläger[2], Kathrin Koch[2], Franziska Knolle[2]*

[1]) Division of Neurobiology, Faculty of Biology, LMU Munich, Martinsried 82152, Germany; Graduate School of Systemic Neurosciences, LMU Munich, Martinsried 82152, Germany

[2]) Department of Diagnostic and Interventional Neuroradiology, School of Medicine, Technical University of Munich, Munich, Germany

* corresponding author: franziska.knolle@tum.de

Abstract: 255 Main Text: 3854

Conflict of Interest: None of the authors declares a conflict of interest.

Funding: FK received funding from the European Union's Horizon 2020 [Grant number 754462].

Abstract

Cognitive deficits are present in the majority of psychosis patients. They range across various domains, such as working memory, and executive functioning, and are linked to neurobiological changes, including changes in gray matter volume. In this study, we explored data-driven clustering solutions using behavioural, demographic, psychological and structural brain data, but no clinical data, in a cross-diagnostic sample combining affective (N=51) and non-affective (N=111) psychosis patients as well as healthy controls (N=55). The goal of the study was (1) to test classification sensitivity of diagnostic groups using data without diagnostic information, and (2) to explore identification of potential cognitive phenotypes. We used K-means and spectral clustering, investigating two-, three- and four-cluster solutions, and group-membership matching. We then explored cognitive deficits and symptom expression within the four-cluster solution. Our results revealed best group-cluster matching using PCA-pre-selected only non-brain, mostly cognitive, features. The three groups were clustered with medium sensitivity, correctly identifying between 44% and 78% of individuals per group. More importantly, however, clustering using four clusters allowed the identification of cognitive phenotypes, that significantly varied in clinical and cognitive impairment. We identified one cluster expressing the lowest symptom scores and unimpaired cognition, one other cluster with the highest symptom scores, especially negative symptoms, and global cognitive impairments, across all domains, and two intermediate clusters. In conclusion, these results provide evidence for cognitive phenotypes with specific symptom expressions combining individuals with different overall diagnoses. The clear link between the cognitive deficits and symptoms indicates the need for the development of cognition-based interventions.

Contents

1 Introduction **3**

2 Methods **4**

2.1 Participants 4

2.2 Analysis 5

2.2.1 Clustering Strategy and Preprocessing 5

2.2.2 Correlation and Homogeneity 5

2.2.3 K-Means and Spectral Clustering 6

2.3 Group Statistics 7

3 Results **7**

3.1 Behavioural, Psychological and Demographic Data 7

3.2 Brain Data 10

3.3 Brain and Non-Brain Data 13

3.4 Clinical Data 14

3.5 Cluster Exploration with Clinical and Cognitive Scores 15

4 Discussion **18**

References **21**

Supplementary Materials **29**

4.1 Brain Data 34

4.2 Brain Data 48

1 Introduction

Cognitive alterations are core symptoms of psychosis [1–4]. Deficits are wide ranged and have been described in areas of working memory [5, 6], attention [7, 8], reasoning [9, 10], decision making [11, 12], salience processing [13, 14] and problem solving [15, 16] and across all stages of the disease [17–21]. Cognitive deficits furthermore precede the clinical onset of psychosis [22], and predict functional outcome in later stages of the disease [23, 24], impacting employment status, independent living and social functioning [24, 25]. Although cognitive impairments are present in about 80% of patients suffering from psychotic disorders [26–28], there is great variability between patients, and across different diagnoses, e.g., affective vs non-affective psychosis [29–31]. However, many studies often do not differentiate between affective vs non-affective psychosis when investigating cognitive deficits [32]. In a review, Barch and Sheffield [31] summarised that while the severity of cognitive impairment is stronger in non-affective compared to affective psychosis, the relative impairments across different cognitive domains are very similar. Generally, research has shown that fewer cognitive deficits and higher cognitive reserve during prodromal and first episode psychosis are generally, diagnosis-independently, associated with better functioning and recovery [33, 34]. Therefore, early detection and intervention is crucial, to improve long-term outcome.

Cognitive deficits in psychosis have been linked to alterations in the cortico-cerebellar-thalamic-cortical circuits [35]. Here, dysfunctional GABA (gamma-amino-butyric acid) inter-neurons, the main inhibitory neurons of the central nervous system, may disrupt the balance between excitatory and inhibitory processes in the cortex [36]. These accounts confirm an established hypothesis suggesting that neuronal pruning during adolescence [37] through reorganization of both structural and functional brain networks is dysfunctional in schizophrenia, impacting neural communication and potentially causing cognitive impairments in individuals suffering from the disorder [38]. Further support is provided by imaging studies showing reduced gray matter volume and altered network organisation which correlates with cognitive deficits, at illness onset [39, 40], early psychosis [41] and chronic schizophrenia [42]. Cognitive deficits in psychosis have been described and investigated as intermediate phenotypes [43]. In a recent study, Shafee et al [44] pointed out that cognitive phenotypes may vary grossly depending on specific types of psychosis (e.g. affective vs non-affective), suggesting that certain domains of cognition (e.g., working memory vs face processing) may be more etiologically linked to psychosis than others. Using a K-means clustering approach in a cross-diagnostic sample, Lewandowski et al [45] identified four cognitive subgroups combining different psychosis groups. Importantly, they identified one cognitively intact cluster including healthy controls and patients with different diagnoses, while the other three clusters were dominated by different cognitive impairment profiles [45]. A recent systematic review of data-driven identification of cognitive subtypes [46] highlighted that despite the heterogeneity of clustering methods used and cognitive domains studied,

there is some commonality in the identification of a severe cognitive deficit phenotype, which shows deficits across multiple domains. However, it is unclear whether the severe cognitive deficit phenotype also shows the strongest symptom expression.

In the current study, we aimed to explore whether data-driven clustering using behavioural, demographic, psychological and structural brain data, but no clinical data, can be applied successfully (1) to reproduce diagnostic groups, i.e., affective psychosis patients, non-affective psychosis patients and healthy controls; and (2) to explore whether different cluster solutions would yield the identification of cognitive phenotypes. We used data from the "Human Connectome Project for Early Psychosis" which provides a large battery of behavioural tests and scores assessing cognition, such as working memory, attention, or decision making, as well as psychological data, such as general distress, and brain imaging data, namely grey matter volume. We used a data-driven approach comparing clustering of behavioural, demographic and psychological data, and of brain data separately as well as jointly. Importantly, we first applied component analyses to identify those features that explain the most variance, and subsequently used those features for clustering. In this study, we provide answers to the following questions: (1) Using pre-selected, variance-explaining features, not including diagnostic information, is it possible to reproduce group-membership of affective psychosis patients, non-affective psychosis patients and controls with a three cluster solution; or (2) do clusters form subgroups with specific cognitive profiles (i.e., cognitive phenotypes) using a four cluster solution?

2 Methods

2.1 Participants

In this paper, we analysed data collected by the "Human Connectome Project for Early Psychosis" (HCP-EP) [47]. The HCP-EP 1.1 release (August 2021 HCP-EP Release 1.1 on NDA) contains 251 subjects with 57 affective and 126 non-affective psychotic patients, both within the first three years of the onset of psychotic symptoms, and 68 matched healthy controls with comprehensive demographic (e.g., age, education), psychological (e.g., emotional survey, distress scale), and behavioural (e.g., auditory discrimination task, decision making task). Brain imaging data (e.g., structural data, functional task data) is available for 183 of those subjects with 28 in the affective psychosis group and 93 in the non-affective psychosis group and 57 healthy controls. After selection (Supplement, 1.1 Features Selection) analysis was performed on 217 subjects (55 healthy controls, 51 affective psychosis group, 111 non-affective psychosis group, Tab. 1).

Table 1: Group demographics and clinical scores of final sample (N=217).

	Control	Affective	Non-Affective	Group Comparison KW-chi2/P-ch2, chi2(df), p-value
N	55	51	111	
Age mean/std (range)	24.4/4.25 (17-36)	23.8/3.79 (18-34)	22.7/3.23 (18-34)	6.79 (2), 0.034
Gender female	19	30	32	13.68 (2), 0.0012
Race	White (37)	White (36)	White (50), Black or African American (49)	41.63 (12), <0.001
PANSS total	na	44.35 (13.00)	50.28 (14.83)	12.54 (1), <0.001
PANSS positiv	na	10.1 (3.92)	12.38 (4.64)	11,48 (1), <0.001
PANSS negative	na	11.53 (4.67)	14.03 (5.66)	9.17 (1), 0.003
General psycho-pathology	na	23.43 (6.28)	25.51 (5.77)	6.86 (1), 0.009

2.2 Analysis

2.2.1 Clustering Strategy and Preprocessing

The feature selection strategies are described in detail in the supplementary section 1.1 Feature Selection. Clustering on selected features was performed separately on non-brain and brain data (Fig. 1A). The behavioural, psychological and demographic data consisted of continuous or numeric and categorical features, hence discrete data (Supplement, Tab. 3), while brain data were continuous. We therefore split our dataset according to numerical and discrete data types, normalised the data and performed Multiple Correspondence Analysis (MCA) on the categorical data, that maps the categorical data to a continuous space [48], and a Principle Component Analysis (PCA) to the numeric data (Fig. 1B). See supplementary section 1.2 Clustering of Mixed Data Types and section 1.3-1.5 on preprocessing for details.

2.2.2 Correlation and Homogeneity

To describe the association between subjects, we performed a Pearson's correlation between all subjects across all features. We then tested homogeneity between groups, we used a Levene's test for equal variances for between groups. Furthermore, we used a Pearson's correlation to quantify the association between features across all subjects.

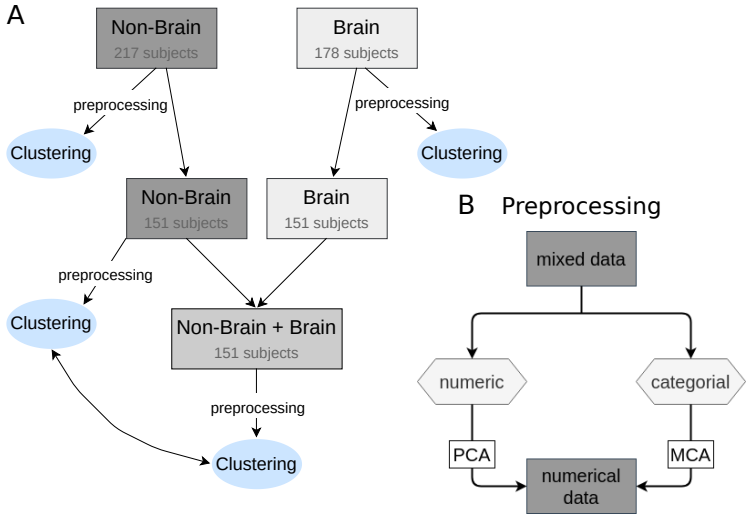

Figure 1: **Workflow and preprocessing of mixed data.** (**A**) Clustering was performed separately on non-brain and brain data. Of 217 subjects in the non-brain data and 178 subjects in the brain data, clustering was performed on the overlap of 151 subjects. To compare clustering on non-brain data and combined non-brain and brain data, clustering was performed on non-brain data with overlapping subjects again. (**B**) Before performing the clustering analysis, non-brain data had to be transformed to be fully continuous. The data was split into numeric and categorical for performing PCA and MCA, respectively. The significant components resulting from these transformations were joint for clustering analysis.

2.2.3 K-Means and Spectral Clustering

We performed the clustering analysis three times: (1) on the behavioural, psychological and demographic data only (N=217), (2) on the brain data only (N=178), and (3) on the combined behavioural, psychological and demographic, and brain data (N=151). To account for the large difference in number of subjects in behavioural psychological and demographic data, and the combined data, we performed the clustering with behavioural data only again with a reduced sample, only containing 151 subjects matching between the brain and non-brain data (Fig. 1, B). As a control analysis, we performed clustering of clinical data (i.e. all items of PANSS) in a two-cluster solution, to investigate the overlap with patient group-membership.

K-Means and spectral clustering on behavioural, psychological and demographic data was performed on joint significant components of numerical and categorical data. For brain data, K-Means was performed on the significant components of the PCA. Spectral clustering on brain data was performed with the first 10 components of the PCA to allow for more flexibility in the nonlinear embedding that is performed in the spectral clustering. For the combined data, both clustering analysis was done on the significant components of behavioural, demographic and brain data. For clinical data, K-Means and spectral clustering

was also performed on the significant components of the PCA.

Performance of the clustering analysis was determined by the ratio of the subjects of one group (healthy controls, affective psychosis group and non-affective psychosis group) in each cluster. The ratio of subjects in cluster j for group i in {control, affective and non-affective} is defined as:

$$\text{ratio}_{i,j} = \frac{\text{number of subjects of group i in cluster j}}{\text{total number of subjects of group i}}$$

To compare ratios between different cluster analyses, e.g. clustering behavioural, psychological, demographic data vs. clustering behavioural, psychological, demographic data combined with brain data, we took the mean of rations over 100 random initialization of the K-Means (spectral) algorithm. This way, we can rule out that possible performance differences were the result of random initialization.

2.3 Group Statistics

In order to test for statistical significance in the differences between the correlations of subjects groups, we used SciPy's T-test for the means of two independent samples of scores, and Levene's test for equal variance to test for differences in variance of correlations. For pairwise comparison between the three subject groups, we corrected the p-value for multiple testing with the Bonferroni method.

To compare features across clusters, we used SciPy's Kruskal-Wallis test for a group comparison and Post hoc pairwise test for pairwise comparisons [49]. We furthermore conducted a partial correlation analysis, correcting for cluster membership, to investigate correlations between cognitive scores and symptom expression across all clusters. We used Pingouin 0.5.2 [50] pairwise correlation with the Holm–Bonferroni method to correct for multiple testing.

3 Results

3.1 Behavioural, Psychological and Demographic Data

Correlations within the control group were stronger than in both the affective (T-test: t_{5624}=23.23, p<0.001) and non-affective (t_{15344}=46.81, p<0.001) group. Non-affective subjects showed higher heterogeneity compared to controls (Levene's test for equal variance: $F_{1,164}$=93.40, p<0.001) or affective subjects ($F_{1,160}$=64.16, p<0.001) (Fig. 2A). There was no significant difference between affective and control group. All p-values were controlled for multiple comparisons.

Correlations of features across subject groups revealed that behavioral features correlated significantly more than psychological features (T-test: t_{1400}=4.55, p<0.001). There was significantly less correlation within features in the affective group compared to control (t_{2900}=5.63, p<0.001) and non-affective group (t_{2900}=5.27, p<0.001) (Supplement, Fig. 1).

We performed K-Means clustering on joint components of the PCA on numerical data (six significant components, Fig. 2 B, blue components) and MCA on categorical data (two significant components). Results for all PCAs and MCAs are presented in the supplementary section Results PCA and MCA. Clustering with only two clusters did not successfully separate patients and controls. Using three clusters, resulted in 87% of controls (Note: Percentages are presented in relation to specific group, and in relation to the individuals of the cluster) classified in cluster 1, however, also patients were assigned to this cluster (34% of affective and 19% of non-affective patients). Performing K-Means clustering with four clusters, yielded a cluster that comprised 83% of control subjects, with fewer patients, specifically 28% of affective and 16% of non-affective patients. K-means clustering revealed that separation of affective and non-affective patients was challenging (Fig. 2E,G upper). However, non-affective patients were more clearly separated from controls compared to affective patients: for three (four) clusters 34% (28%) of affective and 19% (16%) of non-affective patients were situated in cluster 1 with 87% (83%) of controls.

Performing spectral clustering on the preprocessed data allowed a more successful separation of patients and control when clustering with two clusters (Fig. 2F,G lower). Although a separation of affective and non-affective patients was still not possible with spectral clustering, clustering with four clusters identified a subgroup containing 32% of the non-affective patients which has minimal overlap with the affective patients (2%) and controls (2%).

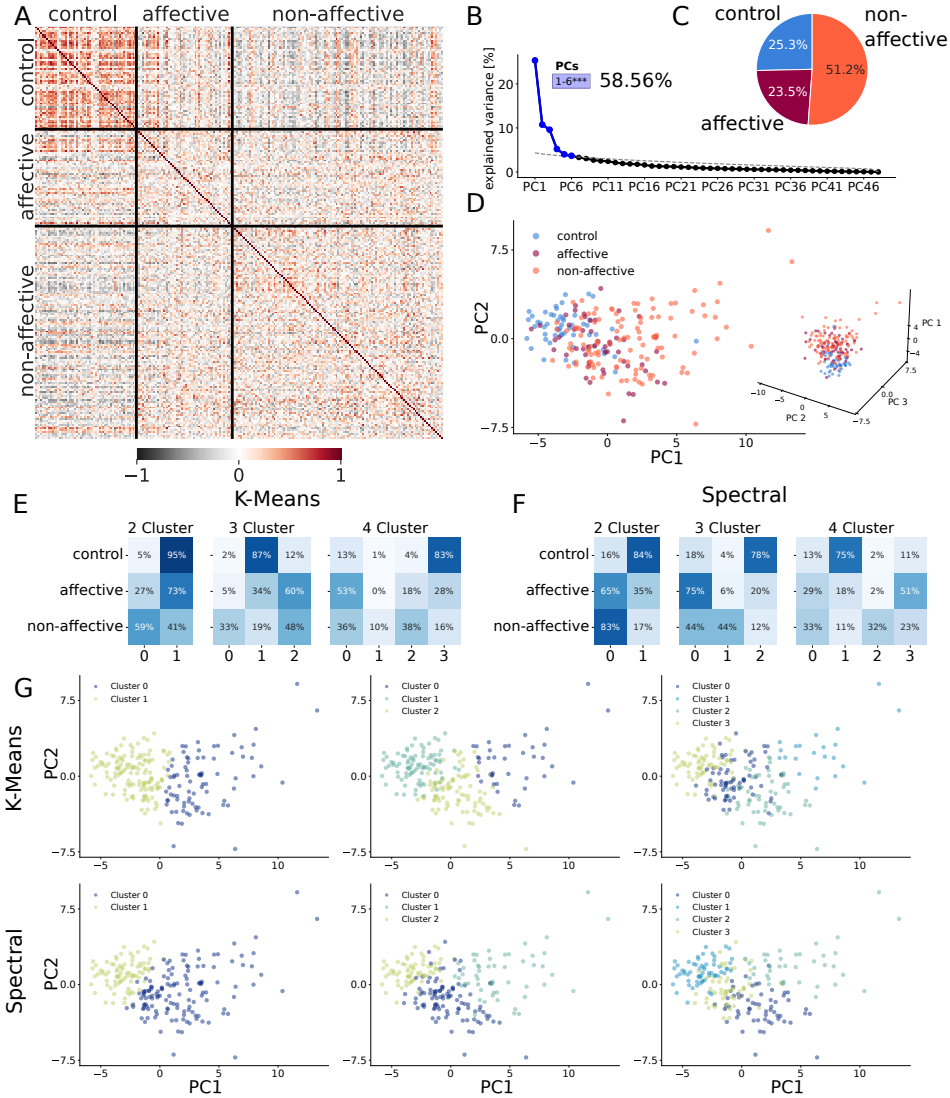

Figure 2: **Clustering of behavioural, psychological and demographic data.** (**A**) Subject-based correlation across numeric features (48). Squares on diagonal show correlation within groups, squares off diagonal show correlations between groups. (**B**) Explained variance by each principal component (PC) in %, with PCA performed on all 48 numeric features across 217 subjects. The first six PCs (blue) survived permutation testing ($p<0.05$, 5000 permutations) and captured 58.6% of total variance. (**C**) Ratio of the three subject groups. Of 217 subjects in total, 25.3% were part of control, 23.5% of affective and 51.2% of non-affective group. (**D**) The first two (left) and first three (right) PCs are used to visualize the data. Data points correspond to subjects and are coloured according to group. (**E**) Percentage of subjects of a group in each cluster, for K-Means clustering analysis with two, three and four clusters on first six PCs from PCA on numeric data and first two components from MCA on categorical data. (**F**) Same as E for spectral clustering. (**G**) Clustering result for K-Means (upper) and spectral (lower) analysis visualized on first two PCs for two, three and four clusters (from left to right).

3.2 Brain Data

We performed subject-based correlations of the raw grey matter volume and examined the variance of subject-based correlations for each group (Fig. 3, brain regions included in morphometric networks are described in supplementary section 4.2 Brain Data).

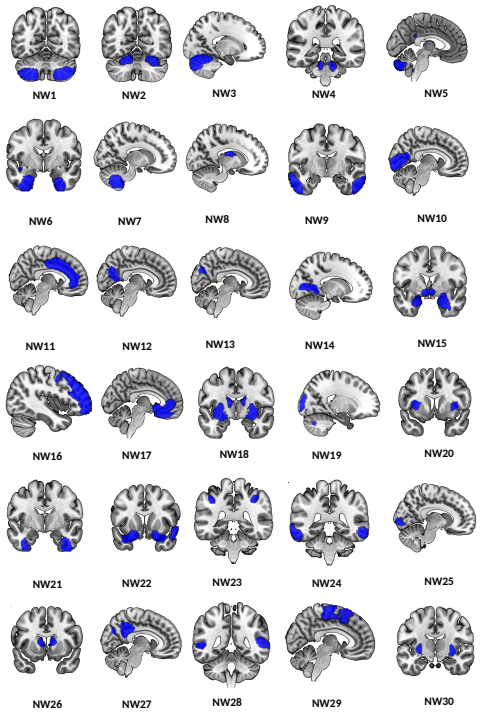

Figure 3: **Morphometric grey matter networks.** This figure shows the 30 anatomically derived morphometric networks from the ICA thresholded at Z>3.5, each network identifies structural grey matter covariance across all subjects.

Fig. 4A indicates that subject groups show a similar homogeneity. Variance between groups did not differ significantly. Mean correlation coefficients were only significantly different between control and affective subjects, with correlations significantly smaller within controls (T-test: t_{11896}=6.88, p<0.001). Feature correlations of grey matter values within each group revealed highest correlation within non-affective compared to controls (T-test: t_{1798}=22.09, p<0.001) and affective (t_{1798}=22.09, p<0.001) and lowest correlations of features within controls (t_{1798}=11.50, p<0.001) (Supplement, Fig. 3).

The PCA analysis (supplementary section 2 Results) did not allow a separation of the three groups, which is also reflected in the clustering analysis. K-Means clustering was conducted on the significant principal components (four) and spectral clustering was con-

ducted on the first 10 principal components (see Methods). Neither K-Means clustering nor spectral clustering were able to separate the groups based on the brain data (Fig. 4E-G).

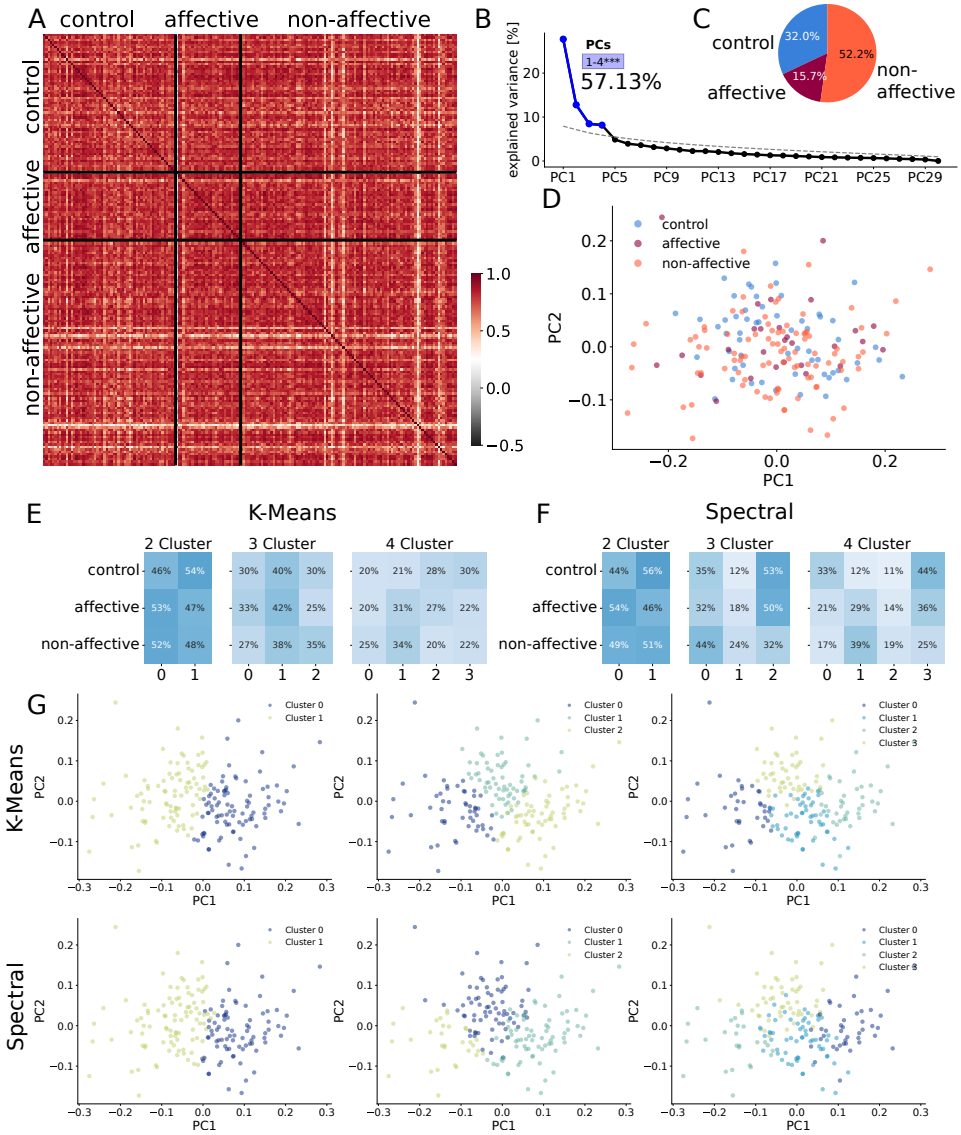

Figure 4: **Clustering subject based on brain data.** (**A**) Correlation between subjects on raw grey matter values. Correlations within groups are displayed on diagonal squares and correlations between groups on off-diagonal squares. (**B**) Variance explained by each of the principal components (PCs) in % of a PCA performed on all 30 corrected gray matter volumes across 178 subjects. The first four PCs (blue) survived permutation testing ($p<0.05$, 5000 permutations). Significant components captured 57.1% of all variance. (**C**) Ratio of the three subject groups. Of 178 subjects in total, 32.0% were part of control, 15.7% of affective and 52.2% of non-affective group. (**D**) The first two PCs are visualized and colored according to groups. (**E**) Percentage of subjects of a group in each cluster, for K-Means clustering analysis with two, three and four clusters on significant PCs. (**F**) Same as E for spectral clustering on first ten PCs, which yielded better results than four. (**G**) Clustering result for K-Means (upper) and spectral (lower) analysis visualized on first two PCs for two, three and four clusters.

3.3 Brain and Non-Brain Data

Finally, we investigated whether the clustering performance would improve when combining behavioural, psychological, and demographic data with brain data. Grey matter values correlated with cognitive features more strongly compared to psychological features (T-test: t_{1558}=10.41, p<0.001) (Fig. 5A).

Non-brain data with fewer subjects still showed no significant difference in mean or variance. We studied the subject-based correlations of combined data. Subject correlations of combined data were significantly higher compared to non-brain data only (T-test: t_{45600}=62.18, p<0.001). The correlations are highest within controls compared to affective (T-test: t_{2690}=18.55, p<0.001) and non-affective (t_{8675}=26.03, p<0.001) and lowest in the non-affective group (t_{7135}=6.29, p<0.001) compared to the affective psychosis group (Fig. 5B). The highest inhomogeneity was again found within non-affective subjects, compared to controls (Levene's test for equal variance: $F_{1,125}$=125.97, p<0.001) and affective subjects $F_{1,103}$=58.76, p<0.001).

Because of fewer subjects in the combined data set, we performed the clustering analysis on the behavioural, psychological and demographic data with matching subjects again, to compare the clustering performance based on added features only (Supplement, Fig. 6). Clustering on combined data with K-Means or Spectral clustering showed no improvement of performance (Fig. 5D).

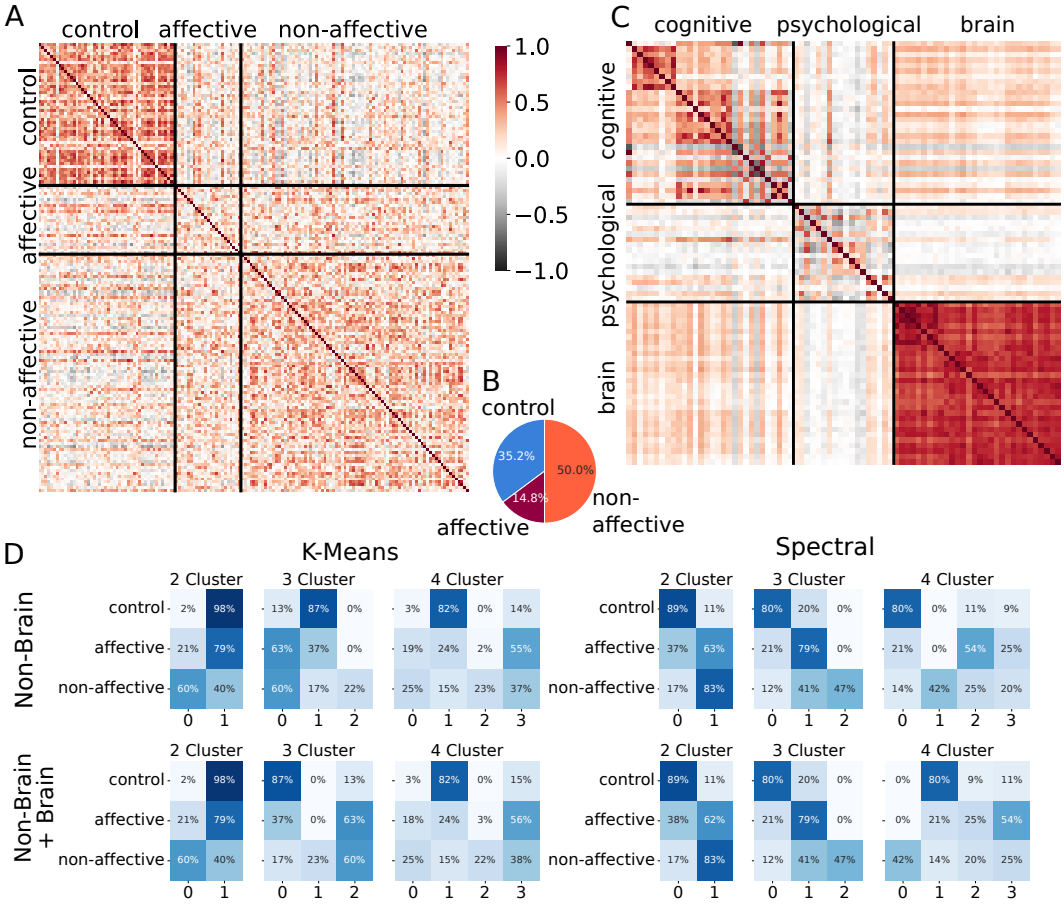

Figure 5: **Clustering on combined brain and non-brain data.** (**A**) Correlations between subjects on combined numerical features of non-brain and raw brain data. Correlations within groups displayed on squares on diagonal, and correlations between groups displayed on off-diagonal squares. (**B**) Ratio of the three subject groups. Of 151 subjects in total, 35.2% were part of control group, 14.8% of affective and 50.0% of non-affective group. (**C**) Correlation between 48 numerical non-brain data, and 30 raw grey matter values across all subjects. Squares on diagonal show correlations within a feature group, squares off-diagonal show correlations between feature groups. (**D**) Plots show results of clustering non-brain data and combined data with matched subjects. Results of K-Means clustering are displayed on the left and results of spectral clustering on the right.

3.4 Clinical Data

We identified three significant principal components in numerical features, that captured 47.9% of total variance (permutation test $p<0.05$, 5000 permutations). We performed K-Means clustering on three significant PCA components. Clustering with two clusters separated affective and non-affective patients with medium sensitivity. One cluster consisted of 63% of non-affective patients, while the other cluster consisted of 78% of the affective

psychosis patients (Supplements, Fig 8 D, E). Spectral clustering on significant components did not result in better separation of patients, one cluster comprising 76% of affective and the other one 63% of non-affective patients 10).

3.5 Cluster Exploration with Clinical and Cognitive Scores

To identify possible cognitive subtypes represented in the clusters, we explored clinical and cognitive scores (Tab. 2) for the four clusters of the spectral clustering analysis for the complete data set of 217 subjects for non-brain data (see Fig. 2G lower right). For clinical scores, only patient data were included. See supplementary section 2.5 Cluster Exploration for further detail.

We found clear differences in PANSS scores comparing clusters. Patients in cluster 1, which included the majority of controls, 75%, (N non-affective: 12, N affective: 9, N controls: 41), showed reduced clinical scores in all domains. Cluster 1 was chosen as the clinical reference cluster, revealing lowest overall clinical scores. Cluster 3, which included the same number of affective and non-affective patients (N non-affective: 26, N affective: 26, N controls: 6) showed the greatest distributions in clinical scores, but only showed significantly higher scores in the general psychopathology compared to cluster 1. Patients in cluster 2, which included mainly non-affective subjects (N non-affective: 36, N affective: 1, N controls: 1), showed significantly higher clinical scores in all scores but PANSS positive. However, cluster 2 showed especially high negative symptoms scores. Patients in cluster 0 (N non-affective: 37, N affective: 15, N controls: 7) showed increased scores in all except the negative symptom scores and furthermore showed particularly high scores in the positive scores (Fig. 6C, see Tab. 2 for group comparison).

Cognitive scores were compared across patients in specific clusters and control subjects, see Table 2 for statistical comparisons and for boxplots and description see Supplement (Fig. 11).

We furthermore conducted a partial correlation analysis, correcting for cluster membership, to investigate correlations between cognitive scores and symptom expression across all clusters (Fig. 6B). The results revealed negative correlations between negative symptoms and crystallized cognition, impulsive decision making, oral reading recognition, picture vocabulary test, total IQ, verbal IQ and auditory attention; negative correlations between general psychopathology and fluid cognition, crystallized cognition, executive functioning, selective attention and oral reading recognition; and finally, negative correlations between positive symptoms and fluid cognition and executive functioning.

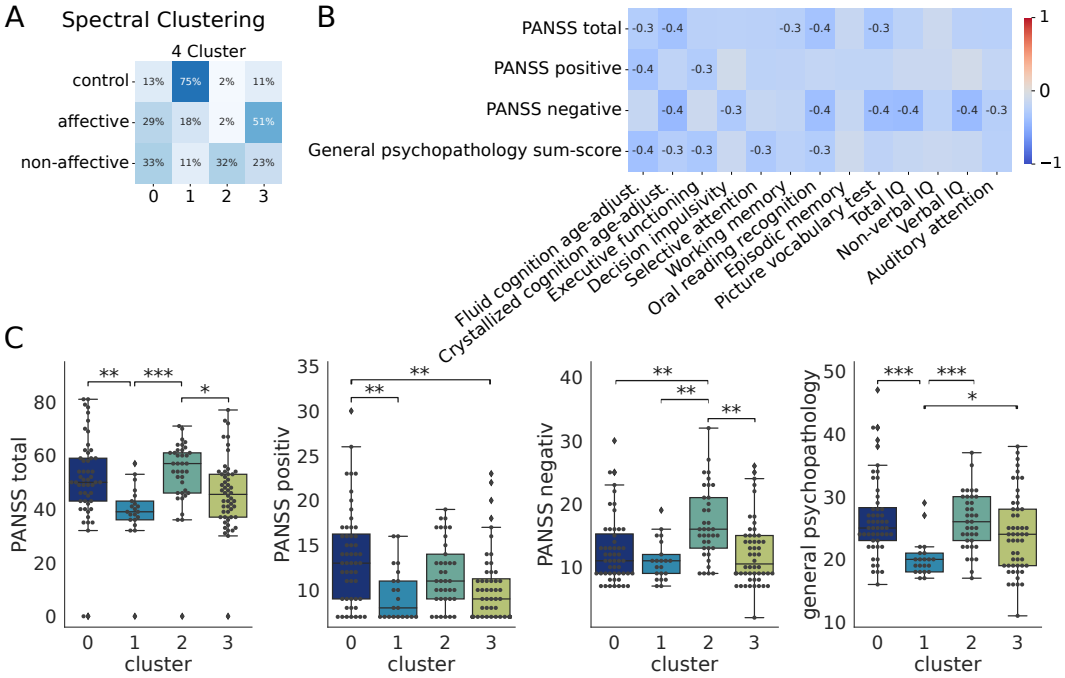

Figure 6: **Exploration of clinical and cognitive features of patients clustered based on behavioural data.** (**A**) Percentage of group in each cluster, resulting from spectral clustering of behavioural data of all subjects with four clusters, as described in Fig. 2F). (**B**) Correlation of clinical and cognitive scores across all patients (affective and non-affective) controlled for assigned clusters. Only significant correlation scores are indicated. (**C**) Clinical scores of patients (affective and non-affective) displayed in boxplots for each cluster.

Table 2: Cognitive and clinical scores of patients in clusters of non-brain data compared to cognitive scores of all controls. The table contains the Kruskal-Wallis H statistic and p-values, corrected for multiple comparisons.

Features	Group comparison	Cluster 0 vs 1	Cluster 0 vs 2	Cluster 0 vs 3	Cluster 0 vs C	Cluster 1 vs 2	Cluster 1 vs 3	Cluster 1 vs C	Cluster 2 vs 3	cluster 2 vs C	Cluster 3 vs C
Cognitive											
Fluid cognition age-adjusted	79.99	13.23	28.03	ns	27.37	28.57	7.15	ns	34.1	47.9	14.81
	<0.001	0.004	<0.001		<0.001	<0.001	0.041		<0.001	<0.001	0.002
Crystallized cognition age-adjusted	67.08	ns	28.62	9.01	9.24	29.88	ns	ns	47.01	43.29	ns
	<0.001		<0.001	0.014	0.009	<0.001			<0.001	<0.001	
Executive functioning	51.34	ns	24.0	ns	7.79	18.75	ns	ns	24.97	41.06	7.63
	<0.001		<0.001		0.009	<0.001			<0.001	<0.001	0.011
Decision impulsivity (40000)	99.69	37.13	7.9	71.01	50.1	9.6	ns	ns	27.36	11.97	9.64
	<0.001	<0.001	0.005	<0.001	<0.001	0.004			<0.001	0.001	0.01
Selective attention, executive control	55.49	5.34	14.79	ns	17.83	19.21	ns	ns	22.72	40.78	11.77
	<0.001	0.024	<0.001		<0.001	<0.001			<0.001	<0.001	0.002
Working memory	61.76	ns	29.73	ns	6.69	25.25	ns	ns	40.04	47.45	ns
	<0.001		<0.001		0.02	<0.001			<0.001	<0.001	
Oral reading recognition	72.5	ns	26.81	10.25	10.03	30.17	ns	ns	50.31	49.38	ns
	<0.001		<0.001	0.002	0.003	<0.001			<0.001	<0.001	
Episodic memory	58.63	5.43	11.62	ns	25.53	17.32	ns	ns	17.24	42.58	14.38
	<0.001	0.02	0.006		<0.001	<0.001			<0.001	<0.001	<0.001
Picture vocabulary test	69.07	ns	28.82	9.46	8.34	28.3	ns	ns	49.28	45.47	ns
	<0.001		<0.001	0.004	0.007	<0.001			<0.001	<0.001	
Total IQ	78.84	ns	36.81	5.55	11.32	28.97	ns	ns	51.71	56.31	ns
	<0.001		<0.001	0.03	0.002	<0.001			<0.001	<0.001	
Non-verbal IQ	55.34	ns	19.07	ns	15.87	16.19	ns	ns	29.76	40.42	5.11
	<0.001		<0.001		<0.001	<0.001			<0.001	<0.001	0.05
Verbal IQ	66.07	ns	27.28	6.83	4.33	25.81	ns	ns	48.44	48.48	ns
	<0.001		<0.001	0.01	0.044	<0.001			<0.001	<0.001	
Auditory attention	87.75	9.21	36.13	ns	22.89	36.32	ns	ns	40.08	58.08	12.87
	<0.001	0.011	<0.001		<0.001	<0.001			<0.001	<0.001	0.001
Clinical											
PANSS total	26.94	13.79	ns	ns	-	20.56	ns	-	11.19	-	-
	<0.001	0.005				<0.001			0.011		
PANSS positive	22.58	11.72	ns	14.12	-	ns	ns	-	ns	-	-
	<0.001	0.008		0.004							
PANSS negative	24.3	ns	14.07	ns	-	16.59	ns	-	15.75	-	-
	<0.001		0.005			0.004			0.002		
General psychopathology sum-score	24.64	20.38	ns	ns	-	21.32	6.25	-	ns	-	-
	<0.001	<0.001				0.001	0.03				

4 Discussion

This study investigated data-driven clustering in a cross-diagnostic sample combining affective and non-affective psychosis patients as well as healthy controls using behavioural, demographic, psychological and structural brain data, but no information regarding diagnostics. We found that best group-cluster matching was achieved using PCA-pre-selected, mostly behavioural features (i.e., features that describe cognitive functioning and cognition in general) with non-brain data spectral clustering. The results showed that the three groups (i.e., affective psychosis, non-affective psychosis, controls) were clustered according to their group membership with medium sensitivity (i.e., correct identification of group membership), correctly identifying between 44% and 78% of individuals. More importantly, however, clustering using four clusters allowed the identification of cognitive subtypes, which significantly varied in clinical and cognitive impairment. With one cluster expressing the lowest symptom scores across all four categories (i.e., PANSS total, PANSS positive, PANSS negative and general psychopathology) and unimpaired cognition (i.e., similar to controls), and one other cluster with the highest symptom scores, especially negative symptoms, and global cognitive impairments, across all domains.

In this cognitively intact cluster, 13% (21/162) of the total number of patients were included. Interestingly, this cluster held only 10.8% of all non-affective psychosis (12/111) patients, but 17.7% of all affective psychosis patients (9/51). Given the link between reduced overall symptoms and maintained cognition, this might not be surprising, as non-affective psychosis patients in this study showed stronger overall symptoms. Using only cognitive data Lewandowski and colleagues [45] found a four-cluster solution to provide the best fit to their data, with one globally impaired cluster, for which cognitive deficits were associated with symptom severity and poorer functioning [45]. Moreover, in our data, we found one cluster (cluster 2) for which patients showed global cognitive impairment, for all scores except decision impulsivity. This cluster also showed the highest values in the negative symptoms (i.e. negative PANSS), and general psychopathology. Interestingly, for positive symptoms, we did not find a significant difference between cluster 2 and cluster 1, the cognitively intact reference cluster, indicating that the cognitive deficits occur in those subjects with strong negative symptoms [51], but importantly independent of diagnosis (i.e., affective vs. non-affective psychosis). The two remaining clusters showed mixed cognitive and symptom profiles. Cluster 0 was characterized by significant cognitive impairment across all cognitive scores compared to controls, however, patients in this cluster performed significantly better in all categories except for impulsive decision making compared to those in cluster 2. Furthermore, cluster 0 patients performed worse than those in the cognitively preserved cluster 1 in measures of fluid cognition, decision impulsivity, episodic memory and auditory attention. Importantly, however, patients in cluster 0 showed significantly enhanced positive symptom scores compared to cluster 1 and 3. Finally, patients of cluster 3

were similar to the cognitively intact cluster 1, but were significantly different from controls across several domains, especially in fluid cognition, episodic memory, selective attention and auditory attention. Patients of cluster 3 showed similar symptom expressions to the cognitively preserved cluster 1, except for increased general psychopathology. The overall structure of cognitive clusters identified in this study supports findings discussed in a recent meta analysis of data driven identification of cognitive phenotypes in schizophrenia [46]. Characteristic to all cluster solutions is the presence of a cognitively spared, one or multiple intermediate and a deficit subgroup [46]. Our work extends this research by showing (1) that these cognitive phenotypes exist independent from the primary diagnosis (i.e., affective or non-affective psychosis); (2) that different clustering methods (i.e., k-means and spectral clustering) yielded highly similar cluster representations with specific symptoms profiles, also in the intermediate phenotypes. The inclusion of other data types, such as grey matter data, which is also associated with cognitive deficits, did not improve the clustering (see discussion below), confirming that the heterogeneity of cognitive profiles can be reliably captured with cognitive task data. Our results indicate that patient classification and diagnosis solely based on International Statistical Classification of Diseases (ICD)/Diagnostic and Statistical Manual of Mental Disorders (DSM) [52, 53] may provide a description of patients which is too general, potentially ignoring options for interventions and treatment development, such as cognitive remediation.

Our correlation analysis revealed that stronger cognitive deficits are linked to stronger negative symptoms, and symptoms or general psychopathology. Although it has been suggested that negative and cognitive symptoms are separable domains which require independent treatment development [54], many studies show an association between cognitive deficits and negative symptoms across all stages of the disease [27, 55–59]. Interestingly, our correlation analysis also revealed specific links between positive symptoms and cognitive deficits, specifically between fluid cognition and executive functioning. Other studies did not find associations between positive symptoms and cognitive deficits [59, 60], although a common mechanism within the dopaminergic system has been discussed [61]. A recent study by Horwood and colleagues [62] showed that lower IQ was linked to positive psychotic symptoms in adolescents, where hallucinations and delusions were only present in those with significantly reduced IQ. Guillem and colleagues [63] specifically investigated the relationship between positive symptoms and deficits in executive functioning. Their findings showed that increased inference sensitivity was associated with hallucinations potentially revealing deficits in thought monitoring, and that disorganization and lack of inhibition was linked to delusional thinking. These studies in combination with our findings may further emphasize the need for interventions targeted at cognitive deficits, as they in turn may improve positive and negative symptoms, as indicated in single studies on cognitive remediation therapy [64, 65].

Furthermore, our study shows that clustering on behavioural data alone resulted in good separation of patients and controls, as well as subtypes. Incorporating structural brain data into the clustering analysis of behavioural, psychological and demographic data did not improve the performance significantly. This is interesting as alterations in grey matter volume have been associated with an increased risk for psychosis and disease development [66–68], and provided the basis for good classification in a recent multicohort-study [69], as well as in earlier studies [70, 71], although classification results are inconsistent [72] and these previous studies contained different patient groups. Future research should investigate whether functional brain data, either resting state or task, or white matter connectivity data may increase the specificity of the clustering. Additionally, further improvement of clustering and identification of cognitive phenotypes may be achieved through the selection of specific behavioural tasks and cognitive domains. Moreover, advanced analysis strategies, e.g., computational modelling may improve clustering [73, 74], as it provides the opportunity to identify and mathematically differentiate behaviourally parameters, which were found to be reliable and unique across individuals [75].

In conclusion, these results provide evidence for the presence of cognitive phenotypes with specific symptom expressions combining individuals with different overall diagnoses, potentially overcoming traditional diagnoses. Considering the clear link between the cognitive deficits and symptom expression, this study is drawing attention to the need for the development of cognition based treatment and interventions that goes beyond traditional diagnoses classifications.

References

[1] P. Fusar-Poli, G. Deste, R. Smieskova, S. Barlati, A. R. Yung, O. Howes, R.-D. Stieglitz, A. Vita, P. McGuire, and S. Borgwardt. "Cognitive Functioning in Prodromal Psychosis: A Meta-analysis". *Archives of General Psychiatry* 69.6 (June 2012), pp. 562–571. ISSN: 0003-990X. DOI: `10.1001/archgenpsychiatry.2011.1592`.

[2] D. M. Barch and A. Ceaser. "Cognition in schizophrenia: core psychological and neural mechanisms". *Trends in cognitive sciences* 16.1 (2012), pp. 27–34.

[3] S. J. Dienel and D. A. Lewis. "Alterations in cortical interneurons and cognitive function in schizophrenia". *Neurobiology of Disease* 131 (2019), p. 104208. ISSN: 0969-9961. DOI: `https://doi.org/10.1016/j.nbd.2018.06.020`.

[4] M. F. Green and P. D. Harvey. "Cognition in schizophrenia: Past, present, and future". *Schizophrenia Research: Cognition* 1.1 (2014), e1–e9. ISSN: 2215-0013.

[5] A. Zanello, L. Curtis, M. B. Bâ, and M. C. Merlo. "Working memory impairments in first-episode psychosis and chronic schizophrenia". *Psychiatry research* 165.1-2 (2009), pp. 10–18.

[6] S. J. Wood, C. Pantelis, T. Proffitt, L. Phillips, G. Stuart, J.-A. Buchanan, K. Mahony, W. Brewer, D. Smith, and P. McGorry. "Spatial working memory ability is a marker of risk-for-psychosis". *Psychological medicine* 33.7 (2003), pp. 1239–1247.

[7] B. A. Cornblatt and A. K. Malhotra. "Impaired attention as an endophenotype for molecular genetic studies of schizophrenia". *American journal of medical genetics* 105.1 (2001), pp. 11–15.

[8] I. Aase, K. Kompus, J. Gisselgård, I. Joa, J. O. Johannessen, and K. Brønnick. "Language lateralization and auditory attention impairment in young adults at ultra-high risk for psychosis: A dichotic listening study". *Frontiers in psychology* 9 (2018), p. 608.

[9] P. Garety, E. Joyce, S. Jolley, R. Emsley, H. Waller, E. Kuipers, P. Bebbington, D. Fowler, G. Dunn, and D. Freeman. "Neuropsychological functioning and jumping to conclusions in delusions". *Schizophrenia Research* 150.2-3 (2013), pp. 570–574. ISSN: 09209964. DOI: `10.1016/j.schres.2013.08.035`.

[10] A. O. Ermakova, N. Gileadi, F. Knolle, A. Justicia, R. Anderson, P. C. Fletcher, M. Moutoussis, and G. K. Murray. "Cost evaluation during decision-making in patients at early stages of psychosis". *Computational Psychiatry* 3 (2019), pp. 18–39. ISSN: 2379-6227.

[11] P. Sterzer, M. Voss, F. Schlagenhauf, and A. Heinz. "Decision-making in schizophrenia: A predictive-coding perspective". *NeuroImage* 190 (2019), pp. 133–143. ISSN: 1053-8119.

[12] A. O. Ermakova, F. Knolle, A. Justicia, E. T. Bullmore, P. B. Jones, T. W. Robbins, P. C. Fletcher, and G. K. Murray. "Abnormal reward prediction-error signalling in antipsychotic naive individuals with first-episode psychosis or clinical risk for psychosis". *Neuropsychopharmacology* 43.8 (2018), pp. 1691–1699.

[13] F. Knolle, A. O. Ermakova, A. Justicia, P. C. Fletcher, N. Bunzeck, E. Düzel, and G. K. Murray. "Brain responses to different types of salience in antipsychotic naıve first episode psychosis: An fMRI study". *Translational psychiatry* 8.1 (2018), pp. 1–13.

[14] J. P. Kesby, G. K. Murray, and F. Knolle. "Neural circuitry of salience and reward processing in psychosis". *Biological Psychiatry Global Open Science* (2021). ISSN: 2667-1743.

[15] R. Morris, T. Rushe, P. Woodruffe, and R. Murray. "Problem solving in schizophrenia: A specific deficit in planning ability". *Schizophrenia Research* 14.3 (1995), pp. 235–246. ISSN: 0920-9964. DOI: `https://doi.org/10.1016/0920-9964(94)00044-9`.

[16] N. Revheim, I. Schechter, D. Kim, G. Silipo, B. Allingham, P. Butler, and D. C. Javitt. "Neurocognitive and symptom correlates of daily problem-solving skills in schizophrenia". *Schizophrenia research* 83.2-3 (2006), pp. 237–245.

[17] Y. Liu, G. Wang, H. Jin, H. Lyu, Y. Liu, W. Guo, C. Shi, J. Meyers, J. Wang, J. Zhao, et al. "Cognitive deficits in subjects at risk for psychosis, first-episode and chronic schizophrenia patients". *Psychiatry Research* 274 (2019), pp. 235–242.

[18] J. M. Sheffield, N. R. Karcher, and D. M. Barch. "Cognitive deficits in psychotic disorders: a lifespan perspective". *Neuropsychology review* 28.4 (2018), pp. 509–533.

[19] J. M. Gold and P. D. Harvey. "Cognitive deficits in schizophrenia". *Psychiatric Clinics of North America* 16.2 (1993), pp. 295–312.

[20] M. Fioravanti, O. Carlone, B. Vitale, M. E. Cinti, and L. Clare. "A meta-analysis of cognitive deficits in adults with a diagnosis of schizophrenia". *Neuropsychology review* 15.2 (2005), pp. 73–95.

[21] S. Kelly, S. Guimond, A. Lyall, W. S. Stone, M. E. Shenton, M. Keshavan, and L. J. Seidman. "Neural correlates of cognitive deficits across developmental phases of schizophrenia". *Neurobiology of disease* 131 (2019), p. 104353.

[22] M. F. Green. "Impact of cognitive and social cognitive impairment on functional outcomes in patients with schizophrenia". *The Journal of clinical psychiatry* 77.suppl 2 (2016), p. 8569. ISSN: 0160-6689.

[23] G. Fervaha, H. Takeuchi, G. Foussias, M. K. Hahn, O. Agid, and G. Remington. "Achievement motivation in early schizophrenia: Relationship with symptoms, cognition and functional outcome". eng. *Early Interv Psychiatry* 12.6 (2018), pp. 1038–1044. DOI: `10.1111/eip.12405`.

[24] K. Allott, P. Liu, T.-M. Proffitt, and E. Killackey. "Cognition at illness onset as a predictor of later functional outcome in early psychosis: systematic review and methodological critique". *Schizophrenia research* 125.2-3 (2011), pp. 221–235. ISSN: 0920-9964.

[25] R. Van Winkel, I. Myin-Germeys, M. De Hert, P. Delespaul, J. Peuskens, and J. Van Os. "The association between cognition and functional outcome in first-episode patients with schizophrenia: mystery resolved?" *Acta Psychiatrica Scandinavica* 116.2 (2007), pp. 119–124.

[26] A. Reichenberg, P. D. Harvey, C. R. Bowie, R. Mojtabai, J. Rabinowitz, R. K. Heaton, and E. Bromet. "Neuropsychological function and dysfunction in schizophrenia and psychotic affective disorders". *Schizophrenia bulletin* 35.5 (2009), pp. 1022–1029.

[27] A. McCleery and K. H. Nuechterlein. "Cognitive impairment in psychotic illness: prevalence, profile of impairment, developmental course, and treatment considerations". *Dialogues in clinical neuroscience* (2022).

[28] R. S. Keefe and W. S. Fenton. "How should DSM-V criteria for schizophrenia include cognitive impairment?" *Schizophrenia bulletin* 33.4 (2007), pp. 912–920.

[29] B. Fagerlund, A. K. Pagsberg, and R. P. Hemmingsen. "Cognitive deficits and levels of IQ in adolescent onset schizophrenia and other psychotic disorders". *Schizophrenia research* 85.1-3 (2006), pp. 30–39.

[30] M. Aas, P. Dazzan, H. L. Fisher, C. Morgan, K. Morgan, A. Reichenberg, J. Zanelli, P. Fearon, P. B. Jones, R. M. Murray, et al. "Childhood trauma and cognitive function in first-episode affective and non-affective psychosis". *Schizophrenia research* 129.1 (2011), pp. 12–19.

[31] D. M. Barch and J. M. Sheffield. "Cognitive impairments in psychotic disorders: common mechanisms and measurement". *World Psychiatry* 13.3 (2014), pp. 224–232.

[32] N. Craddock, M. C. O'Donovan, and M. J. Owen. "Psychosis genetics: modeling the relationship between schizophrenia, bipolar disorder, and mixed (or "schizoaffective") psychoses". *Schizophrenia bulletin* 35.3 (2009), pp. 482–490.

[33] S. Amoretti, A. R. Rosa, G. Mezquida, B. Cabrera, M. Ribeiro, M. Molina, M. Bioque, A. Lobo, A. González-Pinto, D. Fraguas, and et al. "The impact of cognitive reserve, cognition and clinical symptoms on psychosocial functioning in first-episode psychoses". *Psychological Medicine* 52.3 (2022), pp. 526–537. DOI: 10.1017/S0033291720002226.

[34] O. Santesteban-Echarri, M. Paino, S. Rice, C. González-Blanch, P. McGorry, J. Gleeson, and M. Alvarez-Jimenez. "Predictors of functional recovery in first-episode psychosis: a systematic review and meta-analysis of longitudinal studies". *Clinical Psychology Review* 58 (2017), pp. 59–75.

[35] D. A. Lewis. "Cortical circuit dysfunction and cognitive deficits in schizophrenia–implications for preemptive interventions". *European Journal of Neuroscience* 35.12 (2012), pp. 1871–1878.

[36] M.-y. Xu and A. H. Wong. "GABAergic inhibitory neurons as therapeutic targets for cognitive impairment in schizophrenia". *Acta Pharmacologica Sinica* 39.5 (2018), pp. 733–753.

[37] S.-J. Blakemore and S. Choudhury. "Development of the adolescent brain: implications for executive function and social cognition". *Journal of child psychology and psychiatry* 47.3-4 (2006), pp. 296–312.

[38] I. Feinberg. "Schizophrenia: caused by a fault in programmed synaptic elimination during adolescence?" *Journal of psychiatric research* 17.4 (1982), pp. 319–334.

[39] J. Kambeitz, L. Kambeitz-Ilankovic, C. Cabral, D. B. Dwyer, V. D. Calhoun, M. P. Van Den Heuvel, P. Falkai, N. Koutsouleris, and B. Malchow. "Aberrant functional whole-brain network architecture in patients with schizophrenia: a meta-analysis". *Schizophrenia Bulletin* 42.suppl_1 (2016), S13–S21.

[40] P. Fusar-Poli, J. Radua, P. McGuire, and S. Borgwardt. "Neuroanatomical maps of psychosis onset: voxel-wise meta-analysis of antipsychotic-naive VBM studies". *Schizophrenia bulletin* 38.6 (2012), pp. 1297–1307.

[41] T. M. Minatogawa-Chang, M. S. Schaufelberger, A. M. Ayres, F. L. Duran, E. K. Gutt, R. M. Murray, T. M. Rushe, P. K. McGuire, P. R. Menezes, M. Scazufca, et al. "Cognitive performance is related to cortical grey matter volumes in early stages of schizophrenia: a population-based study of first-episode psychosis". *Schizophrenia research* 113.2-3 (2009), pp. 200–209.

[42] N. Banaj, F. Piras, F. Piras, V. Ciullo, M. Iorio, C. Battaglia, D. Pantoli, G. Ducci, and G. Spalletta. "Cognitive and psychopathology correlates of brain white/grey matter structure in severely psychotic schizophrenic inpatients". *Schizophrenia research: cognition* 12 (2018), pp. 29–36.

[43] W. Mark and T. Toulopoulou. "Cognitive intermediate phenotype and genetic risk for psychosis". *Current Opinion in Neurobiology* 36 (2016). Neurobiology of disease, pp. 23–30. ISSN: 0959-4388. DOI: `https://doi.org/10.1016/j.conb.2015.08.008`.

[44] R. Shafee, P. Nanda, J. L. Padmanabhan, N. Tandon, N. Alliey-Rodriguez, S. Kalapurakkel, D. J. Weiner, R. E. Gur, R. S. Keefe, S. K. Hill, et al. "Polygenic risk for schizophrenia and measured domains of cognition in individuals with psychosis and controls". *Translational psychiatry* 8.1 (2018), pp. 1–9.

[45] K. E. Lewandowski, J. T. Baker, J. M. McCarthy, L. A. Norris, and D. Öngür. "Reproducibility of Cognitive Profiles in Psychosis Using Cluster Analysis". *Journal of the International Neuropsychological Society* 24.4 (2018), pp. 382–390. DOI: 10.1017/S1355617717001047.

[46] M. J. Green, L. Girshkin, K. Kremerskothen, O. Watkeys, and Y. Quidé. "A systematic review of studies reporting data-driven cognitive subtypes across the psychosis spectrum". *Neuropsychology Review* 30.4 (2020), pp. 446–460.

[47] A. Breier, M. Shenton, D. Holt, M. Keshavan, D. Öngür, and L. Seidman. *HCP Early Psychosis 1.0 Data Release: Reference Manual 9*. Tech. rep. August. Indianapolis: Boston, MA: Brigham and Women's Hospital, Beth Israel Deaconess-Massachusetts Mental Health Center, McLean Hospital, Massachusetts General Hospital; 2021.

[48] L. Phan, H. Liu, and C. Tortora. "K-Means Clustering on Multiple Correspondence Analysis Coordinates". *Researchgate.Net* 1.1 (2019), pp. 1–17. ISSN: 2510-0564. DOI: 10.5445/KSP/1000085952/05.

[49] M. A. Terpilowski. "scikit-posthocs: Pairwise multiple comparison tests in Python". *Journal of Open Source Software* 4.36 (2019), p. 1169. DOI: 10.21105/joss.01169.

[50] R. Vallat. "Pingouin: statistics in Python". *Journal of Open Source Software* 3.31 (2018), p. 1026. DOI: 10.21105/joss.01026.

[51] M. Carbon and C. U. Correll. "Thinking and acting beyond the positive: the role of the cognitive and negative symptoms in schizophrenia". *CNS Spectrums* 19.S1 (2014), pp. 35–53. DOI: 10.1017/S1092852914000601.

[52] M. J. Kas, B. Penninx, B. Sommer, A. Serretti, C. Arango, and H. Marston. "A quantitative approach to neuropsychiatry: The why and the how." eng. *Neuroscience and biobehavioral reviews* 97 (Feb. 2019), pp. 3–9. ISSN: 1873-7528 (Electronic). DOI: 10.1016/j.neubiorev.2017.12.008.

[53] P. F. Hitchcock, E. I. Fried, and M. J. Frank. "Computational psychiatry needs time and context". *Annual review of psychology* 73 (2022), p. 243.

[54] P. D. Harvey, D. Koren, A. Reichenberg, and C. R. Bowie. "Negative Symptoms and Cognitive Deficits: What Is the Nature of Their Relationship?" *Schizophrenia Bulletin* 32.2 (Apr. 2006), pp. 250–258. ISSN: 0586-7614. DOI: 10.1093/schbul/sbj011.

[55] L. Leanza, L. Egloff, E. Studerus, C. Andreou, U. Heitz, S. Ittig, K. Beck, M. Uttinger, and A. Riecher-Rössler. "The relationship between negative symptoms and cognitive functioning in patients at clinical high risk for psychosis". *Psychiatry Research* 268 (2018), pp. 21–27.

[56] C. L. Hovington, M. Bodnar, R. Joober, A. K. Malla, and M. Lepage. "Impairment in verbal memory observed in first episode psychosis patients with persistent negative symptoms". *Schizophrenia research* 147.2-3 (2013), pp. 223–229.

[57] P. M. Grant and A. T. Beck. "Defeatist beliefs as a mediator of cognitive impairment, negative symptoms, and functioning in schizophrenia". *Schizophrenia bulletin* 35.4 (2009), pp. 798–806.

[58] L. Glenthøj, J. R. M. Jepsen, C. Hjorthøj, N. Bak, T. Kristensen, C. Wenneberg, K. Krakauer, M. Nordentoft, and B. Fagerlund. "Negative symptoms mediate the relationship between neurocognition and function in individuals at ultrahigh risk for psychosis". *Acta Psychiatrica Scandinavica* 135.3 (2017), pp. 250–258.

[59] D. S. O'Leary, M. Flaum, M. L. Kesler, L. A. Flashman, S. Arndt, and N. C. Andreasen. "Cognitive correlates of the negative, disorganized, and psychotic symptom dimensions of schizophrenia". *The Journal of neuropsychiatry and clinical neurosciences* 12.1 (2000), pp. 4–15.

[60] Y. Zhu, F. Y. Womer, H. Leng, M. Chang, Z. Yin, Y. Wei, Q. Zhou, S. Fu, X. Deng, J. Lv, et al. "The relationship between cognitive dysfunction and symptom dimensions across schizophrenia, bipolar disorder, and major depressive disorder". *Frontiers in psychiatry* 10 (2019), p. 253.

[61] O. D Howes, P. Fusar-Poli, M. Bloomfield, S. Selvaraj, and P. McGuire. "From the prodrome to chronic schizophrenia: the neurobiology underlying psychotic symptoms and cognitive impairments". *Current pharmaceutical design* 18.4 (2012), pp. 459–465.

[62] J. Horwood, G. Salvi, K. Thomas, L. Duffy, D. Gunnell, C. Hollis, G. Lewis, P. Menezes, A. Thompson, D. Wolke, and et al. "IQ and non-clinical psychotic symptoms in 12-year-olds: results from the ALSPAC birth cohort". *British Journal of Psychiatry* 193.3 (2008), pp. 185–191. DOI: 10.1192/bjp.bp.108.051904.

[63] F. Guillem, M. Rinaldi, T. Pampoulova, and E. Stip. "The complex relationships between executive functions and positive symptoms in schizophrenia". *Psychological Medicine* 38.6 (2008), pp. 853–860. DOI: 10.1017/S0033291707002577.

[64] L. Lecardeur, E. Stip, M. Giguere, G. Blouin, J.-P. Rodriguez, and M. Champagne-Lavau. "Effects of cognitive remediation therapies on psychotic symptoms and cognitive complaints in patients with schizophrenia and related disorders: a randomized study". *schizophrenia Research* 111.1-3 (2009), pp. 153–158.

[65] J. Ventura, K. L. Subotnik, D. Gretchen-Doorly, L. Casaus, M. Boucher, A. Medalia, M. D. Bell, G. S. Hellemann, and K. H. Nuechterlein. "Cognitive remediation can improve negative symptoms and social functioning in first-episode schizophrenia: a randomized controlled trial". *Schizophrenia research* 203 (2019), pp. 24–31.

[66] R. Gallardo-Ruiz, B. Crespo-Facorro, E. Setién-Suero, and D. Tordesillas-Gutierrez. "Long-term grey matter changes in first episode psychosis: a systematic review". *Psychiatry Investigation* 16.5 (2019), p. 336.

[67] A. Vita, L. De Peri, G. Deste, and E. Sacchetti. "Progressive loss of cortical gray matter in schizophrenia: a meta-analysis and meta-regression of longitudinal MRI studies". *Translational psychiatry* 2.11 (2012), e190–e190.

[68] P. Fusar-Poli, R. Smieskova, G. Serafini, P. Politi, and S. Borgwardt. "Neuroanatomical markers of genetic liability to psychosis and first episode psychosis: a voxelwise meta-analytical comparison". *The World Journal of Biological Psychiatry* 15.3 (2014), pp. 219–228.

[69] F. Knolle, S. S. Arumugham, R. A. Barker, M. W. Chee, A. Justicia, N. Kamble, J. Lee, S. Liu, A. Lenka, S. J. Lewis, et al. "Grey matter morphometric biomarkers for classifying early schizophrenia and PD psychosis: a multicentre study". *medRxiv* (2022).

[70] S. Iwabuchi, P. F. Liddle, and L. Palaniyappan. "Clinical utility of machine-learning approaches in schizophrenia: improving diagnostic confidence for translational neuroimaging". *Frontiers in psychiatry* 4 (2013), p. 95.

[71] H. G. Schnack, M. Nieuwenhuis, N. E. van Haren, L. Abramovic, T. W. Scheewe, R. M. Brouwer, H. E. H. Pol, and R. S. Kahn. "Can structural MRI aid in clinical classification? A machine learning study in two independent samples of patients with schizophrenia, bipolar disorder and healthy subjects". *Neuroimage* 84 (2014), pp. 299–306.

[72] S. Vieira, Q.-y. Gong, W. H. Pinaya, C. Scarpazza, S. Tognin, B. Crespo-Facorro, D. Tordesillas-Gutierrez, V. Ortiz-Garcıa, E. Setien-Suero, F. E. Scheepers, et al. "Using machine learning and structural neuroimaging to detect first episode psychosis: reconsidering the evidence". *Schizophrenia bulletin* 46.1 (2020), pp. 17–26.

[73] F. Knolle, E. F. Sterner, M. Moutoussis, R. A. Adams, J. D. Griffin, J. Haarsma, H. Taverne, I. M. Goodyer, P. C. Fletcher, G. K. Murray, et al. "Action selection in early stages of psychosis: an active inference approach". *medRxiv* (2022).

[74] M. Montagnese, F. Knolle, J. Haarsma, J. D. Griffin, A. Richards, P. E. Vertes, B. Kiddle, P. C. Fletcher, P. B. Jones, M. J. Owen, et al. "Reinforcement learning as an intermediate phenotype in psychosis? Deficits sensitive to illness stage but not associated with polygenic risk of schizophrenia in the general population". *Schizophrenia research* 222 (2020), pp. 389–396.

[75] A. Mkrtchian, V. Valton, and J. P. Roiser. "Reliability of Decision-Making and Reinforcement Learning Computational Parameters". *bioRxiv* (2021). DOI: 10.1101/2021.06.30.450026. eprint: https://www.biorxiv.org/content/early/2021/11/06/2021.06.30.450026.full.pdf.