Title: Exploring Heterogeneity on the Wisconsin Card Sorting Test in Schizophrenia Spectrum Disorders: A Cluster Analytical Investigation.

Authors: Sean P. Carruthers[a*], Caroline T. Gurvich[b], Denny Meyer[a,c], Australian Schizophrenia Research Bank[d], Chad Bousman[e,f,g], Ian P. Everall[e,g,h,i,j], Erica Neill[a,g,h], Christos Pantelis[e,g,h,j,k], Philip J. Sumner[a], Eric J. Tan[a,l], Elizabeth H. X. Thomas[b], Tamsyn E. Van Rheenen[a,k], Susan L. Rossell[a,l]

Address:
[a]Centre for Mental Health, Faculty of Health, Arts and Design, Swinburne University of Technology, Victoria, 3122, Australia.
[b]Monash Alfred Psychiatry research centre (MAPrc), Monash University Central Clinical School and The Alfred Hospital, Melbourne 3004 Australia.
[c]Department of Statistics, Data Science and Epidemiology, Swinburne University of Technology, Victoria, 3122, Australia.
[d]Schizophrenia Research Institute, 405 Liverpool Street, Darlinghurst, Sydney, NSW 2010, Australia
[e]Cooperative Research Centre (CRC) for Mental Health, 161 Barry Street, Carlton, Victoria, 3053.
[f]Departments of Medical Genetics, Psychiatry, and Physiology & Pharmacology, University of Calgary, Calgary, AB, Canada.
[g]Department of Psychiatry, The University of Melbourne, Parkville, Victoria, 3052, Australia
[h]Florey Institute of Neuroscience and Mental Health, Parkville, Victoria 3052, Australia.
[i]Institute of Psychiatry, Psychology and Neuroscience, King's College London, UK.
[j]Centre for Neural Engineering, The University of Melbourne, Carlton, VIC, Australia
[k]Melbourne Neuropsychiatry Centre, Department of Psychiatry, The University of Melbourne and Melbourne Health, Carlton South, Victoria 3053, Australia.
[l]Psychiatry, St Vincent's Hospital, Melbourne, Victoria 3065, Australia.

* Corresponding Author: Sean Carruthers
scarruthers@swin.edu.au
Centre for Mental Health,
Swinburne University of Technology
Hawthorn VIC 3122
Australia

Word Count:
Abstract: 232
Text Body: 3748 (including in-text references and acknowledgements)

Abstract

Background. The Wisconsin Card Sorting Test (WCST) is a complex measure of executive function that is frequently employed to investigate the schizophrenia spectrum. Using traditional group-averaged comparisons, impaired performance on the WCST is commonly attributed to cognitive inflexibility amongst patients relative to healthy controls. It is now known that considerable cognitive heterogeneity exists amongst schizophrenia spectrum participants and there is substantive evidence of patient subgroups exhibiting distinct cognitive profiles. Critically, the within-group performance heterogeneity of schizophrenia spectrum patients on the WCST has yet to be investigated. A data-driven cluster analysis was performed on 210 schizophrenia spectrum outpatients to characterise WCST performance heterogeneity.

Method. Subjects with schizophrenia (*n*=172) or schizoaffective disorder (*n*=38) were assessed using the WCST. Hierarchical cluster analysis (Wards Method, squared Euclidean distance), with *k*-means optimization was employed to identify homogenous subgroups. Emergent clusters were then compared to each other and a group of 194 healthy controls on WCST performance and demographic/clinical variables.

Results. Three clusters emerged and were validated via altered design iterations. Clusters were deemed to reflect a relatively intact patient subgroup, a moderately impaired patient subgroup and a severely impaired patient subgroup. Clusters differed on several demographic/clinical variables.

Conclusions. Considerable within-group heterogeneity exists on the WCST. Identification of subgroups of patients who exhibit homogenous performance on measures of executive functioning may assist in optimizing cognitive interventions. Previous associations found using the WCST amongst schizophrenia spectrum participants should be reappraised.

Executive functions (EF) have been conceptualized as a set of high-level control processes that enable an individual to adapt to diverse situations, inhibit inappropriate responses, formulate, initiate and persevere with plans, and mediate the organisation of goal-directed thoughts and actions (Jurado and Rosselli, 2007, Miyake and Friedman, 2012). While no single neuropsychological measure currently exists that can provide a holistic quantitative account of the complex composition of EF, the Wisconsin Card Sorting Test (WCST) is one of the most commonly administered tasks employed to investigate EF in neuropsychologically impaired populations, particularly in the schizophrenia spectrum. The card-based attentional set-shifting paradigm provides metrics of a respondent's ability to develop, maintain and update a response strategy, whilst inhibiting irrelevant or inappropriate responses. This makes it a versatile and popular measure of EF. Deficits on key WCST outcome variables that proxy poor concept formation and cognitive inflexibility have been consistently reported amongst people with a schizophrenia spectrum diagnosis (Li, 2004, Polgár *et al.*, 2010, Waford and Lewine, 2010). In turn, executive dysfunction, as quantified through impaired WCST performance, has been associated with a broad range of constructs in schizophrenia spectrum research, including symptom dimensions (Donohoe and Robertson, 2003, Nieuwenstein *et al.*, 2001, Polgár *et al.*, 2010), functional outcome and learning proficiency (Green *et al.*, 2000, Lysaker *et al.*, 1995, Rempfer *et al.*, 2006, Rempfer *et al.*, 2017), insight and theory of mind (Croca *et al.*, 2018, Rossell *et al.*, 2003), genetic variability (Rybakowski *et al.*, 2006, Scarr *et al.*, 2012) and cortical activation and structure (Sasabayashi *et al.*, 2017, Wilmsmeier *et al.*, 2010).

Investigations using the WCST to study EF in schizophrenia spectrum disorders typically employ the traditional group-averaged comparison approach, whereby inferences are made from the results of contrasting patient performance against that of healthy controls

(HC) or another clinical group of interest (Kim *et al.*, 2014, Li, 2004, Rady *et al.*, 2011, Rossell and David, 1997). However, such group-averaged comparisons fail to account for the considerable within-group cognitive heterogeneity that exists within the schizophrenia spectrum and other psychotic disorders (Bora, 2016, Joyce and Roiser, 2007, Seaton *et al.*, 2001). Several studies utilising neuropsychological batteries have attempted to characterise this within-group variability using a data-driven, cluster-analytic approach (Lewandowski *et al.*, 2018, Lewandowski *et al.*, 2014, Van Rheenen *et al.*, 2017). Through identifying distinct clusters of participants who share a homogenous cognitive profile, several subgroups have been consistently identified that characterise the within-group cognitive variability prevalent amongst people with psychosis. Broadly, two anchoring cognitive profiles repeatedly emerge, a severely impaired subgroup and a cognitively intact subgroup, with a varying number of intermediate profiles of mixed/specific cognitive deficits emerging in-between (Burdick *et al.*, 2014, Lewandowski *et al.*, 2018, Van Rheenen *et al.*, 2017, Weickert *et al.*, 2000, Wells *et al.*, 2015). These subgroups can be identified at first episode (Reser *et al.*, 2015, Uren *et al.*, 2017), transcend the familial and psychosis spectrum (Burdick *et al.*, 2014, Hoti *et al.*, 2004, Lewandowski *et al.*, 2014, Van Rheenen *et al.*, 2017), exhibit distinctive brain structure (Czepielewski *et al.*, 2017, Van Rheenen *et al.*, 2018, Woodward and Heckers, 2015) and display altered response to treatment and functional outcomes (Gilbert *et al.*, 2014, Uren *et al.*, 2017). Together, this research indicates that whilst most schizophrenia spectrum patients experience mild-to-severe cognitive dysfunction, a subset perform at similar-to-HC levels.

To date, investigations using cluster analytical techniques to characterise the cognitive heterogeneity within the schizophrenia spectrum have typically been restricted to batteries of general cognition. The extent of the within-group variability on a solitary complex measure of EF has yet to be examined. In a recent cross-diagnostic latent class analysis of social

cognitive abilities in schizophrenia and bipolar disorder, two subgroups emerged exhibiting profiles of executive dysfunction in contrast to a HC group, neuropsychologically normal patient subgroup and a selective theory-of-mind deficit patient subgroup (Bora *et al.*, 2016). Using a cognitive battery that included a number of simple EF measures, Lewandowski and colleagues reported four emergent cognitive clusters with varying levels of executive dysfunction (Lewandowski *et al.*, 2018, Lewandowski *et al.*, 2014). Two anchoring clusters were identified to represent neuropsychologically normal and global cognitive impairment patient subgroups, with two intermediate clusters exhibiting intact and impaired EF amongst a combination of other cognitive impairments. Furthermore, several studies have utilized combinations of EF measures to characterise emergent cognitive subgroups post-cluster analysis and reported that subgroups differed in their executive abilities (Gilbert *et al.*, 2014, Weickert *et al.*, 2000). It is therefore becoming evident that considerable executive heterogeneity within the psychosis spectrum exists.

Due to the multidimensional composition of EF, additional examination of the cognitive heterogeneity within the schizophrenia spectrum is required. Successful completion of the more complex EF measures relies on multiple interdependent cognitive processes; therefore, it is likely that considerable variability exists amongst patients at the task level. Consequently, a cluster analytic approach was employed here to highlight the influential effects of EF heterogeneity in the schizophrenia spectrum, through characterising the within-group variability of outpatient performance on the WCST. The aim of the present study was to identify performance subgroups on the WCST using cluster analysis. It was hypothesised that two anchoring homogenous clusters of intact and severely impaired ability would emerge. An exploratory aim was to identify and characterise any additional clusters that emerge.

Method

Participants

Data from 210 patients with a diagnosis of schizophrenia or schizoaffective disorder and 192 HCs were obtained from the Cognitive and Genetic Explanations of Mental Illnesses (CAGEMIS) and Cooperative Research Centre (CRC) for Mental Health bio-databanks (Table 1). All participants had given prior informed consent for the analysis of their stored data and were recruited from metropolitan-based outpatient and community clinics. Psychiatric diagnosis and HC eligibility was confirmed using the MINI-International Neuropsychiatric Interview (Hergueta *et al.*, 1998). At time of testing, all patients were on stable doses of anti-psychotic medication. Patient symptomology was assessed at time of testing with either the Brief Psychiatric Rating Scale (BPRS; Overall and Gorham, 1962) or the Positive and Negative Syndrome Scale (PANSS; Kay *et al.*, 1987). As the use of the BPRS and PANSS was mixed during data collection, a consistent metric of symptomology across all participants was created by using the PANSS items that are encompassed within the BPRS for those participants without BPRS scores (see Supplementary Table 1). These scales have been shown to have similar scoring formats, display good correspondence in terms of treatment response (Leucht *et al.*, 2006, Leucht *et al.*, 2013) and have previously been combined in similar research studies (Van Rheenen *et al.*, 2017). Items were combined and averaged to reflect the following three BPRS symptom subscales; affect (anxiety, guilt, depression, somatic concern), positive (unusual thought content, conceptual disorganisation, hallucinatory behaviour, grandiosity) and negative (blunted affect, emotional withdrawal, motor retardation; Shafer, 2005). All participants were fluent in English, between the ages of 18 and 65 years old, and had an estimated premorbid intelligence quotient (IQ) > 70 as scored by either the Wechsler Test of Adult Reading (Wechsler, 2001) or the National Adult Reading Test (Blair and Spreen, 1989). Participants with significant visual or verbal

impairments, a known neurological disorder and/or current substance/alcohol abuse or dependence were excluded.

The Wisconsin Card Sorting Test

Each participant completed the 128-card version of the WCST, which was administered and scored in standard fashion (Heaton, 1993). Variables selected for cluster analysis included trials administered (TA), categories achieved (CAT), perseverative errors (PE), non-perseverative errors (NPE) and trials to first category (TFC). Total Correct (TC) and failure to maintain set (FMS) were used to characterise the subgroups emerging from the cluster analysis. This combination of variables was selected as to provide a holistic summary of WCST task performance. TA, CAT, and TFC provide an indication of overall task success and efficiency. PE reflect the extent of perseveration, with higher scores indicative of greater cognitive inflexibility. Higher NPE indicates impaired task performance not attributable to cognitive inflexibility.

Statistical Analysis

Several hierarchical cluster analyses were conducted using TA, CAT, NPE, PE and TFC Z-scores (based on the HC mean and *SD*) to identify and validate homogenous performance subgroups within our sample. Similarity between schizophrenia spectrum cases was computed using hierarchical agglomerative clustering, with Ward's minimum variance method and Squared Euclidean distance. Collaborative inspection of the agglomeration schedule and dendrogram was used to establish the appropriate number of clusters to retain and confirmed by discriminant function analysis. A *k*-means iterative partitioning technique was then employed to optimise the retained clusters, with initial partitions in the *k*-means solution defined using the cluster means obtained from the initial clustering procedure. The

stability of the final cluster solution was evaluated through split-sample and alternate method replication via Cohen's κ analysis, with high agreement over multiple design iterations required to validate the final clustering solution obtained ($\kappa > 0.80$; Landis and Koch, 1977).

Emergent clusters were compared on premorbid IQ, demographic and clinical variables, and WCST variables using analysis of variance (ANOVA) or Chi-square (χ^2) analyses as appropriate. Brown-Forsythe F-ratio was used when appropriate. *Post-hoc p*-values were Games-Howell corrected for unequal sample sizes/unequal variances. Eta-Squared (η^2) was calculated for omnibus tests (large effect, $\eta^2 > 0.26$), Cohen's *d* for *post-hoc* pairwise comparisons (large effect, $d > 0.8$), and Cramer's *V* for χ^2 analyses (large effect, $\chi^2 > 0.3$), as measures of effect size. For significant χ^2 results, adjusted standardised residuals and their calculated significance were reported.

Results

Cluster analysis

Three clusters emerged representing three distinct WCST performance profiles (see Table 2). Table 3 displays the membership agreement and κ scores between the final clustering solution and four alternate design iterations. An almost perfect level of agreement (Landis and Koch, 1977) was detected between the final clustering solution and the solutions emerging from a random 75%, 50% and 25% subset of the original patient sample using the same method, as well as the solution that emerged using the average linkage method on the whole patient sample.

Cluster one outperformed the remaining clusters on all WCST variables except for FMS (see below). Compared to the HC group, Cluster One exhibited a slightly less efficient, albeit successful performance on the WCST, scoring more TC, NPE and FMS and therefore requiring additional TA to reach the same number of categories completed as the HC group.

Cluster One was therefore labelled as the relatively intact subgroup. Cluster Two was deemed to represent a moderate level of impairment, distinguished by control equivalent performance on TC, in the absence of successfully achieving a high number of categories. This patient subgroup appeared to grasp the conceptual aspects of the task to permit the successful completion of some categories, however failed to maintain set to enable successful completion of the WCST. Despite requiring a relatively small number of trials to reach the first category, the performance of Cluster Two was impaired by a moderate number of both NPE and PE. Taken together, the moderate impairment subgroup displayed a profile of inefficient, unsuccessful, yet relatively flexible performance on the WCST. The third and final cluster, labelled as the severe impairment subgroup, was distinguished by a high TFC, PE, combined with an impaired profile of high TA, NPE and low TC, CAT. The severe impairment subgroup was significantly impaired on all variables compared to both patient subgroups and HCs except for FMS. The non-differentiated performance on FMS was interpreted to reflect an inability to acquire a correct response set and therefore maintain it, as opposed to the inability to maintain an acquired set.

Non-WCST Results

Table 4 summarises the demographic characteristics of the three emergent patient subgroups and HCs. Subgroups were not significantly different in illness duration, age of symptom onset, CPZ equivalent, BPRS affect and BPRS positive. The severe impairment subgroup was detected to have significantly higher BPRS negative scores compared to the remaining patient subgroups, with the relatively intact subgroup detected to have significantly higher estimated premorbid IQ compared to the moderate and severe impairment subgroups. A significant main effect of age was detected, however *post-hoc* comparisons revealed no significant pairwise differences. A significant difference in the

distribution of sex across all groups was detected, with a significant difference in the distribution of schizophrenia spectrum diagnosis also detected across the three patient subgroups. Follow up examination of any sex-by-group or schizophrenia spectrum diagnosis-by-group interactions was performed. Several significant main effects of sex and schizophrenia spectrum diagnosis were returned across the WCST variables; with females and patient participants with a diagnosis of schizoaffective disorder displaying a better overall performance compared to their male and schizophrenia counterparts. However, no significant sex-by-group or schizophrenia spectrum diagnosis-by-subgroup interactions were detected (see Supplementary Tables 2 and 3).

Discussion

The present study sought to characterise cognitive heterogeneity within the schizophrenia spectrum by examining the performance variability of a group if outpatients on a complex and popular measure of EF, the WCST. Using a data-driven cluster analysis optimised with *k*-means iterative partitioning, we were able to detect three homogenous subgroups of schizophrenia spectrum patients that exhibited distinct performance profiles. Of the three emergent clusters, a patient subgroup exhibiting a relatively intact level of performance on the WCST was identified, in addition to two subgroups of impairment; a moderate impairment subgroup and a severe impairment subgroup.

Impaired performance on the WCST is frequently reported in the schizophrenia spectrum literature and is most commonly attributed to cognitive inflexibility (Li, 2004). Contrary to this, we were able to identify a subgroup of participants who exhibited a control-equivalent level of overall performance. The relatively intact patient subgroup demonstrated cognitive flexibility and the ability to form and maintain concepts, enabling the majority of the subgroup to successfully complete the WCST, albeit less efficiently than the HC group.

All relatively intact subgroup participants achieved five or more categories (one category = 10 successive correct sorts), with 83% (n = 60) reaching the six categories achieved task discontinuation criterion to successfully complete the WCST. This contrasts with the compromised performances of the moderate and severe impairment subgroups, in which no participant achieved more than four or two categories respectively (see Supplementary Table 4).

The severe impairment subgroup represented an amplified version of what is typically reported within the available literature, a profile of poor concept formation paired with cognitive inflexibility. Distinguished by a high number of trials to reach the first category and a high number of perseverative errors, participants in this subgroup failed to grasp the requirements of the task. Members of the severe impairment subgroup displayed an inability to form and maintain the concepts required to successfully complete multiple, if any categories despite ongoing negative feedback. In contrast, the profile of the moderate impairment subgroup is viewed to reflect an unsuccessful attempt at adapting to the demands of the task. The moderate impairment subgroup exhibited a degree of cognitive flexibility and relative success, requiring only a modest number of trials to reach the first category and correctly sorting more trials and achieving more categories than the severely impaired subgroup. However, when compared against the performances of the relatively intact patient subgroup and HC group, the moderately impaired participants were less effective when it came to concept formation and demonstrated a compromised ability to maintain set.

To our knowledge, this is the first time that cluster analysis has been applied to a solitary complex measure of EF to identify homogenous subgroups in a conventional research sample of schizophrenia spectrum participants. The emergence of three distinct subgroups on the WCST coincides with previous cluster analytic studies that have documented multiple cognitive phenotypes within the psychosis spectrum (Lewandowski *et*

al., 2018, Lewandowski *et al.*, 2014, Reser *et al.*, 2015, Van Rheenen *et al.*, 2017). Such studies have consistently identified two anchoring cognitive subgroups; one characterised by severe cognitive impairments and the other a control-equivalent or neuropsychologically normal level of functioning, with a varying number of profiles of mixed impairments typically emerging in-between. The present results extend this literature, identifying two subgroups that differ in severity of executive dysfunction and a third subgroup exhibiting a relatively intact level of functioning as quantified through the WCST. Relatively intact subgroups of schizophrenia spectrum participants have previously been characterised (Van Rheenen *et al.*, 2017), which exhibit an overarching profile of intact cognitive functioning, especially in comparison with the remaining patient subgroups that emerge, however present with one or two slight cognitive impairments relative to HCs. Considering that the WCST requires multiple interdependent cognitive processes to complete, it is not surprising that a group of schizophrenia spectrum participants identified as demonstrating an intact level of overall performance on the complex measure of EF were found to do so with a degree of inefficiency compared to HCs. Whilst the relatively intact and two impaired subgroups were found to differ in estimations of premorbid IQ, such a distinction was not detected between relatively intact patient and HC participants. This suggests that the subtle differences between the HC and relatively intact groups, and the differences in severity of impairment between the two remaining patient subgroups are not attributable to differences in premorbid functioning. Similarly, medication appears not to be a contributing factor in this instance, with no significant difference in chlorpromazine equivalence detected across patient subgroups.

The identification of three distinct profiles of cognitive flexibility in the present study, encourages the consideration of EF heterogeneity when developing, implementing and evaluating psychosocial and cognitive interventions (Bryce *et al.*, 2018, Cella *et al.*, 2014, Wykes *et al.*, 2007). Several studies report the functional benefits of interventions targeting

cognitive flexibility (Farreny *et al.*, 2012, Penadés *et al.*, 2010), which could likely be influenced by the within-group variability identified here. For example, patients exhibiting flexible, yet ineffective adaptions to the environment may benefit from interventions based in the real-world, such as Cognitive Adaptation Training, which utilises environmental supports and compensatory strategies to improve environmental adjustment and vocational recovery (Allott *et al.*, 2017, Fredrick *et al.*, 2015). Highly perseverative patients may benefit from an intervention combining both cognitive remediation and functional adaptation training (Bowie *et al.*, 2012). Alternatively, the effectiveness of interventions improving cognitive and functional outcomes in impaired patients may be negatively influenced by the inclusion of cognitively intact individuals. Such patients would likely show minimal improvements after standard interventions designed to improve impaired cognitive capacities, negatively impacting the outcome indices used to assess the effectiveness of the intervention. Limited in its cognitive scope in contrast to a multidimensional cognitive battery, the WCST, in combination with cluster analysis could be a useful and brief way to identify schizophrenia spectrum patients better suited to a particular intervention therapy.

The present study is not without its limitations. Whilst speculations were offered concerning the ecological and functional benefits of pairing cluster analysis with the WCST within the context of schizophrenia spectrum disorders, no empirically-based elaboration was possible. Considering the functional consequences associated with cognitive inflexibility (Green *et al.*, 2004, Velligan *et al.*, 2000), more elaborate and targeted examination of the association between EF heterogeneity and outcomes is needed to permit further insight into the real-world applicability of the present findings. Due to the present study's specific investigative focus, any implications are restricted to the EF of cognitive flexibility and concept formation. The field will therefore benefit from a latent variable approach, investigating EF heterogeneity within the schizophrenia spectrum using the fundamental EFs

of Updating, Shifting and Inhibition (Miyake and Friedman, 2012, Miyake *et al.*, 2000). Furthermore, the present study only had access to verbal learning based estimates of premorbid functioning and whilst no unexpected differences across subgroups were found, the use of a premorbid functioning estimate measure based on visuospatial ability would have been more appropriate given the WCST reliance on visuospatial reasoning. Whilst there has been a renewed interested in the use of data-driven cluster analytic approaches to characterise homogenous performance profiles in psychosis, limitations inherently lie in the statistical method that impact the validity and generalisability of the results of the present study. Emergent clusters are determined by the characteristics of the sample examined and by the measures employed. Therefore, despite establishing a high level of agreement between the final clustering solution and multiple design iterations, external replication is necessary to establish the validity and utility of the subgroups identified here (Lewandowski *et al.*, 2018).

In conclusion, cluster analysis was employed to investigate the within group performance variability on a popular measure of EF, the WCST in a sizeable schizophrenia spectrum sample. Three clusters emerged, reflecting relatively intact, moderately and severely impaired profiles of performance. The identification of three distinct subgroups highlights the importance of considering sample heterogeneity when administering complex measures of EF to schizophrenia spectrum participants. Interventions targeting EF and functional outcomes should consider the strengths and weaknesses of heterogeneous subgroups within their sample before commencing. Due to the distinct performance profiles identified here, the results promote the reappraisal of findings from previous group-based association studies using the WCST to investigate the schizophrenia spectrum.

Table 1. Demographic summary

	Schizophrenia Spectrum (n = 210)	Healthy Control (n = 192)	Test Statistic[a]
Age	39.1 (9.7)	36.0 (13.5)	$F_{1,276.8} = 11.5, p \leq .001$
% Male	75.2	54.9	$\chi^2 = 18.4, p \leq .001$
Estimated Premorbid IQ	104.9 (13.3)	110.2 (10.9)	$\mathbf{F_{1,385.4} = 18.6, p \leq .001}$
WCST			
TA	118.7 (17.4)	88.7 (18.7)	$\mathbf{F_{1,391.5} = 274.2, p \leq .001}$
TC	68.2 (15.4)	69.4 (6.7)	$F_{1,290.2} = 1.0, p = .31$
CAT	3.2 (2.1)	5.8 (0.8)	$\mathbf{F_{1,264.3} = 260.7, p \leq .001}$
NPE	21.7 (11.8)	9.0 (7.2)	$\mathbf{F_{1,350.4} = 174.4, p \leq .001}$
PE	28.8 (18.3)	10.5 (9.5)	$\mathbf{F_{1,319.6} = 163.1, p \leq .001}$
TFC	33.1 (36.5)	14.5 (9.1)	$F_{1,237.0} = 57.3, p \leq .001$
FMS	1.7 (2.0)	0.4 (0.7)	$\mathbf{F_{1,267.2} = 81.8, p \leq .001}$

Data is reported as mean (SD) unless otherwise stated; Wisconsin Card Sorting Test; TA, trials administered; TC, total correct; CAT, categories achieved; NPE, non-perseverative errors; PE, perseverative errors; TFC, trials to first category; FMS, failure to maintain set; [a] Brown-Forythe F-ratio reported; Bold font denotes large effect size. Estimated Premorbid IQ data unavailable for 10 participants.

Table 2. Comparison of WCST z-scores across emergent cluster subgroups.

	Cluster 1	Cluster 2	Cluster 3	HC	Test Statistic*	1 *v.* 2		1 *v.* 3		2 *v.* 3		1 *v.* HC		2 *v.* HC		3 *v.* HC	
	Relatively Intact (n = 72)	Moderate Impairment (n = 114)	Severe Impairment (n = 24)	(n=192)		*d*	*p*	*d*	*p*	*d*	*p*	*d*	*p*	*d*	*p*	*d*	*p*
WCST																	
TA	0.65 (1.1)	2.09 (0.0)	2.09 (0.0)	-0.01 (1.0)	$\mathbf{F_{3,152.2} = 245.9, p \leq .001}$	2.1	≤ .001	1.5	≤ .001	-	-	0.6	≤ .001	2.7	≤ .001	2.2	≤ .001
CAT	0.08 (0.5)	-4.66 (1.3)	-7.37 (0.6)	0.04 (0.9)	$\mathbf{F_{3,232.9} = 1138.8, p \leq .001}$	4.5	≤ .001	14.3	≤ .001	2.3	≤ .001	0.1	0.945	4.4	≤ .001	8.5	≤ .001
NPE	-0.36 (0.9)	2.39 (1.4)	3.03 (1.7)	-0.02 (1.0)	$\mathbf{F_{3,76.3} = 104.6, p \leq .001}$	2.3	≤ .001	3.0	≤ .001	0.4	0.327	0.4	≤ .05	2.1	≤ .001	2.8	≤ .001
PE	0.20 (0.6)	2.58 (1.6)	4.05 (1.9)	-0.01 (1.0)	$\mathbf{F_{3,69.6} = 123.0, p \leq .001}$	1.8	≤ .001	3.6	≤ .001	0.9	≤ .01	0.2	0.174	2.1	≤ .001	3.6	≤ .001
TFC	0.03 (0.6)	1.42 (2.2)	12.13 (1.3)	-0.04 (0.9)	$\mathbf{F_{3,157.1} = 585.7, p \leq .001}$	0.8	≤ .001	14.8	≤ .001	5.2	≤ .001	0.1	0.905	1.0	≤ .001	12.9	≤ .001
TC[#]	1.13 (1.7)	-0.38 (2.0)	-3.16 (2.2)	0.00 (1.0)	$\mathbf{F_{3,98.9} = 33.5, p \leq .001}$	0.8	≤ .001	2.4	≤ .001	1.4	≤ .001	0.9	≤ .001	0.3	0.240	2.7	≤ .001
FMS[#]	0.91 (1.9)	2.38 (2.7)	2.04 (3.7)	-0.03 (0.9)	$F_{3,64.5} = 20.3, p \leq .001$	0.6	≤ .001	0.5	0.485	0.1	0.972	0.8	≤ .001	1.4	≤ .001	1.4	≤ .05

Data is reported as mean (SD) unless otherwise stated; HC, healthy control; WCST, Wisconsin Card Sorting Test; TA, trials administered; CAT, categories achieved; NPE, non-perseverative errors; PE, perseverative errors; TFC, trials to first category; TC, total correct; FMS, failure to maintain set; [#] not included in cluster analysis; [*] Brown-Forythe F-ratio reported; Bold font denotes large effect size for omnibus tests.

Table 3. Agreement and Kappa coefficient scores between final clustering solution and alternate replications.

	Cluster 1	Cluster 2	Cluster3	Kappa coefficient [95% CI]
	Relatively Intact	Moderate Impairment	Severe Impairment	
Random 75% subset				
Cluster 1	55 (100%)	0	0	$\kappa = 0.94$ [0.91, 0.96], $p \leq .001$
Cluster 2	0	77 (92.8%)	0	
Cluster 3	0	6 (7.2%)	19 (100%)	
Random 50% subset				
Cluster 1	30 (100%)	1 (1.6%)	0	$\kappa = .98$ [0.96, 1.0], $p \leq .001$
Cluster 2	0	63 (98.4%)	0	
Cluster 3	0	0	11 (100%)	
Random 25% subset				
Cluster 1	19 (100%)	0	0	$\kappa = 0.85$ [0.79, 0.91], $p \leq .001$
Cluster 2	0	23 (82.1%)	0	
Cluster 3	0	5 (17.9%)	6 (100%)	
Average linkage solution[a]				
Cluster 1	72 (100%)	0	0	$\kappa = 1.00$ [1.0, 1.0], $p \leq .001$
Cluster 2	0	114 (100%)	0	
Cluster 3	0	0	24 (100%)	

Note: CI = confidence interval. [a] All remaining analysis steps are as per main analysis described in method.

Table 4. Demographic characteristics of emergent cluster subgroups.

	Cluster 1	Cluster 2	Cluster 3			
	Relatively Intact (n = 72)	Moderate Impairment (n = 114)	Severe Impairment (n = 24)	Test Statistic	Significant *post-hoc* testing	
Age	38.5 (9.7)	39.0 (9.1)	41.4 (12.3)	$F_{3,134.1} = 0.83, p = .43$		
Premorbid IQ	110.4 (10.5)	102.1 (14.0)	102.1 (12.5)	$F_{3,143.8} = 13.5, p \leq .001$[#]	1 *v.* 2: $p \leq .001$	1 *v.* 3: $p \leq .05$
% Male	75	79.8	54.2	$\chi^2 = 7.01, p \leq .05$	3_{ASR}: -2.5, $p \leq .01$	
% Schizophrenia	69.4	89.5	81.9	$\chi^2 = 12.0, p \leq .01$	1_{ASR}: 3.4, $p \leq .001$	2_{ASR}: 3.1, $p \leq .01$
% Unemployed/Not Studying	47.1	75.4	79.2	$\chi^2 = 17.55, p \leq .001$	1_{ASR}: -4.17, $p \leq .001$	2_{ASR}: 3.06, $p \leq .01$
Illness duration[a]	16.1 (9.0)	15.7 (9.4)	16.8 (9.8)	$F_{2,198} = 0.2, p = .86$		
Age of symptom onset	21.4 (6.9)	23.2 (7.0)	24.6 (9.9)	$F_{2,55.4} = 1.6, p = .22$[#]		
CPZ equivalent	561.4 (449.0)	569.6 (428.7)	490.9 (287.0)	$F_{2,196} = 0.3, p = .72$		
BPRS						
Affect	2.1 (1.1)	1.9 (1.0)	2.2 (1.1)	$F_{2,209} = 1.7, p = .19$		
Positive	2.2 (1.2)	2.3 (1.2)	2.9 (1.2)	$F_{2,209} = 2.7, p = .07$		
Negative	1.6 (0.8)	1.8 (1.0)	2.4 (0.8)	$F_{2,114.2} = 6.4, p \leq .01$[#]	**1 *v.* 3: $p \leq .001$**	2 *v.* 3: $p \leq .05$

Data is reported as mean (SD) unless otherwise stated; BRPS, Brief Psychotic Rating Scale; $_{ASR}$, adjusted standardised residuals; [a] Years total duration of active and residual periods; [#] Brown-Forythe F-ratio reported; Bold indicates large effect sizes. Illness duration data unavailable for 11 participants. Age of symptom onset data unavailable for 18 participants. CPZ data unavailable for 13 participants.

Acknowledgements

This work was supported by Australian Postgraduate Awards (S.C, P.S and E.T) and by the Australian National Health and Medical Research Council (NHMRC; fellowships to C.G (ID: 5467262), T.V.R (1088785) and C.P (628386 & 1105825); and a project grant to S.L.R (ID: 1060664)). The authors acknowledge the support of the Monash Alfred Psychiatry Research Centre and the financial support of the Cooperative Research Centre (CRC) for Mental Health. The CRC programme is an Australian Government Initiative. The authors wish to acknowledge the CRC Scientific Advisory Committee, in addition to the contributions of study participants, clinicians at recruitment services, staff at the Murdoch Children's Research Institute, staff at the Australian Imaging, Biomarkers and Lifestyle Flagship Study of Aging, and research staff at the Melbourne Neuropsychiatry Centre, including coordinators Phassouliotis, C., Merritt, A., and research assistants, Burnside, A., Cross, H., Gale, S., and Tahtalian, S. None of the funding sources played any role in the study design; collection, analysis or interpretation of data; in the writing of the report; or in the decision to submit the paper for publication. Participants for this study were sourced, in part, through the Australian Schizophrenia Research Bank (ASRB), which is supported by the National Health and Medical Research Council of Australia (Enabling Grant ID: 386500), the Pratt Foundation, Ramsay Health Care, the Viertel Charitable Foundation and the Schizophrenia Research Institute. We thank the Chief Investigators and ASRB Manager: Carr, V., Schall, U., Scott, R., Jablensky, A., Mowry, B., Michie, P., Catts, S., Henskens, F., Pantelis, C., Loughland, C. We acknowledge the help of Jason Bridge for ASRB database queries. The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

Conflicts of Interest

None.

References

Allott, K. A., Killackey, E., Sun, P., Brewer, W. J. & Velligan, D. I. (2017). Improving vocational outcomes in first-episode psychosis by addressing cognitive impairments using Cognitive Adaptation Training. *Work* **56**, 581-589.

Blair, J. R. & Spreen, O. (1989). Predicting premorbid IQ: a revision of the National Adult Reading Test. *The Clinical Neuropsychologist* **3**, 129-136.

Bora, E. (2016). Differences in cognitive impairment between schizophrenia and bipolar disorder: Considering the role of heterogeneity. *Psychiatry and Clinical Neurosciences* **70**, 424-433.

Bora, E., Veznedaroğlu, B. & Vahip, S. (2016). Theory of mind and executive functions in schizophrenia and bipolar disorder: A cross-diagnostic latent class analysis for identification of neuropsychological subtypes. *Schizophrenia Research* **176**, 500-505.

Bowie, C. R., McGurk, S. R., Mausbach, B., Patterson, T. L. & Harvey, P. D. (2012). Combined cognitive remediation and functional skills training for schizophrenia: Effects on cognition, functional competence, and real-world behavior. *American Journal of Psychiatry* **169**, 710-718.

Bryce, S. D., Rossell, S. L., Lee, S. J., Lawrence, R. J., Tan, E. J., Carruthers, S. P. & Ponsford, J. L. (2018). Neurocognitive and Self-efficacy Benefits of Cognitive Remediation in Schizophrenia: A Randomized Controlled Trial. *Journal of the International Neuropsychological Society* **23**, 1-14.

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**, 3083-3096.

Cella, M., Bishara, A. J., Medin, E., Swan, S., Reeder, C. & Wykes, T. (2014). Identifying cognitive remediation change through computational modelling - Effects on reinforcement learning in schizophrenia. *Schizophrenia Bulletin* **40**, 1422-1432.

Croca, M., Lagodka, A., Gadel, R., Bourdel, M. C., Bendjemaa, N., Gaillard, R., Olié, J. P., Champagne-Lavau, M., Krebs, M. O. & Amado, I. (2018). Theory of mind and schizophrenia in young and middle-aged patients: Influence of executive functions. *Psychiatry Research* **259**, 532-537.

Czepielewski, L. S., Wang, L., Gama, C. S. & Barch, D. M. (2017). The Relationship of Intellectual Functioning and Cognitive Performance to Brain Structure in Schizophrenia. *Schizophrenia bulletin* **43**, 355-364.

Donohoe, G. & Robertson, I. H. (2003). Can specific deficits in executive functioning explain the negative symptoms of schizophrenia? A review. *Neurocase* **9**, 97-108.

Farreny, A., Aguado, J., Ochoa, S., Huerta-Ramos, E., Marsà, F., López-Carrilero, R., Carral, V., Haro, J. M. & Usall, J. (2012). REPYFLEC cognitive remediation group training in schizophrenia: Looking for an integrative approach. *Schizophrenia Research* **142**, 137-144.

Fredrick, M. M., Mintz, J., Roberts, D. L., Maples, N. J., Sarkar, S., Li, X. & Velligan, D. I. (2015). Is cognitive adaptation training (CAT) compensatory, restorative, or both? *Schizophrenia Research* **166**, 290-296.

Gilbert, E., Mérette, C., Jomphe, V., Émond, C., Rouleau, N., Bouchard, R. H., Roy, M. A., Paccalet, T. & Maziade, M. (2014). Cluster analysis of cognitive deficits may

mark heterogeneity in schizophrenia in terms of outcome and response to treatment. *European Archives of Psychiatry and Clinical Neuroscience* **264**, 333-343.

Green, M. F., Kern, R. S., Braff, D. L. & Mintz, J. (2000). Neurocognitive deficits and functional outcome in schizophrenia: Are we measuring the 'right stuff'? *Schizophrenia Bulletin* **26**, 119-136.

Green, M. F., Kern, R. S. & Heaton, R. K. (2004). Longitudinal studies of cognition and functional outcome in schizophrenia: Implications for MATRICS. *Schizophrenia Research* **72**, 41-51.

Heaton, R. K. (1993). Wisconsin Card Sorting Test manual, revised and expanded. *Psychological Assessment Services, Inc*.

Hergueta, T., Baker, R. & Dunbar, G. C. (1998). The Mini-International Neuropsychiatric Interview (MINI): the development and validation of a structured diagnostic psychiatric interview for DSM-IVand ICD-10. *J clin psychiatry* **59**, 2233.

Hoti, F., Tuulio-Henriksson, A., Haukka, J., Partonen, T., Holmström, L. & Lönnqvist, J. (2004). Family-based clusters of cognitive test performance in familial schizophrenia. *BMC Psychiatry* **4**.

Joyce, E. M. & Roiser, J. P. (2007). Cognitive heterogeneity in schizophrenia. *Current Opinion in Psychiatry* **20**, 268-272.

Jurado, M. B. & Rosselli, M. (2007). The elusive nature of executive functions: A review of our current understanding. *Neuropsychology Review* **17**, 213-233.

Kay, S. R., Fiszbein, A. & Opler, L. A. (1987). The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophrenia Bulletin* **13**, 261-276.

Kim, H. S., An, Y. M., Kwon, J. S. & Shin, M. S. (2014). A preliminary validity study of the cambridge neuropsychological test automated battery for the assessment of

executive function in schizophrenia and bipolar disorder. *Psychiatry Investigation* **11**, 394-401.

Landis, J. R. & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *biometrics*, 159-174.

Leucht, S., Kane, J. M., Etschel, E., Kissling, W., Hamann, J. & Engel, R. R. (2006). Linking the PANSS, BPRS, and CGI: Clinical implications. *Neuropsychopharmacology* **31**, 2318-2325.

Leucht, S., Rothe, P., Davis, J. M. & Engel, R. R. (2013). Equipercentile linking of the BPRS and the PANSS. *European Neuropsychopharmacology* **23**, 956-959.

Lewandowski, K. E., Baker, J. T., McCarthy, J. M., Norris, L. A. & Öngür, D. (2018). Reproducibility of cognitive profiles in psychosis using cluster analysis. *Journal of the International Neuropsychological Society* **24**, 382-390.

Lewandowski, K. E., Sperry, S. H., Cohen, B. M. & Öngür, D. (2014). Cognitive variability in psychotic disorders: A cross-diagnostic cluster analysis. *Psychological Medicine* **44**, 3239-3248.

Li, C. S. R. (2004). Do schizophrenia patients make more perseverative than non-perseverative errors on the Wisconsin Card Sorting Test? A meta-analytic study. *Psychiatry Research* **129**, 179-190.

Lysaker, P., Bell, M. & Beam-Goulet, J. (1995). Wisconsin card sorting test and work performance in schizophrenia. *Psychiatry Research* **56**, 45-51.

Miyake, A. & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science* **21**, 8-14.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A. & Wager, T. D. (2000). The Unity and Diversity of Executive Functions and Their Contributions to

Complex "Frontal Lobe" Tasks: A Latent Variable Analysis. *Cognitive Psychology* **41**, 49-100.

Nieuwenstein, M. R., Aleman, A. & De Haan, E. H. F. (2001). Relationship between symptom dimensions and neurocognitive functioning in schizophrenia: A meta-analysis of WCST and CPT studies. *Journal of Psychiatric Research* **35**, 119-125.

Overall, J. E. & Gorham, D. R. (1962). The brief psychiatric rating scale. *Psychological reports* **10**, 799-812.

Penadés, R., Catalán, R., Puig, O., Masana, G., Pujol, N., Navarro, V., Guarch, J. & Gastó, C. (2010). Executive function needs to be targeted to improve social functioning with Cognitive Remediation Therapy (CRT) in schizophrenia. *Psychiatry Research* **177**, 41-45.

Polgár, P., Réthelyi, J. M., Bálint, S., Komlósi, S., Czobor, P. & Bitter, I. (2010). Executive function in deficit schizophrenia: What do the dimensions of the Wisconsin Card Sorting Test tell us? *Schizophrenia Research* **122**, 85-93.

Rady, A., Elsheshai, A., el Wafa, H. A. & Elkholy, O. (2011). Wisconsin Card Sort Test (WCST) performance in schizophrenia and severe depression with psychotic features. *German Journal of Psychiatry* **14**, 91-94.

Rempfer, M., Hamera, E., Brown, C. & Bothwell, R. J. (2006). Learning proficiency on the Wisconsin Card Sorting Test in people with serious mental illness: What are the cognitive characteristics of good learners? *Schizophrenia Research* **87**, 316-322.

Rempfer, M. V., McDowd, J. M. & Brown, C. E. (2017). Measuring learning potential in people with schizophrenia: A comparison of two tasks. *Psychiatry Research* **258**, 316-321.

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**, 819-827.

Rossell, S. L., Coakes, J., Shapleske, J., Woodruff, P. W. R. & David, A. S. (2003). Insight: Its relationship with cognitive function, brain volume and symptoms in schizophrenia. *Psychological Medicine* **33**, 111-119.

Rossell, S. L. & David, A. S. (1997). Improving performance on the WCST: Variations on the original procedure. *Schizophrenia Research* **28**, 63-76.

Rybakowski, J. K., Borkowska, A., Czerski, P. M., Dmitrzak-Weglarz, M., Skibinska, M., Kapelski, P. & Hauser, J. (2006). Performance on the Wisconsin Card Sorting Test in schizophrenia and genes of dopaminergic inactivation (COMT, DAT, NET). *Psychiatry Research* **143**, 13-19.

Sasabayashi, D., Takayanagi, Y., Nishiyama, S., Takahashi, T., Furuichi, A., Kido, M., Nishikawa, Y., Nakamura, M., Noguchi, K. & Suzuki, M. (2017). Increased frontal gyrification negatively correlates with executive function in patients with first-episode schizophrenia. *Cerebral Cortex* **27**, 2686-2694.

Scarr, E., Sundram, S., Deljo, A., Cowie, T. F., Gibbons, A. S., Juzva, S., Mackinnon, A., Wood, S. J., Testa, R., Pantelis, C. & Dean, B. (2012). Muscarinic M1 receptor sequence: Preliminary studies on its effects on cognition and expression. *Schizophrenia Research* **138**, 94-98.

Seaton, B. E., Goldstein, G. & Allen, D. N. (2001). Sources of Heterogeneity in Schizophrenia: The Role of Neuropsychological Functioning. *Neuropsychology Review* **11**, 45-67.

Shafer, A. (2005). Meta-analysis of the brief psychiatric rating scale factor structure. *Psychological Assessment* **17**, 324-335.

Uren, J., Cotton, S. M., Killackey, E., Saling, M. M. & Allott, K. (2017). Cognitive clusters in first-episode psychosis: Overlap with healthy controls and relationship to concurrent and prospective symptoms and functioning. *Neuropsychology* **31**, 787-797.

Van Rheenen, T. E., Cropley, V., Zalesky, A., Bousman, C., Wells, R., Bruggemann, J., Sundram, S., Weinberg, D., Lenroot, R. K., Pereira, A., Shannon Weickert, C., Weickert, T. W. & Pantelis, C. (2018). Widespread Volumetric Reductions in Schizophrenia and Schizoaffective Patients Displaying Compromised Cognitive Abilities. *Schizophrenia Bulletin* **44**, 560-574.

Van Rheenen, T. E., Lewandowski, K. E., Tan, E. J., Ospina, L. H., Ongur, D., Neill, E., Gurvich, C., Pantelis, C., Malhotra, A. K., Rossell, S. L. & Burdick, K. E. (2017). Characterizing cognitive heterogeneity on the schizophrenia-bipolar disorder spectrum. *Psychological Medicine* **47**, 1848-1864.

Velligan, D. I., Bow-Thomas, C. C., Mahurin, R. K., Miller, A. L. & Halgunseth, L. C. (2000). Do specific neurocognitive deficits predict specific domains of community function in schizophrenia? *Journal of Nervous and Mental Disease* **188**, 518-524.

Waford, R. N. & Lewine, R. (2010). Is perseveration uniquely characteristic of schizophrenia? *Schizophrenia Research* **118**, 128-133.

Wechsler, D. (2001). *Wechsler Test of Adult Reading: WTAR*. Psychological Corporation.

Weickert, T. W., Goldberg, T. E., Gold, J. M., Bigelow, L. B., Egan, M. F. & Weinberger, D. R. (2000). Cognitive impairments in patients with schizophrenia displaying preserved and compromised intellect. *Archives of General Psychiatry* **57**, 907-913.

Wells, R., Swaminathan, V., Sundram, S., Weinberg, D., Bruggemann, J., Jacomb, I., Cropley, V., Lenroot, R., Pereira, A. M. & Zalesky, A. (2015). The impact of

premorbid and current intellect in schizophrenia: cognitive, symptom, and functional outcomes. *npj Schizophrenia* **1**, 15043.

Wilmsmeier, A., Ohrmann, P., Suslow, T., Siegmund, A., Koelkebeck, K., Rothermundt, M., Kugel, H., Arolt, V., Bauer, J. & Pedersen, A. (2010). Neural correlates of set-shifting: Decomposing executive functions in schizophrenia. *Journal of Psychiatry and Neuroscience* **35**, 321-329.

Woodward, N. D. & Heckers, S. (2015). Brain Structure in Neuropsychologically Defined Subgroups of Schizophrenia and Psychotic Bipolar Disorder. *Schizophrenia Bulletin* **41**, 1349-1359.

Wykes, T., Reeder, C., Landau, S., Everitt, B., Knapp, M., Patel, A. & Romeo, R. (2007). Cognitive remediation therapy in schizophrenia: Randomised controlled trial. *British Journal of Psychiatry* **190**, 421-427.

Exploring Heterogeneity on the Wisconsin Card Sorting Test in Schizophrenia-Spectrum Disorders. A Cluster Analytical Investigation

Sean P. Carruthers, Caroline T. Gurvich, Denny Meyer, Australian Schizophrenia Research Bank, Chad Bousman, Ian P. Everall, Erica Neill, Christos Pantelis, Philip J. Sumner, Eric J. Tan, Elizabeth H. X. Thomas, Tamsyn E. Van Rheenen, Susan L. Rossell

Supplementary Material

Supplementary Table 1. BPRS items with corresponding PANSS item.

BPRS Item	Corresponding PANSS Item
Affect	
Somatic Concern	G1. Somatic Concern
Anxiety	G2. Anxiety
Depression	G6. Depression
Guilt	G3. Guilty Feelings
Positive	
Grandiosity	P5. Grandiosity
Hallucinations	P3. Hallucinatory Behaviour
Unusual Thought Content	G9. Unusual Thought Content
Conceptual Disorganisation	P2. Conceptual Disorganisation
Negative	
Blunted Affect	N1. Blunted Affect
Emotional Withdrawal	N2. Emotional Withdrawal
Motor Retardation	G7. Motor Retardation

Supplementary Table 2. Comparison of WCST z-scores across sex and group.

		Relatively Intact (54/18)	Moderate Impairment (91/23)	Severe Impairment (13/11)	Healthy Control (106/86)	Total		Test Statistic	Sex x Group
						Male (n = 264)	Female (n = 139)		
WCST									
TA	Male	0.63 (1.1)	2.09 (0.0)	2.09 (0.0)	0.07 (1.0)	0.98 (1.2)	0.55 (1.3)	$F_{1,402} = 11.2, p \leq .001$	$F_{3,402} = 0.4, p = .75$
	Female	0.69 (1.0)	2.09 (0.0)	2.09 (0.0)	-0.11 (1.0)				
CAT	Male	0.06 (0.5)	-4.64 (1.4)	-7.49 (0.4)	0.03 (0.9)	-1.95 (2.7)	-1.35 (2.7)	$F_{1,402} = 4.3, p \leq .05$	$F_{3,402} = 0.2, p = .88$
	Female	0.16 (0.4)	-4.73 (1.3)	-7.2 (0.9)	0.04 (0.9)				
NPE	Male	0.34 (0.9)	2.33 (1.4)	3.24 (2.0)	0.04 (1.0)	1.05 (1.6)	0.68 (1.6)	$F_{1,402} = 4.9, p \leq .05$	$F_{3,402} = 0.9, p = .43$
	Female	0.41 (0.8)	2.6 (1.4)	2.80 (1.4)	-0.09 (0.9)				
PE	Male	0.18 (0.6)	2.51 (1.7)	3.90 (2.0)	-0.01 (1.0)	1.09 (1.8)	0.84 (1.8)	$F_{1,402} = 1.7, p = .19$	$F_{3,402} = 0.4, p = .73$
	Female	0.26 (0.6)	2.84 (1.3)	4.24 (1.8)	-0.02 (1.0)				
TFC	Male	0.06 (0.7)	1.26 (2.0)	12.12 (1.4)	0.05 (1.1)	1.06 (3.0)	1.26 (3.5)	$F_{1,402} = 0.3, p = .56$	$F_{3,402} = 2.3, p = .07$
	Female	-0.05 (0.5)	2.05 (2.6)	12.16 (1.1)	-0.14 (0.5)				
TC[#]	Male	1.15 (1.7)	0.22 (2.1)	-3.16 (2.5)	0.14 (1.1)	0.06 (1.9)	-0.39 (1.5)	$F_{1,339.0} = 6.4, p \leq .05$[*]	$F_{3,402} = 0.8, p = .52$
	Female	1.10 (1.6)	-1.0 (1.4)	-3.17 (2.0)	-0.17 (0.8)				
FMS[#]	Male	1.00 (1.9)	2.68 (3.0)	2.41 (4.1)	0.01 (1.0)	1.25 (2.5)	0.40 (1.7)	$F_{1,382.9} = 16.8, p \leq .001$[*]	$F_{3,402} = 2.2, p = .09$
	Female	0.63 (1.8)	1.21 (2.0)	1.59 (3.2)	-0.08 (0.8)				

Data is reported as mean (SD) unless otherwise stated; (Male/Female); WCST, Wisconsin Card Sorting Test; TA, trials administered; CAT, categories achieved; NPE, non-perseverative errors; PE, perseverative errors; TFC, trials to first category; TC, total correct; FMS, failure to maintain set; [#] not included in cluster analysis; [*] Brown-Forythe F-ratio reported Bold indicates large effect sizes

Supplementary Table 3. Comparison of WCST z-scores across schizophrenia-spectrum diagnosis and subgroup.

		Relatively Intact (50/22)	Moderate Impairment (102/12)	Severe Impairment (20/4)	Total		Test Statistic	Schizophrenia-Spectrum Diagnosis x Group
					Schizophrenia (n = 172)	Schizoaffective Disorder (n = 38)		
WCST								
TA	Sz	0.73 (1.1)	2.09 (0.0)	2.09 (0.0)	1.69 (0.9)	1.14 (1.1)	$F_{1,47.3} = 8.6, p \leq .001^{*}$	$F_{2,210} = 0.8, p = .47$
	SzA	0.45 (1.0)	2.09 (0.0)	2.09 (0.0)				
CAT	Sz	0.04 (0.5)	-4.65 (1.3)	-7.39 (0.6)	-3.61 (2.7)	-2.16 (3.0)	$F_{1,209} = 8.5, p \leq .01$	$F_{2,210} = 0.2, p = .85$
	SzA	0.18 (0.4)	-4.74 (1.4)	-7.26 (0.7)				
NPE	Sz	0.32 (0.9)	2.34 (1.3)	3.24 (1.7)	1.86 (1.6)	1.35 (1.6)	$F_{1,209} = 3.1, p = .82$	$F_{2,210} = 2.2, p = .12$
	SzA	0.44 (0.7)	2.80 (1.7)	2.04 (1.5)				
PE	Sz	0.27 (0.7)	2.59 (1.7)	4.08 (1.9)	2.09 (1.9)	1.2 (1.8)	$F_{1,209} = 6.4, p \leq .01$	$F_{2,210} = 0.2, p = .99$
	SzA	0.05 (0.6)	2.47 (1.2)	3.91 (2.1)				
TFC	Sz	0.03 (0.7)	1.29 (2.1)	12.07 (1.4)	2.18 (4.0)	2.12 (4.1)	$F_{1,209} = 0.01\ p = .94$	$F_{2,210} = 1.7, p = .18$
	SzA	0.03 (0.6)	2.53 (2.8)	12.42 (0.2)				
TC[#]	Sz	1.33 (1.7)	-0.34 (2.0)	-3.42 (2.2)	-0.21 (2.4)	-0.01 (1.9)	$F_{1,209} = 0.3\ p = .62$	$F_{2,210} = 1.8, p = .18$
	SzA	0.69 (1.4)	-0.66 (2.0)	-1.89 (2.1)				
FMS[#]	Sz	1.15 (2.1)	2.55 (2.9)	1.87 (3.8)	2.06 (2.9)	0.82 (1.9)	$F_{1,80.7} = 11.1, p \leq .001^{*}$	$F_{2,210} = 1.2, p = .30$
	SzA	0.36 (1.1)	0.97 (2.0)	2.88 (3.6)				

Data is reported as mean (SD) unless otherwise stated; (schizophrenia/schizoaffective disorder); WCST, Wisconsin Card Sorting Test; TA, trials administered; CAT, categories achieved; NPE, non-perseverative errors; PE, perseverative errors; TFC, trials to first category; TC, total correct; FMS, failure to maintain set; Sz, schizophrenia; SzA, schizoaffective disorder; [#] not included in cluster analysis; [*] Brown-Forythe F-ratio reported Bold indicates large effect sizes

Categories Achieved	Relatively Intact (n=72)		Moderate Impairment (n=114)		Severe Impairment (n=24)		Healthy Control (n=192)	
	Count	%	Count	%	Count	%	Count	%
0	-	-	-	-	21	87.5	-	-
1	-	-	34	29.8	2	8.3	-	-
2	-	-	34	29.8	1	4.2	1	0.5
3	-	-	32	28.1	-	-	5	2.6
4	-	-	14	12.3	-	-	7	3.6
5	12	16.7	-	-	-	-	6	3.1
6	60	83.3	-	-	-	-	173	90.1

Supplementary Table 4. Frequency of categories achieved across subgroups.

Figure S1. Agglomeration schedule

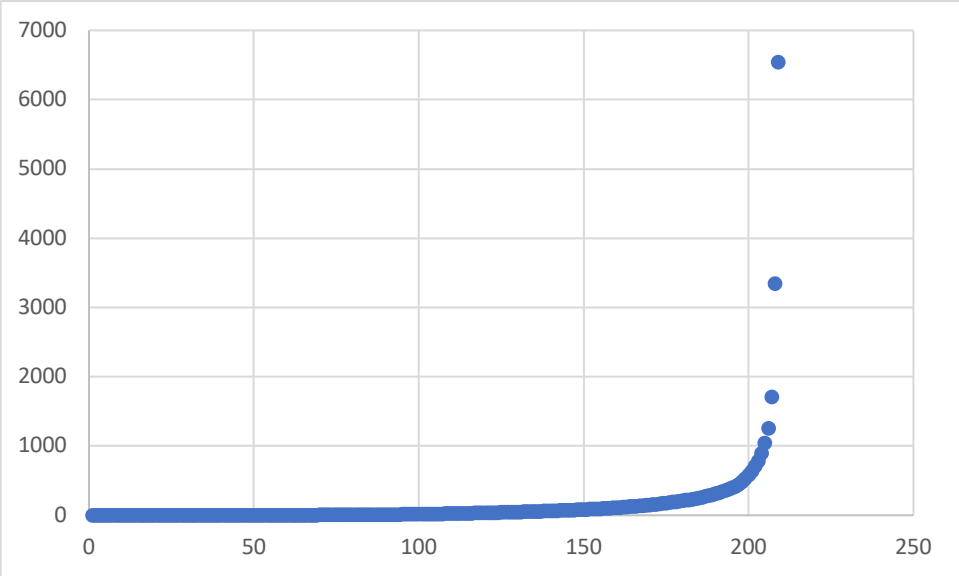

Figure S2. Dendrogram of cluster solution.