

# Reproducibility of Cognitive Profiles in Psychosis Using Cluster Analysis

Kathryn E. Lewandowski,<sup>1,2</sup> Justin T. Baker,<sup>1,2</sup> Julie M. McCarthy,<sup>1,2</sup> Lesley A. Norris,<sup>1</sup> AND Dost Öngür<sup>1,2</sup>

<sup>1</sup>McLean Hospital, Schizophrenia and Bipolar Disorder Program, Belmont, Massachusetts

<sup>2</sup>Harvard Medical School, Department of Psychiatry, Boston, Massachusetts

(RECEIVED January 12, 2017; FINAL REVISION July 20, 2017; ACCEPTED August 26, 2017; FIRST PUBLISHED ONLINE October 18, 2017)

## Abstract

**Objectives:** Cognitive dysfunction is a core symptom dimension that cuts across the psychoses. Recent findings support classification of patients along the cognitive dimension using cluster analysis; however, data-derived groupings may be highly determined by sampling characteristics and the measures used to derive the clusters, and so their interpretability must be established. We examined cognitive clusters in a cross-diagnostic sample of patients with psychosis and associations with clinical and functional outcomes. We then compared our findings to a previous report of cognitive clusters in a separate sample using a different cognitive battery. **Methods:** Participants with affective or non-affective psychosis ( $n = 120$ ) and healthy controls ( $n = 31$ ) were administered the MATRICS Consensus Cognitive Battery, and clinical and community functioning assessments. Cluster analyses were performed on cognitive variables, and clusters were compared on demographic, cognitive, and clinical measures. Results were compared to findings from our previous report.

**Results:** A four-cluster solution provided a good fit to the data; profiles included a neuropsychologically normal cluster, a globally impaired cluster, and two clusters of mixed profiles. Cognitive burden was associated with symptom severity and poorer community functioning. The patterns of cognitive performance by cluster were highly consistent with our previous findings. **Conclusions:** We found evidence of four cognitive subgroups of patients with psychosis, with cognitive profiles that map closely to those produced in our previous work. Clusters were associated with clinical and community variables and a measure of premorbid functioning, suggesting that they reflect meaningful groupings: replicable, and related to clinical presentation and functional outcomes. (*JINS*, 2018, 24, 382–390)

**Keywords:** Cognition, Cluster, Schizophrenia, Bipolar disorder, Psychosis, Replication

## INTRODUCTION

Cognitive dysfunction is a core and disabling symptom dimension that cuts across the psychoses. Deficits in patients with schizophrenia (SZ) and related disorders, including affective psychosis, have been reported in multiple cognitive domains including attention, verbal processing, working memory, and executive functioning; however, substantial cognitive heterogeneity exists both within and between diagnostic categories, and cognitive profiles characteristic of diagnostic groups have not emerged (Goldstein, 1990; Seaton, Goldstein, & Allen, 2001).

A renewed interest in data-driven grouping of participants along the cognitive symptom dimension suggests that across diagnostic groups patients can be classified by cognitive profiles using cluster analysis (Burdick et al., 2014; Lewandowski, Sperry, Cohen, & Öngür, 2014; Van Rheenen

et al., 2017), and that such groupings may be more closely related to neurobiology than diagnostic categories (Clementz et al., 2016). Separable cognitive clusters can be identified by the first episode (Reser, Allott, Killackey, Farhall, & Cotton, 2015), indicating that profiles may reflect underlying illness traits. Broadly, results from these recent studies are consistent with earlier reports in patients with SZ, which tended to find distinct cognitive subgroups of patients including a “neuropsychologically normal” group, a group with significant global deficits, and one to three additional groups with mixed cognitive profiles (Allen, Goldstein, & Warnick, 2003; Goldstein, Allen, & Seaton, 1998; Goldstein, Beers, & Shemansky, 1996; Goldstein & Shemansky, 1995; Heinrichs & Awad, 1993; Hill, Ragland, Gur, & Gur, 2002; Palmer et al., 1997; Seaton et al., 2001).

Cognitive clusters are associated with distinct clinical features. Patients with predominantly positive symptom SZ (e.g., paranoia) are more likely to be in the “neuropsychologically normal” cluster (Goldstein, Shemansky, & Allen, 2005; Hill et al., 2002), whereas patients with more prominent negative or disorganized features are more likely to fall

Correspondence and reprint requests to: Kathryn E. Lewandowski, McLean Hospital, 115 Mill Street, AB S349, Belmont, MA 02478. E-mail: klewandowski@mclean.harvard.edu

in cognitively impaired clusters. In terms of functional outcomes, subjects with neuropsychologically normal profiles exhibit increased socialization, fewer hospitalizations, and better overall community functioning than patients in clusters characterized by cognitive impairment (Allen et al., 2003; Lewandowski et al., 2014; Palmer et al., 1997).

While data-driven approaches to classification may produce homogeneous groupings along a dimension of interest, the interpretability and validity of findings is open to debate, due in part to potential problems with reproducibility in psychological research generally (Open Science Collaboration, 2015; Pashler & Harris, 2012; Simmons, Nelson, & Simonsohn, 2011), and based on limitations inherent in data-driven approaches themselves. In cluster analysis, emergent clusters are highly determined by sampling characteristics and the measures used to derive the clusters, making reproducibility critical to establishing the validity and utility of groups. In cognitive cluster analytic studies in SZ, for instance, intermediate clusters in particular vary considerably from study to study, and studies reporting the greatest comparability of emergent clusters tend to be those that use the same measures (Goldstein et al., 1998; Heinrichs, Ruttan, Zakzanis, & Case, 1997; Seaton, Allen, Goldstein, Kelley, & van Kammen, 1999). Conceptually, replication of cluster solutions in separate samples using different measures that tap the same underlying constructs but may have different psychometric properties may be particularly important for establishing the validity and generalizability of emergent cluster solutions (Stroebe & Strack, 2014).

In the present study, we used cluster analysis to examine cognitive clusters in a cross-diagnostic sample of patients with psychosis and the associations among clusters and clinical and functional variables, and to assess the reproducibility of our previous cluster solution (Lewandowski et al., 2014) in this new sample using a different (but related) cognitive battery and the same clustering approach. It was hypothesized that (1) a four-cluster solution would provide a good fit to the data, (2) cognitive profiles would map closely to our previous findings, and (3) associations between clusters and clinical and functional variables would replicate our previous findings, including (a) diagnoses distributed across clusters, and (b) greater symptom severity and poorer community functioning associated with greater neuropsychological burden.

## METHOD

### Participants

Participants with diagnoses of affective or non-affective psychosis ( $n = 120$ ) and healthy controls ( $n = 31$ ) were recruited through the Schizophrenia and Bipolar Disorder Program (SBDP) and *via* fliers posted at McLean Hospital. Participants were recruited in the context of several separate but related studies: (1) cognitive remediation in SZ or bipolar disorder (BD) ( $n = 42$ ), (2) neuroimaging ( $n = 33$ ), or (3) clinical characterization of psychosis ( $n = 76$ ). For subjects recruited into longitudinal studies, baseline measures were used. Inclusion

criteria included a DSM-IV diagnosis of SZ, schizoaffective disorder (SZA), schizophreniform disorder, psychosis not otherwise specified (NOS), BD I with psychosis, or major depressive disorder (MDD) with psychosis, ages 18 to 65. Exclusion criteria for all participants included history of head trauma with loss of consciousness, history of seizure disorder, and current substance abuse or dependence. Healthy controls had no personal or first-degree family history of a psychiatric diagnosis, and no history of substance abuse or dependence. This study and all associated procedures comply with the Helsinki Declaration of 1975, as revised in 2008, and were approved by the McLean Hospital Institutional Review Board.

### Materials

Diagnosis was determined using the SCID-IV-TR through patient interview, medical record review, and consultation with the participants' treatment providers. Clinical assessment included the Young Mania Rating Scale (YMRS; Young, Biggs, Ziegler, & Meyer, 1978), the Montgomery-Asberg Depression Rating Scale (MADRS; Montgomery & Asberg, 1979), and the Positive and Negative Syndrome Scale (PANSS; Kay, Fiszbein, & Opler, 1987). Community functioning was measured using an abbreviated version of the Multnomah Community Ability Scale (MCAS; Barker, Barron, McFarland, Bigelow, & Carnahan, 1994) as described in Lewandowski et al. (Lewandowski, Cohen, Keshavan, Sperry, & Ongür, 2013). Premorbid IQ was measured with the North American Adult Reading Test (NAART; Uttl, 2002).

Cognition was measured using the MATRICS Consensus Cognitive Battery (MCCB; Nuechterlein et al., 2008). The MCCB is comprised of 10 subtests that make up seven domain scores and a composite. Domains include: Speed of Processing (Trail Making Test A; Brief Assessment of Cognition in Schizophrenia: Symbol Coding; Category Fluency); Attention/Vigilance (Continuous Performance Test: Identical Pairs); Working Memory (Wechsler Memory Scale Spatial Span; Letter Number Span); Visual Learning (Brief Visuospatial Memory Test); Verbal Learning (Hopkins Verbal Learning Test); Reasoning/Problem Solving (Neuropsychological Assessment Battery: Mazes); and Social Cognition (Mayer-Salovey-Caruso Emotional Intelligence Test: Managing Emotions). Subtest, domain, and composite scores are converted to T-scores based on MCCB scoring software using age and gender adjusted norms. The MCCB takes approximately 60–90 min to complete.

### Procedure

Neuropsychological and clinical data were collected in one to two sessions. Patient-reported information regarding medication was collected, and chlorpromazine (CPZ) equivalents were calculated using guidelines described by Baldessarini (2012).

### Statistical Approach

Cluster analysis was performed in STATA 12.1 (StataCorp LP, USA) first with hierarchical cluster analysis using Ward's

method including patients only to replicate our previous approach. The nine MCCB neuropsychological subtest scores were entered into the cluster analyses; we excluded the Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT) test of social cognition because our original cluster analysis did not include a social cognition measure. Age- and gender-adjusted T-scores were reported for all MCCB variables for ease of comparison and so that each variable contributed equally to the distance measure. Cluster fit was determined by visual inspection of the dendrogram and “elbow test,” and acceptability was confirmed using discriminant function analysis to assess classification accuracy. While one of our aims was to replicate our previous findings of a four-cluster solution, hierarchical clustering was undertaken first to evaluate the appropriateness of forcing a four-cluster solution, and to more thoroughly explore whether a different number of clusters clearly offered a better fit to these data.

Next, a K-means cluster analysis was performed entering the above nine MCCB subscale scores, specifying the number of clusters selected based on our previous report and/or the number supported by the hierarchical approach. K-means was added as a second step because it is an iterative cluster approach which creates more stable cluster solutions. Linear discriminant function analysis was performed to evaluate classification accuracy, and the final K-means clusters were used for all subsequent group analyses.

Clusters were compared on demographic and clinical variables, and MCCB cognitive domain and composite scores using analysis of variance (ANOVA) or  $\chi^2$  as appropriate. *Post hoc* paired *t* tests were conducted for all significant ANOVA effects, with Bonferroni correction. Linear regressions controlling for clinical variables, CPZ equivalents, and diagnosis were conducted predicting cognition by cluster; similar regressions controlling for clinical variables, CPZ equivalents, diagnosis, and education were conducted predicting community functioning.

Finally, to compare the present cluster profiles to those of our previous report (Lewandowski et al., 2014), we attempted to create comparable cognitive domains using the same standardization across the two data sets. We combined our previous cognitive subtests into domains comparable to MCCB domains, including Processing Speed (Stroop, Trails A, and Category Fluency), Verbal (Hopkins Verbal Learning Test Total Recall), Visual (Brief Visuospatial Memory Test Total Recall), Executive Functions (Stroop Interference, Trails B), and a Composite (average of all cognitive testing scores). Previously reported Z-scores were converted to T-scores for ease of comparison.

The MCCB Working Memory and Problem Solving Domains from the present data were averaged to create an “Executive Functions” domain, comparable to our previous report. Cognitive data from the combined (previous and current) sample was entered into a K-means cluster analysis with a forced four-cluster solution. Variables entered into the cluster analysis included the domain scores described above: Processing Speed, Verbal, Visual, and Executive Functions.

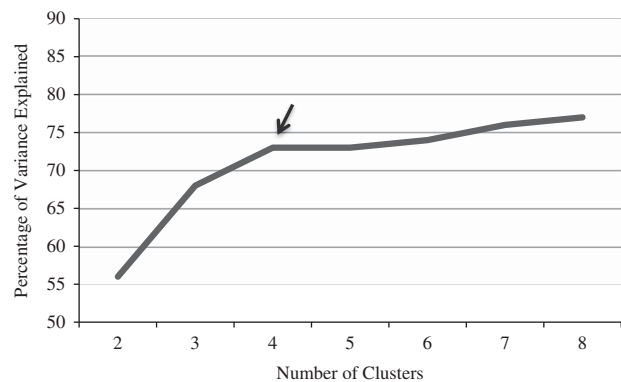
Stability of cluster assignment across the total sample was then evaluated by examining the proportion of the combined sample that was re-classified into the same cluster as their original cluster assignment.

## RESULTS

### Cluster Solution

MCCB subtest scores were entered into a hierarchical cluster analysis using Ward’s linkage with squared Euclidean distance measure. Visual inspection of the resulting dendrogram provided support for two, three, or four clusters. The two-cluster solution resulted in broadly “intact” and “impaired” groups (composite means (sd):  $T = 54.9$  (4.8) and  $T = 34.3$  (10.6), respectively). The three-cluster solution included broadly “intact,” “moderately impaired,” and “significantly impaired” groups (composite means (sd):  $T = 54.9$  (4.8),  $T = 38.1$  (7.0), and  $T = 19.6$  (9.2), respectively). The four-cluster solution further divided the “moderately impaired” group into two mixed-profile clusters; the “intact” and “significantly impaired” groups were unchanged. Linear discriminant function analysis showed significant differentiation and a high degree of classification accuracy in the two-, three-, and four-cluster models (95%, 95%, and 93%, respectively). We then conducted an “elbow test,” plotting the percentage of variance explained in the Composite by two to eight clusters (Figure 1), which shows a clear “elbow” at four clusters. Thus, we found adequate support for the specification of a four-cluster solution.

Next, MCCB subtest scores (excluding MSCEIT) were entered into a K-means cluster analysis specifying a four-cluster solution. Fit of the four-cluster solution using linear discriminant analysis indicated good fit of the model, with 94% classification accuracy (Cluster 1 ( $n = 39$ ): 97%; Cluster 2 ( $n = 42$ ): 91%, Cluster 3 ( $n = 18$ ): 89%, Cluster 4 ( $n = 21$ ): 100%). Cognitive profiles by cluster are described below.



**Fig. 1** . Hierarchical cluster elbow test. Proportion of variance (between-group variance/total variance) explained as a function of the number of clusters. Using the “elbow method,” a cluster solution is chosen based on the inflection point at which adding additional clusters does not substantially improve modeling of the data.

**Table 1.** MCCB Domain Scores by Cluster

Cognitive Domain	Control (n = 31)	Cluster 1 (n = 39)	Cluster 2 (n = 42)	Cluster 3 (n = 18)	Cluster 4 (n = 21)	F-statistic df (4, 146)	Post-hoc t-test <sup>a</sup>
Processing Speed	59.6 (10.5)	54.7 (6.8)	44.0 (6.3)	32.2 (5.6)	30.0 (10.3)	69.06****	4 < 3 < 2 < 1
Attention/Vigilance	53.0 (9.0)	52.3 (8.8)	43.2 (8.9)	34.5 (8.7)	26.1 (8.9)	42.88****	4 < 3 < 2 < 1
Working Memory	53.4 (10.3)	56.1 (7.8)	44.0 (7.4)	43.3 (7.7)	26.2 (8.4)	50.70****	4 < 3, 2 < 1
Verbal Learning	53.7 (10.1)	53.4 (10.1)	42.3 (8.3)	38.5 (6.6)	36.7 (6.7)	21.89****	4 < 3, 2 < 1
Visual Learning	52.2 (8.5)	51.6 (7.7)	32.1 (9.2)	47.4 (9.7)	28.6 (7.2)	38.14****	4 < 2 < 3, 1
Problem Solving	49.9 (8.7)	52.7 (8.2)	45.1 (9.8)	42.8 (7.8)	36.7 (7.1)	14.3****	4 < 2, 1; 3, 2 < 1
Social Cognition	53.6 (10.7)	52.1 (9.1)	49.3 (10.6)	43.9 (12.9)	43.4 (11.4)	4.69**	4 < 1
COMPOSITE	55.7 (9.6)	55.1 (4.4)	40.3 (5.3)	34.2 (6.1)	20.9 (8.6)	121.86****	4 < 3 < 2 < 1

<sup>a</sup>Clusters 2, 3 and 4 differed from Controls on all measures; Cluster 1 did not differ from controls on any measure. All reported post-hoc group differences were significant after Bonferroni correction.

\*\*  $p < .01$  \*\*\*\*  $p < .0001$

### Cognitive Characteristics by Cluster

Neurocognitive profiles by K-means cluster were examined using the seven MCCB domain scores and the composite (Table 1). Profiles included a neuropsychologically normal cluster (Cluster 1; Composite T = 55.1) in which all scores were at or slightly above the mean, a globally and significantly impaired cluster (Cluster 4; Composite T = 20.9) in which all scores fell between 1 and 3 SDs below the mean, and two clusters with variable cognitive profiles (Clusters 2 and 3). Cluster 2 exhibited mild or no impairment in processing speed, attention, verbal learning, working memory, and problem solving, and selective visual learning and memory impairment in the moderate range; social cognition was intact. The Composite score for this group was mildly impaired (T = 40.3). Cluster 3 exhibited intact visual learning and memory, mild or no impairment in working memory, problem solving, verbal learning, and social cognition, and moderate impairment in processing speed and attention; the Composite score for this group was moderately impaired (T = 34.2).

The cognitive composite differed significantly by cluster ( $F_{(4,146)} = 121.86$ ;  $p < .0001$ ); Bonferroni-corrected pairwise

comparisons indicated that all clusters differed from each other ( $p < .05$  to  $p < .0001$ ); Clusters 2, 3, and 4 differed from Controls ( $p < .0001$ ). Cluster 1 did not differ from Controls on any MCCB domain or the composite.

### Demographic and Clinical Characteristics by Cluster

Participants differed on several demographic and clinical variables, including education and premorbid IQ (Table 2). Patient clusters differed from each other on state mania (4 > 1, 2), positive symptoms (4 > 1), negative symptoms (4 > 1), and CPZ equivalents (4 > 1, 2).

We conducted a series of linear regressions to examine effects of cluster membership on cognitive outcomes after controlling for the potential effects of state clinical variables and medication (YMRS, PANSS Positive, PANSS Negative, CPZ equivalents) and diagnosis to examine the possibility that differences in cognitive performance by cluster were better accounted for by overall illness severity. Cluster membership remained a significant predictor of

**Table 2.** Demographic and Clinical Characteristics by Cluster

Variable	Control (n = 31)	Cluster 1 (n = 39)	Cluster 2 (n = 42)	Cluster 3 (n = 18)	Cluster 4 (n = 21)	F-statistic df (4, 146)
Age, years	33.1 (10.9)	31.1 (8.6)	32.2 (11.5)	29.6 (11.3)	32.8 (13.9)	0.38 <sup>ns</sup>
Education, years	16.1 (1.9)	16.1 (1.7)	14.6 (1.7)	14.1 (1.3)	13.1 (1.9)	15.24****
Sex, % female	45%	38%	40%	11%	38%	Chi2 = 6.32 <sup>ns</sup>
Race, % Caucasian	81%	90%	76%	94%	76%	Chi2 = 5.06 <sup>ns</sup>
NAART <sup>a</sup>	114.2 (5.5)	117.2 (6.9)	112.5 (10.1)	112.5 (7.4)	105.8 (12.5)	4.21**
YMRS	–	4.6 (4.9)	5.1 (5.0)	5.6 (7.1)	10.2 (7.9)	4.42**
MADRS	–	10.3 (9.2)	12.0 (9.9)	8.9 (6.8)	11.2 (10.0)	0.53 <sup>ns</sup>
PANSS positive	–	10.2 (4.1)	11.4 (4.9)	11.2 (5.5)	15.0 (6.5)	4.13**
PANSS negative	–	10.6 (3.8)	12.7 (4.8)	13.9 (6.9)	15.8 (7.0)	4.52**
PANSS general	–	21.9 (5.6)	26.7 (6.8)	26.3 (7.5)	28.3 (9.6)	1.97 <sup>ns</sup>
CPZ Equivalents	–	126.4 (173.2)	212.0 (257.8)	283.6 (185.4)	408.5 (349.7)	6.10****
MCAS	54.6 (1.2)	48.3 (4.3)	46.1 (5.1)	44.8 (6.6)	41.1 (9.0)	11.88****

<sup>a</sup>NAART: North American Adult Reading Test

\*\*  $p < .01$  \*\*\*\*  $p < .0001$

all neurocognitive domains and the composite ( $t = -17.19$  to  $-5.93$ ;  $p < .001$ ), but not social cognition ( $t = -1.07$ ;  $p = .29$ ).

Groups differed on community functioning, with all four patient clusters scoring worse than Controls; Cluster 4 scored worse than clusters 1 and 2. Linear regression showed that cluster remained a significant predictor of community functioning after accounting for clinical variables and CPZ equivalents, diagnosis, and education ( $\beta = -1.36$ ;  $t = -2.26$ ;  $p = .03$ ).

**Diagnostic Distribution by Cluster**

The distribution of diagnoses by cluster was examined by collapsing patients into three main diagnostic groups: mood disorders with psychosis (MDP; BD,  $n = 54$ ; MDD,  $n = 10$ ); SZA ( $n = 28$ ); and SZ ( $n = 21$ ; schizophreniform,  $n = 5$ ; psychosis NOS,  $n = 2$ ) to mirror our previous approach and to avoid cells with five or fewer participants. All diagnoses were represented in each cluster, although not evenly (Table 3). Patients with MDP were overrepresented in Cluster 1 and slightly underrepresented in Cluster 4; patients with SZ were underrepresented in Cluster 1 and overrepresented in Clusters 3 and 4; patients with SZA were overrepresented in Cluster 2 and underrepresented in Cluster 1. A spring-loaded graph depicting the similarity (inverse distance) between each participant in nine-dimensional space based on MCCB subtest scores, then overlaid by diagnoses, demonstrates that diagnoses do not map to the cognitive clusters but rather are distributed throughout the space (Figure 2).

*Post hoc* analyses of cognitive domains by diagnosis showed that diagnoses differed significantly on the cognitive composite score, with scores in keeping with the published

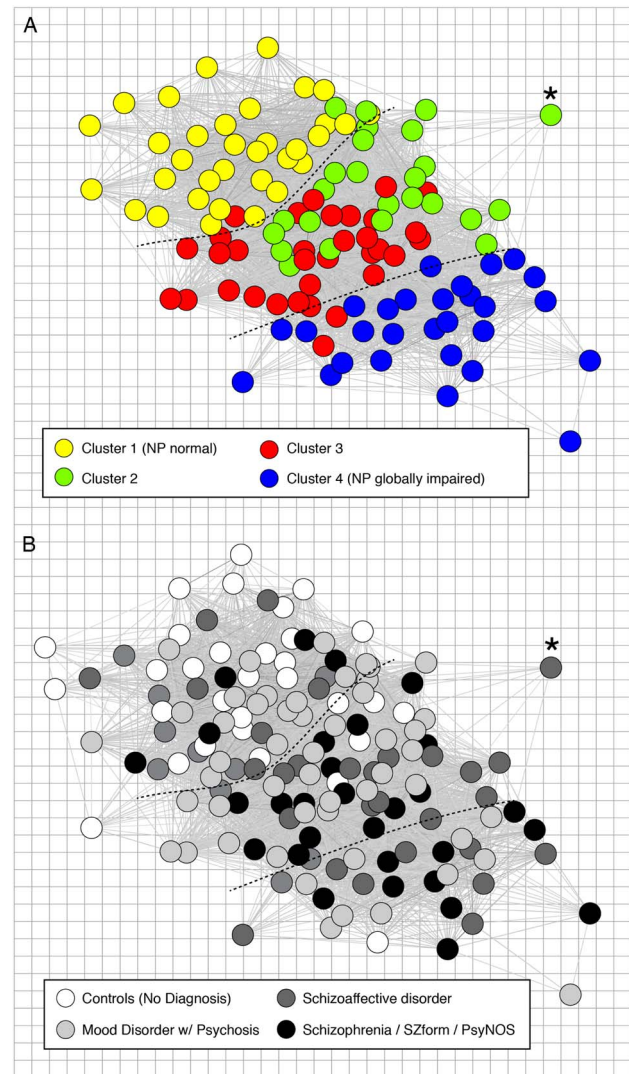
**Table 3.** Distribution of Diagnoses by Cluster <sup>a</sup>

	Diagnosis			
	MDP	SZ	SZA	Total
<b>Cluster 1</b>				
Observed frequency	28	5	6	39
Expected frequency	20.8	9.1	9.1	39.0
$\chi^2$ Contribution	2.5	1.8	1.1	5.4
<b>Cluster 2</b>				
Observed frequency	21	8	13	42
Expected frequency	22.4	9.8	9.8	42.0
$\chi^2$ Contribution	0.1	0.3	1.0	1.5
<b>Cluster 3</b>				
Observed frequency	7	8	3	18
Expected frequency	9.6	4.2	4.2	18.0
$\chi^2$ Contribution	0.7	3.4	0.3	4.5
<b>Cluster 4</b>				
Observed frequency	8	7	6	21
Expected frequency	11.2	4.9	4.9	21.0
$\chi^2$ Contribution	0.9	0.9	0.2	2.1

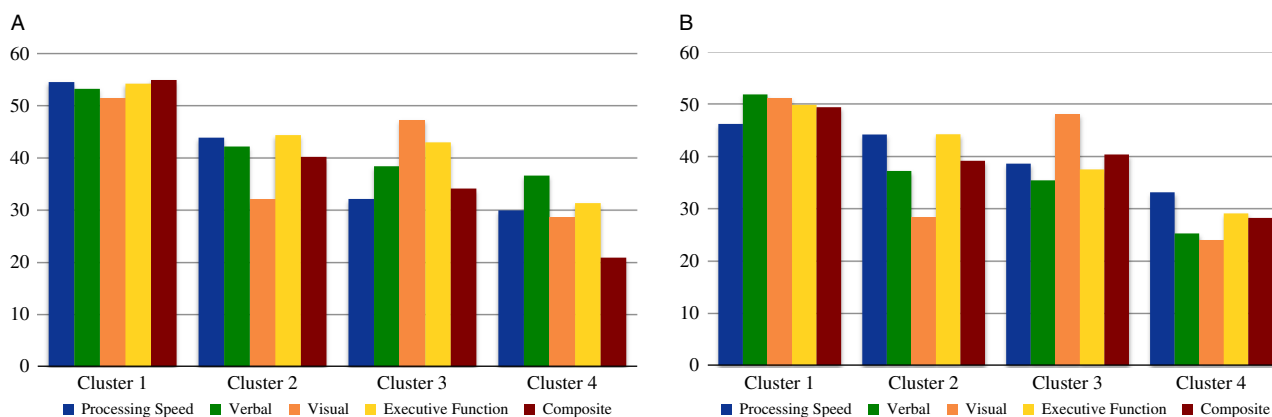
MDP, Mood Disorder with Psychosis; SZ, Schizophrenia; SZA, Schizoaffective Disorder

<sup>a</sup>Pearson  $\chi^2_{(6)} = 13.4$  ( $p < .05$ )

literature (MDP mean = 44.4; SZA mean = 38.0; SZ mean = 35.4; MDP > SZ) (Hill et al., 2013; Sperry et al., 2015). Multinomial logistic regression entering all MCCB domain scores as predictors of either cluster membership or diagnosis revealed that both models were statistically significant ( $p < .0001$ ); however, cognition explained considerably more variance in the cluster groupings (Pseudo  $R^2 = 0.74$ ) than in diagnosis (Pseudo  $R^2 = 0.22$ ).



**Fig. 2.** Community structure of cognition in psychotic illness and relation to diagnostic classification. Each panel shows a spring-loaded graph of the study participants based on the similarity (inverse distance) between each participant in a nine-dimensional space defined by their scores on nine cognitive tasks, with location of patients held constant in the upper and lower panels. The upper panel (A) shows patients only ( $n = 120$ ) with colors assigned based on cognitive cluster assignment, with more neuropsychologically intact individuals on the upper left (Cluster 1) and more globally impaired individuals on the lower right (Cluster 4). The lower panel (B) shows the same patients with colors assigned based on SCID-IV-TR diagnosis, and showing control participants in white ( $n = 31$ ). Asterisks and dotted lines provided for spatial reference between the two panels.



**Fig. 3.** New and previous cognitive profiles by cluster. Cognitive domain and composite scores from the present cluster analysis (Panel 3a) and our previous cluster analysis (Lewandowski et al., 2014; Panel 3b). Panel 3a: MCCB domain scores were combined to reflect the domains available from our previous analysis (Panel 3b) by averaging the Working Memory and Problem Solving Domains into an “Executive Functions” domain. Panel 3b: previous data (Lewandowski et al., 2014) were combined into domains comparable to MCCB domains for more direct comparison between the original data set and the present data. Domains included Processing Speed (Stroop, Trails A, and Category Fluency), Verbal (HVL T Total Recall), Visual (BVM T Total Recall), Executive Functions (Stroop Interference, Trails B), and a Composite (average of all cognitive testing scores). Previously-reported Z-scores were converted to T-scores for ease of comparison.

### Reproducibility of Cognitive Clusters

Figure 3 depicts cognitive profiles by cluster from our previous study and the current data presenting the modified domain scores described above. To compare the stability of the cluster assignments, the previous and new data were combined and K-means cluster analysis conducted. Not surprisingly, the new cluster profiles mirrored those of the two original samples. Original cluster assignments were then compared to the newly generated clusters. The proportion of the sample that was re-classified into the same cluster as their original cluster assignment was highest for Cluster 1 (93%), followed by Cluster 4 (89%), Cluster 3 (82%), and Cluster 2 (64%).

### DISCUSSION

The present study examined cognitive profiles in a cross-diagnostic sample of patients with psychosis using the MCCB, and the validity and interpretability of cluster analysis-derived cognitive groupings as compared to our previous report. We found support for a four-cluster solution with cognitive profiles that map closely to those produced in our previous work. As in our previous study (and others), we found evidence for a cognitively intact cluster that did not differ from our control sample or normative means, and a globally impaired cluster. We also found two clusters with mixed cognitive profiles: a cluster characterized by intact social cognition, mild impairments in processing speed, attention, verbal learning, and executive functions, and moderate selective impairment in visual learning (Cluster 2), and a cluster characterized by intact visual learning, mild impairment in executive functions and social cognition, and moderate impairments in processing speed and attention (Cluster 3). These findings are strikingly similar to the

profiles from our previous report (Figure 3), with a high degree of classification stability from the original cluster assignment to the new cluster assignment based on the combined sample.

The finding of an intact patient group and a globally impaired patient group is among the most commonly-reported findings using cluster analytic techniques (Heinrichs & Awad, 1993; Lewandowski et al., 2014; Palmer et al., 1997; Van Rheenen et al., 2017). In terms of globally impaired patients, widespread cognitive impairment has been described in schizophrenia for over a century. The term “dementia praecox” reflects the centrality of global cognitive impairment to early conceptualizations of schizophrenia (Bleuler, 1911; Kraepelin, 1919), and classifications such as “dementia subtype” (Heinrichs & Awad, 1993) have been used specifically for cluster-derived groups of patients with SZ with widespread severe cognitive disturbance. This cognitive profile may be reflective of some of the more impressive neurobiological findings in the literature, including widespread reductions in gray matter volume, enlarged ventricles, and evidence of early developmental brain abnormalities (Keefe et al., 1987; Woodward & Heckers, 2015).

Findings of “neuropsychologically normal” patients with psychosis have been intriguing, suggesting that some patients are relatively spared the cognitive effects of illness and subsequent poor outcomes (although it should be noted that patients with intact cognition may also have experienced neurodegenerative effects, but starting from a higher baseline). As would be hypothesized, several recent reports suggest that patients with schizophrenia without cognitive impairment evidence fewer structural brain abnormalities than patients with cognitive impairment (e.g., Woodward & Heckers, 2015).

It is also well-recognized that some patients exhibit more selective impairments both in terms of cognition and

neurobiological abnormalities. Cognitive profiles in Clusters 2 and 3 may reflect selective abnormalities, which map to commonly identified neurobiological findings in psychosis. Cluster 2 showed specific, selective impairment in visuospatial learning and memory, which may be associated with hippocampal abnormalities commonly described in patients with psychosis (e.g., Lawrie & Abukmeil, 1998; Nelson, Saykin, Flashman, & Riordan, 1998). Indeed, a recent report found that BVMT performance was associated with hippocampal volume (Bonner-Jackson, Mahmoud, Miller, & Banks, 2015) in a memory clinic population. Alternatively, poor performance on this task may be related to visual perception impairments, which are also well described in SZ (Butler, Silverstein, & Dakin, 2008).

Cluster 3, on the other hand, showed more pronounced deficits in processing speed and attention, with milder impairment in executive functions and intact visuospatial processing. EEG and diffusion weighted imaging studies have found white matter abnormalities in fronto-striatal connectivity associated specifically with measures of processing speed and attention in psychosis (Kochunov et al., 2016; Morales-Munoz et al., 2017; Seitz et al., 2016), suggesting that Cluster 3 may reflect specific white matter abnormalities along these tracts. Use of cluster analysis may reveal groups of patients who share cognitive performance profiles reflective of underlying neurobiological abnormalities. Explicit tests of the specificity of these clusters to neurobiological findings should be undertaken.

Consistent with our previous findings, all diagnoses were represented in all clusters, although not equally. Diagnoses did not map well to the cognitive clusters (Figure 2), and *post hoc* analyses showed that, while diagnostic groups did differ by cognitive performance, clusters do not simply reflect cognitive differences amongst disorders. These findings support the conceptualization of a cognitive dimension that cuts across diagnostic boundaries, although diagnostic groups may be differentially affected. Also consistent with our previous report, clinical symptom severity was worst in the globally impaired cluster; however, linear regression analyses suggested that differences in cluster membership were not best accounted for by clinical severity.

All four patient clusters were functioning more poorly than controls. The neuropsychologically normal cluster showed better functioning than the three neuropsychologically impaired clusters, consistent with our previous work. However, we previously reported no differences in community functioning amongst the clusters characterized by cognitive impairment, whereas the present findings suggest a “step-wise” association, with the globally impaired cluster showing the most pronounced functional deficits. Again, linear regressions suggested that differences in community outcomes were not better accounted for by illness severity or educational attainment. The present findings suggest that increasing cognitive burden is associated with poorer community functioning.

Social cognition did not follow the pattern of overall neurocognitive performance by cluster that might be

expected based on MCCB Composite scores. In fact, only Clusters 3 and 4 differed significantly from controls, and within patient groups only Clusters 1 and 4 differed significantly from each other. These findings suggest that social and neurocognition may rely on at least partially separate mechanisms.

Lack of reproducibility of findings and concerning rates of replication failure (Open Science Collaboration, 2015; Pashler & Harris, 2012) speak to the need for studies examining the reproducibility of findings. This may be particularly true in data-driven approaches, as relatively small samples and techniques that are highly influenced by study characteristics may generate findings that do not generalize outside of the details of the study design. Thus, this “conceptual replication” (Stroebe & Strack, 2014) was undertaken both to examine the reproducibility of our findings and to attempt to overcome an inherent limitation of cluster analysis techniques. We contend that multiple studies of reproducibility are critical to our ability to evaluate true effects (Maxwell, Lau, & Howard, 2015), and that conceptual replication offers the advantage of emphasizing the underlying constructs rather than simply demonstrating reproducibility of findings bound by identical operationalization of the original variables (Stroebe & Strack, 2014).

The present study has several limitations. First, data were assembled from several separate but related studies, with slightly different study criteria. Additionally, because of the distribution of participants amongst the clusters, groups were not equal in size, with our smallest cluster containing fewer than 20 participants. We were, therefore, limited in terms of power. However, our findings support our hypotheses even with conservative multiple comparisons corrections, supporting continued exploration of data-driven grouping strategies and their associations with clinical and functional correlates in psychosis.

Additionally, while we were able to include CPZ equivalents as a covariate in our analyses, we did not have adequate data on lithium or other mood stabilizers to allow investigation of the potential effects of these medications. The cross-sectional nature of these data limits our ability to draw conclusions about the meaning of these cognitive profiles in terms of the course of illness or prognosis. Longitudinal studies are needed to assess these important questions. Lastly, general cognitive functioning based on the NAART suggested that our sample was higher-functioning than would be expected in patients with psychosis. Thus, these findings may not generalize to a more typical sample of patients with psychosis.

In summary, we found groups with remarkably similar cognitive profiles using cluster analysis, suggesting that this technique produces reliable groupings across diagnoses that are not secondary to unique sampling characteristics or the measures used to derive them. Clusters accounted for considerably more variance in cognition than did diagnosis, suggesting that grouping patients along dimensions of interest may yield greater group homogeneity, which may be more likely to map to specific clinical trajectories or neurobiological mechanisms (Clementz et al., 2016). Validation of

our cluster solution based on reproducibility and associations with clinical and community functioning variables supports the interpretation of these groups as meaningfully distinct, reflective of underlying pathophysiology, and predictive of key outcomes. Data-driven grouping approaches may be particularly effective in facilitating the study of neurobiological mechanisms or etiological processes, and for building predictive models.

## ACKNOWLEDGMENTS

All authors K.E.L., J.T.B., J.M.M., L.A.N., and D.O. declare that they have no conflicts of interest to report in relation to this work. This work was supported by the National Institute of Mental Health (K.E.L., grant number K23MH91210), (D.O., grant number K24MH104449), (J.T.B., grant number K23MH104515); and the National Institute on Drug Abuse (J.M.M., grant number T32DA015036-15).

## REFERENCES

- Allen, D.N., Goldstein, G., & Warnick, E. (2003). A consideration of neuropsychologically normal schizophrenia. *Journal of the International Neuropsychological Society*, 9(1), 56–63. Retrieved from [https://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=12570358](https://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12570358)
- Baldessarini, R.J. (2012). *Chemotherapy in psychiatry*. New York: Springer Verlag.
- Barker, S., Barron, N., McFarland, B.H., Bigelow, D.A., & Carnahan, T. (1994). A community ability scale for chronically mentally ill consumers: Part II. Applications. *Community Mental Health Journal*, 30(5), 459–472. Retrieved from [https://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=7851100](https://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=7851100)
- Bleuler, E. (1911). *Dementia praecox or the group of schizophrenias*. Oxford: International Universities Press.
- Bonner-Jackson, A., Mahmoud, S., Miller, J., & Banks, S.J. (2015). Verbal and non-verbal memory and hippocampal volumes in a memory clinic population. *Alzheimer's Research and Therapy*, 7, 61. doi:10.1186/s13195-015-0147-9
- Burdick, K.E., Russo, M., Frangou, S., Mahon, K., Braga, R.J., Shanahan, M., ... Malhotra, A.K. (2014). Empirical evidence for discrete neurocognitive subgroups in bipolar disorder: Clinical implications. *Psychological Medicine*, 44(14), 3083–3096. doi:10.1017/S0033291714000439
- Butler, P.D., Silverstein, S.M., & Dakin, S.C. (2008). Visual perception and its impairment in schizophrenia. *Biological Psychiatry*, 64(1), 40–47. doi:10.1016/j.biopsych.2008.03.023
- Clementz, B.A., Sweeney, J.A., Hamm, J.P., Ivleva, E.I., Ethridge, L.E., Pearlson, G.D., ... Tamminga, C.A. (2016). Identification of distinct psychosis biotypes using brain-based biomarkers. *American Journal of Psychiatry*, 173(4), 373–384. doi:10.1176/appi.ajp.2015.14091200
- Goldstein, G. (1990). Neuropsychological heterogeneity in schizophrenia: A consideration of abstraction and problem-solving abilities. *Archives of Clinical Neuropsychology*, 5(3), 251–264. Retrieved from [https://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=14589685](https://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=14589685)
- Goldstein, G., Allen, D.N., & Seaton, B.E. (1998). A comparison of clustering solutions for cognitive heterogeneity in schizophrenia. *Journal of the International Neuropsychological Society*, 4, 353–362. Retrieved from [https://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=9656609](https://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=9656609)
- Goldstein, G., Beers, S.R., & Shemansky, W.J. (1996). Neuropsychological differences between schizophrenic patients with heterogeneous Wisconsin Card Sorting Test performance. *Schizophrenia Research*, 21(1), 13–18. Retrieved from [https://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=8998271](https://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=8998271)
- Goldstein, G., & Shemansky, W.J. (1995). Influences on cognitive heterogeneity in schizophrenia. *Schizophrenia Research*, 18(1), 59–69. Retrieved from [https://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=8929762](https://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=8929762)
- Goldstein, G., Shemansky, W.J., & Allen, D.N. (2005). Cognitive function in schizoaffective disorder and clinical subtypes of schizophrenia. *Archives of Clinical Neuropsychology*, 20(2), 153–159. doi:10.1016/j.acn.2004.03.008
- Heinrichs, R.W., & Awad, A.G. (1993). Neurocognitive subtypes of chronic schizophrenia. *Schizophrenia Research*, 9(1), 49–58. Retrieved from [https://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=8096391](https://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=8096391)
- Heinrichs, R.W., Ruttan, L., Zakzanis, K.K., & Case, D. (1997). Parsing schizophrenia with neurocognitive tests: Evidence of stability and validity. *Brain and Cognition*, 35(2), 207–224. doi:10.1006/brcg.1997.0938
- Hill, S.K., Ragland, J.D., Gur, R.C., & Gur, R.E. (2002). Neuropsychological profiles delineate distinct profiles of schizophrenia, an interaction between memory and executive function, and uneven distribution of clinical subtypes. *Journal of Clinical and Experimental Neuropsychology*, 24(6), 765–780. doi:10.1076/jcen.24.6.765.8402
- Hill, S.K., Reilly, J.L., Keefe, R.S., Gold, J.M., Bishop, J.R., Gershon, E.S., ... Sweeney, J.A. (2013). Neuropsychological impairments in schizophrenia and psychotic bipolar disorder: Findings from the Bipolar-Schizophrenia Network on Intermediate Phenotypes (B-SNIP) study. *American Journal of Psychiatry*, 170(11), 1275–1284. doi:10.1176/appi.ajp.2013.12101298
- Kay, S.R., Fiszbein, A., & Opler, L.A. (1987). The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophrenia Bulletin*, 13(2), 261–276. Retrieved from [https://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=3616518](https://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=3616518)
- Keefe, R.S., Mohs, R.C., Losonczy, M.F., Davidson, M., Silverman, J.M., Kendler, K.S., ... Davis, K.L. (1987). Characteristics of very poor outcome schizophrenia. *American Journal of Psychiatry*, 144(7), 889–895.
- Kochunov, P., Rowland, L.M., Fieremans, E., Veraart, J., Jahanshad, N., Eskandar, G., ... Hong, L.E. (2016). Diffusion-weighted imaging uncovers likely sources of processing-speed deficits in schizophrenia. *Proceedings of the National Academy of Sciences of the United States of America*, 113(47), 13504–13509.
- Kraepelin, E. (1919). *Dementia praecox and paraphrenia*. Edinburgh: E & S Livingstone.
- Lawrie, S.M., & Abukmeil, S.S. (1998). Brain abnormality in schizophrenia. A systematic and quantitative review of volumetric magnetic resonance imaging studies. *British Journal of Psychiatry*, 172, 110–120.
- Lewandowski, K.E., Cohen, B.M., Keshavan, M.S., Sperry, S.H., & Ongür, D. (2013). Neuropsychological functioning predicts community outcomes in affective and non-affective psychoses: A



- 6-month follow-up. *Schizophrenia Research*, 148(1-3), 34–37. doi:10.1016/j.schres.2013.05.012
- Lewandowski, K.E., Sperry, S.H., Cohen, B.M., & Ongür, D. (2014). Cognitive variability in psychotic disorders: A cross-diagnostic cluster analysis. *Psychological Medicine*, 44(15), 3239–3248. doi:10.1017/S0033291714000774
- Maxwell, S.E., Lau, M.Y., & Howard, G.S. (2015). Is psychology suffering from a replication crisis? What does “failure to replicate” really mean. *American Psychologist*, 70(6), 487–498. doi:10.1037/a0039400
- Montgomery, S.A., & Asberg, M. (1979). A new depression scale designed to be sensitive to change. *British Journal of Psychiatry*, 134, 382–389. Retrieved from [https://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=444788](https://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=444788)
- Morales-Muñoz, I., Jurado-Barba, R., Fernández-Guinea, S., Álvarez-Alonso, M.J., Rodríguez-Jiménez, R., Jiménez-Arriero, M.A., & Rubio, G. (2017). Cognitive impairments in patients with first episode psychosis: The relationship between neurophysiological and neuropsychological assessments. *Journal of Clinical Neuroscience*, 36, 80–87. doi:10.1016/j.jocn.2016.10.023
- Nelson, M.D., Saykin, A.J., Flashman, L.A., & Riordan, H.J. (1998). Hippocampal volume reduction in schizophrenia as assessed by magnetic resonance imaging: A meta-analytic study. *Archives of General Psychiatry*, 55(5), 433–440.
- Nuechterlein, K.H., Green, M.F., Kern, R.S., Baade, L.E., Barch, D.M., Cohen, J.D., ... Marder, S.R. (2008). The MATRICS Consensus Cognitive Battery, part 1: Test selection, reliability, and validity. *American Journal of Psychiatry*, 165(2), 203–213. doi:10.1176/appi.ajp.2007.07010042
- Open Science Collaboration. (2015). PSYCHOLOGY. Estimating the reproducibility of psychological science. *Science*, 349(6251), aac4716. doi:10.1126/science.aac4716
- Palmer, B.W., Heaton, R.K., Paulsen, J.S., Kuck, J., Braff, D., Harris, M.J., ... Jeste, D.V. (1997). Is it possible to be schizophrenic yet neuropsychologically normal. *Neuropsychology*, 11(3), 437–446. Retrieved from [https://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=9223148](https://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=9223148)
- Pashler, H., & Harris, C.R. (2012). Is the replicability crisis overblown? Three arguments examined. *Perspectives on Psychological Science*, 7(6), 531–536. doi:10.1177/1745691612463401
- Reser, M.P., Allott, K.A., Killackey, E., Farhall, J., & Cotton, S.M. (2015). Exploring cognitive heterogeneity in first-episode psychosis: What cluster analysis can reveal. *Psychiatry Research*, 229(3), 819–827. doi:10.1016/j.psychres.2015.07.084
- Seaton, B.E., Allen, D.N., Goldstein, G., Kelley, M.E., & van Kammen, D.P. (1999). Relations between cognitive and symptom profile heterogeneity in schizophrenia. *Journal of Nervous and Mental Disease*, 187(7), 414–419. Retrieved from [https://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=10426461](https://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=10426461)
- Seaton, B.E., Goldstein, G., & Allen, D.N. (2001). Sources of heterogeneity in schizophrenia: The role of neuropsychological functioning. *Neuropsychology Review*, 11(1), 45–67. Retrieved from [https://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=11392562](https://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=11392562)
- Seitz, J., Zuo, J.X., Lyall, A.E., Makris, N., Kikinis, Z., Bouix, S., ... Kubicki, M. (2016). Tractography analysis of 5 white matter bundles and their clinical and cognitive correlates in early-course schizophrenia. *Schizophrenia Bulletin*, 42(3), 762–771. doi:10.1093/schbul/sbv171
- Simmons, J.P., Nelson, L.D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 22(11), 1359–1366. doi:10.1177/0956797611417632
- Sperry, S.H., O'Connor, L.K., Öngür, D., Cohen, B.M., Keshavan, M.S., & Lewandowski, K.E. (2015). Measuring cognition in bipolar disorder with psychosis using the MATRICS Consensus Cognitive Battery. *Journal of the International Neuropsychological Society*, 21(6), 468–472. doi:10.1017/S1355617715000442
- Stroebe, W., & Strack, F. (2014). The alleged crisis and the illusion of exact replication. *Perspectives on Psychological Science*, 9(1), 59–71. doi:10.1177/1745691613514450
- Uttl, B. (2002). North American Adult Reading Test: Age norms, reliability, and validity. *Journal of Clinical and Experimental Neuropsychology*, 24(8), 1123–1137. doi:10.1076/jcen.24.8.1123.8375
- Van Rheenen, T.E., Lewandowski, K.E., Tan, E.J., Ospina, L.H., Ongur, D., Neill, E., ... Burdick, K.E. (2017). Characterizing cognitive heterogeneity on the schizophrenia-bipolar disorder spectrum. *Psychological Medicine*, 28, 1–17. doi:10.1017/S0033291717000307
- Woodward, N.D., & Heckers, S. (2015). Brain structure in neuropsychologically defined subgroups of schizophrenia and psychotic bipolar disorder. *Schizophrenia Bulletin*, 41(6), 1349–1359. doi:10.1093/schbul/sbv048
- Young, R.C., Biggs, J.T., Ziegler, V.E., & Meyer, D.A. (1978). A rating scale for mania: Reliability, validity and sensitivity. *British Journal of Psychiatry*, 133, 429–435. Retrieved from [https://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list\\_uids=728692](https://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=728692)