Journal of the International Neuropsychological Society (1998), **4**, 353–362.
Copyright © 1998 INS. Published by Cambridge University Press. Printed in the USA.

A comparison of clustering solutions for cognitive heterogeneity in schizophrenia

GERALD GOLDSTEIN,[1] DANIEL N. ALLEN,[1] AND BRENT E. SEATON[2]
[1]VA Pittsburgh Healthcare System and University of Pittsburgh
[2]VA Pittsburgh Healthcare System and University of South Dakota

(RECEIVED January 14, 1997; REVISED October 24, 1997; ACCEPTED November 7, 1997)

Abstract

A cluster analytic solution based upon a battery of tests consisting of the Halstead Category and Tactual Performance Tests, the Trail Making Test, and the Wisconsin Card Sorting Test was compared with a solution based on the subtests of the Wechsler intelligence scales, utilizing a sample of 221 schizophrenic patients. Both analyses permitted four-cluster solutions, and we found a weak but significant degree of association between solutions. Examination of external validity of the two solutions revealed stronger associations with clinical variables for the Wechsler-scale-based solution. The major conclusions were that the existence of cognitive heterogeneity in schizophrenia exists across a broad range of abilities, and appears to reflect a combination of continuity of ability level and existence of possible subtypes requiring further neuropsychological and neurobiological verification. (*JINS*, 1998, *4*, 353–362.)

Keywords: Schizophrenia, Cognitive function, Cognitive heterogeneity

INTRODUCTION

The presence of cognitive heterogeneity in schizophrenia is widely recognized, although there is substantial disagreement with regard to whether it reflects underlying subtypes of the disorder or a continuum of severity (DeLisi & Nasrallah, 1995). However, it is difficult to question the fact that schizophrenic patients vary widely in status of cognitive function, ranging from extreme impairment comparable to patients with cortical dementia to essentially no impairment of thinking detectable by standard cognitive tests (Braff et al., 1991; Goldstein, 1994). Some time ago, we suggested that the matter might be clarified through the use of empirical classification methods, and adopted cluster analysis as our statistical method for classification of patients on the basis of scores on tests of cognitive function (Goldstein, 1990). Subsequently, we did a study that did not utilize cluster analysis, but simply compared cognitive functioning on an extensive test battery between schizophrenic patients who were classified as doing well or poorly on the Wisconsin Card Sorting Test (Goldstein & Shemansky, 1996). All of these studies provided abundant evidence of extreme heterogeneity.

Since cluster analysis will classify cases even with random data, it is necessary to determine the external validity of clustering solutions, typically accomplished through determining if cluster membership is associated with pertinent variables not included in the cluster analysis itself. In the case of our studies, we used demographic and clinical variables such as age, education, length of illness, length of hospitalization, and age of onset of illness, finding significant degrees of association between cluster membership and some of these variables. However, a limitation of our previous work was that the tests of cognitive function used related exclusively to the areas of abstract reasoning and problem solving. These areas were chosen because they reflect cardinal deficits in schizophrenia, and we wished to limit the number of variables used in our studies for statistical purposes. However, cognitive function is often impaired in other areas in schizophrenia, and patterns of heterogeneity may differ substantially in those areas. As a result, it is not clear whether patterns of heterogeneity noted in previous studies were produced by the psychometric characteristics of the tests used and the abilities assessed by those tests. Furthermore, such patterns may be influenced by the demographic and clinical characteristics of the samples selected for study.

In order to address these issues, we determined to classify, using cluster analysis of a different set of tests, a large

Reprint requests to: Gerald Goldstein, VA Pittsburgh Healthcare System (151R), 7180 Highland Drive, Pittsburgh, PA 15206.

https://doi.org/10.1017/S1355617798003531 Published online by Cambridge University Press

sample of schizophrenic patients that were classified in previous studies (Goldstein, 1990; Goldstein & Shemansky, 1995) using abstraction and problem solving tests. These earlier studies identified four stable clusters varying in level and pattern of performance. One of the clusters reflected near-normal performance on the tests relative to available norms. Another cluster was characterized by severe, global impairment comparable to what is seen in patients with dementia. The other two clusters had more moderate impairment, varying from each other in pattern.

Using two sets of test procedures with a common sample allows for assessing the influence of test characteristics while controlling for the demographic and clinical variability that could occur as a result of using different samples. The new tests chosen were the subtests of the Wechsler Adult Intelligence Scales (WAIS and WAIS–R) (Wechsler, 1955, 1981). Since the study was retrospective in nature, we had data from the older WAIS and the WAIS–R. Preliminary analyses revealing no differences in relations among subtests and clustering solutions, and only an anticipated Full Scale IQ WAIS *versus* WAIS–R difference of 6 points (Kaufman, 1990; Wechsler, 1981) encouraged us to combine data from the two versions of the Wechsler scale in order to maximize sample size. The hypothesis was that the Wechsler scales (abbreviated as WAIS in the remainder of the paper) and the abstraction and problem-solving battery (referred to as the *abstraction battery* in the remainder of the paper) should generate comparable clusters. That is, the same pattern of cognitive heterogeneity found using abstraction and problem-solving neuropsychological tests should be found with another set of tests that evaluates a variety of cognitive abilities.

The tests utilized in these earlier studies have been largely associated with "executive" or frontal lobe function (Goldberg & Weinberger, 1988). There has been no such suggestion for the WAIS, particularly since the various subtests have been reported to have localizing value for several brain regions (Reitan & Wolfson, 1993). Furthermore, some of the WAIS subtests assess abilities that may be more highly associated with premorbid function than with the consequences of the schizophrenia, while the abstract reasoning and problem solving tests would appear to be more specifically associated with schizophrenic thought disorder. Kremen et al. (1996) have indicated that the Reading subtest of the Revised Wide Range Achievement Test (WRAT–R; Jastak & Wilkinson, 1984) provides a sound estimate of premorbid function in schizophrenia. In order to evaluate this matter, we used WRAT (Jastak & Jastak, 1965) Reading scores as an external validity criterion for the cluster analyses. The WRAT–R had not been published when the bulk of our data were obtained. This procedure was used to evaluate the influence of premorbid ability on cluster assignment.

A recommended method of evaluating stability of a cluster solution is that of cross-tabulation of the solution derived from one algorithm, such as Ward's method, with another algorithm, such as iterative partitioning. In the case of the present study, we used this cross-tabulation method to compare cluster solutions using the same algorithm but different measures. Failure to find a significant proportion of cases in comparable clusters would indicate that the WAIS and our previously employed battery of abstraction and problem solving tests produce different classification solutions. Finding a significant proportion would indicate that both procedures produce similar classification of cases. Combining the abstraction battery tests with the WAIS into a single cluster analysis could provide additional information concerning whether or not the cluster solution is substantially altered, as well as about the kinds of cognitive tests that produce subtypes with optimal external validity.

METHODS

Research Participants

The sample consisted of the same 221 schizophrenic patients described in Goldstein and Shemansky (1995). In that study two subsamples were compared, one consisting of 136 cases that were diagnosed clinically and the other group consisting of 85 prospectively recruited participants, diagnosed as schizophrenic utilizing the Structured Clinical Interview for DSM–III–R (SCID–P; Spitzer et al., 1989) and expert psychiatric consultation. Data analyses performed in that study revealed no differences between those groups across a broad range of cognitive and clinical variables. Most significantly, the cluster analytic solution for the battery of abstraction and problem solving tests described below was essentially identical in the two subgroups. Thus, they were combined for subsequent studies. All participants, regardless of diagnostic documentation procedure used, met DSM–III–R criteria for schizophrenia (American Psychiatric Association, 1980). They were all tested at a VA hospital in Topeka, Kansas or Pittsburgh, Pennsylvania. Patients with active substance abuse, other psychiatric diagnoses, or inability to cooperate for all of the tests were excluded. The present study was made possible because all of these participants had taken the full WAIS or WAIS–R in addition to the battery of abstraction tests.

Materials

The tests in what we will call the *abstraction battery* included the Halstead Category Test, the Halstead Tactual Performance Test, the Trail Making Test (Reitan & Wolfson, 1993), and the Wisconsin Card Sorting Test (Heaton, 1981). As indicated above, preliminary analyses revealed only the anticipated IQ discrepancy between the WAIS and WAIS–R, with WAIS–R scores being lower, and no significant difference in subtest pattern between the corresponding subtests of the Wechsler Adult Intelligence Scale (WAIS; Wechsler, 1955) and the revised version of that scale (WAIS–R; Wechsler, 1981). Therefore, the two versions were combined for purposes of the present study. The reason for the necessity of this procedure was that data were collected over a lengthy period of time, beginning substantially before the WAIS–R was available. As is pointed out in the WAIS–R manual,

https://doi.org/10.1017/S1355617798003531 Published online by Cambridge University Press

Table 1. Cluster profile for the abstraction battery

	Cluster							
	1 (Moderately Impaired + Psychomotor Deficit)		2 (Near Normal)		3 (Severely Impaired)		4 (Moderately Impaired)	
Test	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Category errors	91.75	24.40	48.67	18.88	102.22	24.87	93.91	18.29
TPT (min)	24.73	5.68	17.73	6.10	27.37	4.64	22.41	6.00
TRB (s)	165.73	25.02	74.98	24.07	292.12	19.63	99.02	15.34
WCST–Categories	2.53	2.16	3.52	2.13	.98	1.21	2.57	2.19
WCST–Cards	122.45	13.97	118.47	17.59	126.85	7.34	122.98	12.26

Note. Clusters have been assigned brief names reflecting cognitive status. TPT = Tactual Performance Test; TRB = Trail Making Test, Part B; WCST = Wisconsin Card Sorting Test.

80% of the items remained the same, with most changes made to eliminate items with antiquated content. We switched over to the WAIS–R because of the desire to use contemporary norms and to avoid inappropriate use of antiquated content. Thus, all participants received the appropriate version of the scale for their normative reference groups. The original sample of 136 had received the WAIS, while the new sample of 85 received the WAIS–R. The dependent measures were the 11 subtest scaled scores; Information, Comprehension, Arithmetic, Similarities, Digit Span, Vocabulary, Picture Completion, Block Design, Picture Arrangement, Object Assembly, and Digit Symbol.

Cluster Analysis Methodology

Cluster analyses were performed for both batteries combined, for the abstraction battery alone, and for the WAIS alone using the SPSS *CLUSTER* program (SPSS, 1986). In the case of the combined analysis, Verbal and Performance total scaled scores were substituted for the 11 subtest scores in order to reduce the number of variables. As in previous studies, we used Ward's method with squared Euclidean distance as the similarity measure. The number of clusters was determined by preliminary inspection of the dendograms followed by discriminant function analyses with plotting of clusters in discriminant function space. This latter method is described in Aldenderfer and Blashfield (1984), and involves producing scatterplots of solutions involving a varying number of clusters in two-dimensional space, and determining the point at which the clusters can be seen clearly, with minimal or no overlap. In order to determine similarity of cluster solutions between the WAIS and abstraction battery analyses cluster solutions were cross-tabulated in contingency tables, and analyzed with χ^2 tests and Kappa coefficients. These tables tabulate the number of cases that retain consistent cluster membership across procedures being compared, and number of cluster reassignments across procedures. A significant result would indicate the presence of a strength of association between the two solutions that exceeded chance. Cross-tabulations were produced based upon the cluster solutions derived from the three cluster analyses (WAIS *vs.* abstraction, combined *vs.* abstraction, and WAIS *vs.* combined).

External Validity

As in previous studies, external validity of the cluster solutions was evaluated by obtaining demographic and clinical variables not included in the cluster analysis. In this study, we examined for intercluster differences in the three solutions for age, education, length of illness, age of onset of illness, length of hospitalization, number of times hospitalized, and number of antipsychotic drugs taken at time of testing. Comparisons among the three cluster analyses would indicate whether or not intercluster differences within each solution are associated to comparable extents with these demographic and clinical variables. Additionally, we evaluated the possible contribution of premorbid level of function to cluster membership through utilization of the WRAT Reading score as an external validity criterion measure.

RESULTS[1]

The Abstraction Battery Cluster Analysis

In the Goldstein and Shemansky (1995) study, using the methods of inspection of the dendogram and discriminant function analysis, we were able to justify a four-cluster solution. That cluster analysis is the same as the one utilized here. The cluster profile is presented in Table 1. We identified a cluster with close to normal function (Cluster 2), a cluster with uniformly severe impairment (Cluster 3), and

[1] Full data concerning all of the cluster analyses and associated statistical procedures are available from the senior author.

https://doi.org/10.1017/S1355617798003531 Published online by Cambridge University Press

two clusters with moderate impairment but different profile configurations (Clusters 1 and 4). In the case of one of these clusters abstract reasoning was moderately impaired in association with very poor performance on tests with a psychomotor component (Cluster 1), while the other cluster (Cluster 4) had comparably impaired abstract reasoning ability, but relatively more intact psychomotor function.

The WAIS Cluster Analysis

Inspection of the cluster dendograms and discriminant function analyses again justified acceptance of a four-cluster solution. Trials with larger numbers of clusters produced unacceptably small clusters or clusters that did not separate well in discriminant function space. As in the case of the abstraction tests, there was one cluster with all average range or above scores (Cluster 2), and one with uniformly poor scores (Cluster 3). Cluster 1 has a profile that is notable for the discrepancy between relatively high verbal scores and low performance scores. Cluster 4, like Cluster 3, has an undifferentiated profile, but with less deficit overall. In particular, Cluster 4 does not have the extremely low performance test scores seen in Cluster 3. The cluster profile is presented in Table 2.

The Combined Cluster Analysis

Utilizing inspection of the dendograms and discriminant function analysis it was determined that a four-cluster solution could again be justified that had satisfactory internal validity. That is, four clusters were readily visualized on the dendogram, and could be localized in separate areas of the discriminant function space. Analyses with larger numbers of clusters did not generate this clear separation. The cluster profile is presented in Table 3. The first cluster is characterized by moderately impaired scores on the abstraction and problem solving tests, WAIS verbal test scores in the average range, and WAIS performance scores that are substantially below average. Cluster 2 obtained average or above mean scores on all of the tests. Cluster 3 obtained uniformly severely impaired scores on all of the tests. Cluster 4 shares some characteristics of Cluster 3, notably very impaired performance on the Category and Tactual Performance Tests, but is more comparable to Cluster 1 on the WCST and WAIS scores.

Comparison of the Cluster Solutions

Table 4 contains the frequencies of cases that were assigned to the clusters by the three cluster analyses (WAIS *vs.* abstraction, abstraction *vs.* combined, and WAIS *vs.* combined). Since there were different numbers of cases in each cluster, row and column percentages are also presented.

The comparison between the abstraction battery alone and the WAIS yielded the following results. In the instance of the cluster that performed in the near-normal range on the abstraction battery, 84% of these cases were grouped in the WAIS-based cluster analysis into the cluster characterized by average or above scores on all of the subtests (WAIS Cluster 2). None of the cases that were grouped into the average or above WAIS based cluster analysis were grouped into the abstraction-battery-based cluster (Cluster 3) marked by generalized severe impairment. In the case of the abstraction battery cluster marked by generalized severe impairment, 52% of these cases were assigned to Cluster 3, the most impaired WAIS based cluster. Abstraction battery Cluster 1 had a large number of placements in WAIS Clus-

Table 2. Cluster profile for the WAIS

	Cluster							
	1 (Average Verbal; Moderately Poor Performance)		2 (Average or Above)		3 (Severely Poor)		4 (Moderately Poor)	
Subtest	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Information	11.68	2.20	13.58	2.34	7.67	1.74	8.79	2.13
Comprehension	10.53	2.72	13.89	3.05	6.07	1.70	7.87	2.14
Arithmetic	10.22	2.43	14.37	1.57	6.44	2.03	7.51	1.89
Similarities	11.42	2.27	13.84	2.07	5.91	2.39	8.17	2.05
Digit Span	10.72	2.71	13.37	3.44	6.93	2.63	7.80	2.39
Vocabulary	11.46	1.84	14.11	2.49	6.98	2.10	8.58	1.47
Digit Symbol	6.71	1.96	9.26	2.64	3.94	1.78	6.30	1.65
Picture Completion	9.43	2.28	11.84	2.63	5.33	1.63	8.50	1.47
Block Design	9.24	2.05	12.58	2.46	4.94	1.88	8.41	2.30
Picture Arrangement	8.16	1.95	11.68	2.79	5.11	1.49	7.66	2.14
Object Assembly	8.67	2.13	11.95	1.93	4.65	1.90	8.84	2.52

Note. Clusters have been assigned brief names reflecting intellectual status.

https://doi.org/10.1017/S1355617798003531 Published online by Cambridge University Press

Table 3. Cluster profile for combined abstraction battery and WAIS

	Cluster							
	1 (Moderately Impaired)		2 (Near Normal)		3 (Severely Impaired)		4 (Moderately Impaired + Psychomotor Deficit)	
Test	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Category Errors	81.59	24.68	44.97	16.64	103.71	25.18	102.31	22.80
TPT (min)	22.21	6.27	16.93	5.78	26.92	4.89	26.68	5.01
TRB (s)	114.31	31.87	66.42	16.71	298.97	5.03	199.38	26.70
WCST–Categories	2.71	2.1	3.77	2.25	0.9	1.19	1.9	1.8
WCST–Cards	123.06	12.98	114.24	19.84	128.00	0.00	127.12	3.49
WAIS Verbal scaled scores	55.09	13.35	66.90	12.97	41.80	11.46	49.31	11.56
WAIS Performance scaled scores	37.95	8.25	49.17	8.52	25.46	8.93	32.00	6.61

Note. Clusters have been assigned brief names reflecting cognitive status. TPT = Tactual Performance Test; TRB = Trail Making Test, Part B; WCST = Wisconsin Card Sorting Test.

ters 3 and 4, both of which had below-average scores. Abstraction battery Cluster 3 had over half of its placements in WAIS Cluster 3, the cluster with the lowest WAIS subtest scores. The association between solutions was highly significant [$\chi^2(9) = 97.52$, $p < .001$]. Kappa was equal to .18 ($p < .05$), suggesting significant, but not strong agreement beyond chance.

With regard to the combined WAIS and abstraction test analysis as compared with the abstraction tests alone, the major shifting of cluster assignment occurred for Clusters 1 and 4 of both solutions. In particular, participants in combined-battery-based Cluster 1, reflecting moderate impairment, were distributed into all of the clusters based on the abstraction battery, with the exception of the cluster reflecting severe impairment (Cluster 3). However, only 1 participant placed in the near-normal cluster by the combined battery shifted out of that cluster. That is, the participants in the cluster characterized by bright-normal or superior general intelligence almost always were placed in the near-normal cluster based on the abstraction battery. Similarly, participants with very low general intelligence were mainly placed in Cluster 3 in both analyses, the clusters characterized by severe impairment. Thus, the general impact of adding the WAIS variables was on the moderately impaired participants, who were redistributed substantially. Participants with average or above performance, and those with severely impaired performance, did not show a substantial change in cluster membership. A test for strength of association between the two solutions yielded a $\chi^2(9)$ of 352.66 ($p < .001$). The Kappa coefficient was .41 ($p < .001$), indicating good agreement between the two solutions.

With regard to the WAIS *versus* combined battery comparison, the large majority of participants grouped into the severely impaired cluster (Cluster 3) based on the combined battery were grouped into a comparable cluster based on WAIS subtest scores. Correspondingly, all of the participants in the near-normal abstraction battery cluster were placed into WAIS-based Clusters 1 or 2, which contained participants with average or above levels of general intelligence. The participants in abstraction-battery-based Cluster 1, reflecting moderate impairment without major psychomotor deficit, were mainly placed into WAIS Clusters 1 and 3. WAIS Cluster 1 was marked by a sharp reduction in performance relative to verbal scores, while WAIS Cluster 3 did not exhibit such a discrepancy. The great majority of the participants in abstraction-battery-based Cluster 4 were placed in WAIS-based Clusters 3 and 4, both of which were characterized by low-average to impaired levels of performance on the WAIS subtests. The test of strength of association between the two clusters yielded a $\chi^2(9)$ of 106.86 ($p < .001$). Kappa was equal to .18 ($p < .05$), indicating statistically significant but not strong agreement between the two solutions.

Examination of Table 4 indicates that we did not find direct correspondences between each cluster. Such a direct correspondence was only found for the two Cluster 2s; the high-functioning clusters. The significant Chi-squares should only be interpreted to mean that the association between the two solutions was sufficiently strong to create differences between observed and expected cell frequencies that exceeded chance expectation. High Kappas would have suggested the presence of a direct cluster-for-cluster correspondence, which was not the case here, except in the instance of the abstraction-battery *versus* combined-battery comparison.

External Validity

External validity data for the three clustering solutions are presented in Table 5. All three solutions generated signifi-

https://doi.org/10.1017/S1355617798003531 Published online by Cambridge University Press

Table 4. Classification matrix comparing abstraction battery, WAIS, and combined battery cluster solutions

WAIS *versus* abstraction battery				
	Abstraction battery			
WAIS	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Cluster 1	14	43	5	14
	18.4	56.6	6.6	18.4
	25.9	50.6	12.5	31.1
Cluster 2	1	16		2
	5.3	84.2		10.5
	1.9	18.8		4.4
Cluster 3	19	1	28	6
	35.2	1.9	51.9	11.1
	35.2	1.2	70.0	13.3
Cluster 4	20	25	7	23
	26.7	33.3	9.3	30.7
	37.0	29.4	17.5	51.1
Abstraction battery *versus* combined battery				
	Abstraction battery			
Combined battery	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Cluster 1	33	27		44
	31.7	26.0		42.3
	61.1	31.8		97.8
Cluster 2	58		1	
	98.3		1.7	
	68.2		2.2	
Cluster 3			35	
			100.0	
			87.5	
Cluster 4	21		5	
	80.8		19.2	
	38.9		12.5	
WAIS *versus* combined battery				
	WAIS			
Combined battery	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Cluster 1	36	3	17	48
	34.6	2.9	16.3	46.2
	47.4	15.8	31.5	64.0
Cluster 2	30	16	1	12
	50.8	27.1	1.7	20.3
	39.5	84.2	1.9	16.0
Cluster 3	5		25	5
	14.3		71.4	14.3
	6.6		46.3	6.7
Cluster 4	5		11	10
	19.2		42.3	30.5
	6.6		20.4	13.3

Note. Top number in each cell is count of number of cases in corresponding clusters, middle number is row percent, and last number is column percent.

cant intercluster differences in age, education, Full Scale IQ score, and Reading grade level score from the WRAT. None of the analyses produced significant differences for age of onset of illness. There were some discrepancies, however, among the other variables. In the case of the combined battery cluster analysis, there was a significant difference for number of times hospitalized, but a Scheffé test ($p < .05$) indicated that none of the clusters were

https://doi.org/10.1017/S1355617798003531 Published online by Cambridge University Press

Table 5. External validity variables for the three cluster analyses

Variable	Cluster 1 *M*	Cluster 1 *SD*	Cluster 2 *M*	Cluster 2 *SD*	Cluster 3 *M*	Cluster 3 *SD*	Cluster 4 *M*	Cluster 4 *SD*	*p*
Combined battery									
Age (years)	41.12	9.88	36.60	10.38	43.37	8.59	43.92	10.48	<.001
Education (years)	12.12	2.18	13.52	2.21	10.46	3.43	11.40	2.22	<.001
Full Scale IQ	90.94	10.55	105.24	11.24	79.17	9.23	88.20	10.23	<.001
WRAT Reading	9.62	3.86	12.21	3.64	8.65	3.63	9.64	4.53	<.01
Age of onset (years)	26.64	8.81	27.70	9.58	29.40	10.12	26.94	11.37	>.05
Months in hospital	24.47	29.29	18.85	35.75	24.63	28.51	44.06	47.33	>.05
Months of illness	185.86	127.94	131.55	105.32	155.80	114.52	202.44	185.91	>.05
Times hospitalized	10.00	8.91	7.11	7.85	6.96	7.50	13.57	13.32	<.05
Number of drugs	2.05	1.53	1.88	1.44	2.43	1.69	2.35	1.73	>.05
Abstraction battery									
Age (years)	42.24	9.59	36.79	9.65	44.63	9.32	41.04	11.24	<.001
Education (years)	11.71	2.25	13.34	2.07	10.61	3.31	12.48	2.48	<.001
Full Scale IQ	89.15	11.94	103.01	11.01	79.85	8.88	93.31	12.04	<.001
WRAT Reading	10.16	4.68	11.75	3.47	8.52	3.57	9.88	4.09	<.05
Age of onset (years)	27.93	9.41	26.68	7.80	27.93	11.08	28.34	9.52	>.05
Months in hospital	31.50	43.32	15.67	17.79	33.83	42.48	26.48	36.53	<.05
Months of illness	178.39	132.74	141.94	105.86	195.03	147.55	170.37	137.59	>.05
Times hospitalized	11.16	10.06	7.40	7.74	8.82	10.89	9.62	7.93	>.05
Number of drugs	2.36	1.55	1.89	1.45	2.54	1.90	2.07	1.59	>.05
WAIS									
Age (years)	41.58	9.93	34.42	10.22	44.93	8.27	37.17	10.24	<.001
Education (years)	13.28	2.24	14.79	2.39	10.78	3.04	11.64	1.78	
Full Scale IQ	102.33	6.95	119.68	6.83	78.15	6.60	89.21	6.86	<.001
WRAT Reading	12.01	2.86	15.12	2.33	7.59	3.96	8.25	3.18	<.001
Age of onset (years)	29.84	10.52	28.08	9.08	26.48	9.14	26.12	8.80	>.05
Months in hospital	20.65	25.37	29.31	63.63	32.64	39.47	22.14	28.30	>.05
Months of illness	162.12	124.12	95.33	106.46	218.21	141.47	142.51	113.62	<.01
Times hospitalized	7.48	6.96	5.92	5.57	12.08	11.20	8.36	9.08	<.05
Number of drugs	1.95	1.36	1.00	1.04	2.47	1.61	2.16	1.62	<.05

significantly different from each other. For the abstraction battery, there was a significant difference for length of hospitalization, but the Scheffé test ($p < .05$) indicated no significant intercluster differences. In the case of the WAIS-based cluster analysis, there were significant differences for months of illness, number of hospitalizations, and number of psychoactive drugs administered. The Scheffé tests indicated that Cluster 3 had a significantly greater length of illness than did Clusters 2 and 3 ($p < .05$), and that Cluster 3 was administered a significantly larger number of psychoactive drugs than were the other clusters ($p < .05$). That is, more members of Cluster 3 required more than one drug than was the case for the other clusters. For number of hospitalizations, the Scheffé test ($p < .05$) indicated that there were no significant between-group comparisons.

DISCUSSION

This study indicates that cognitive heterogeneity may be demonstrated in schizophrenia with tests measuring other than abstraction and problem-solving abilities. The Wechsler intelligence scales provide measures of various verbal skills, attention, and spatial–constructional abilities, as well as measures of abstraction and problem-solving, but also provide evidence of heterogeneous performance among schizophrenic patients. It therefore appears that cognitive heterogeneity is a relatively pervasive phenomenon, affecting a wide range of abilities. That being the case, the question remains as to whether this heterogeneity is reducible to a limited number of subtypes, or whether it represents a continuum of severity. This question is often asked in terms of whether proposed subtypes differ with regard to patterns of cognitive profiles or only with regard to level of performance, with little difference among profile patterns. We will try to demonstrate that the results of previous research (Goldstein, 1994) and the present study suggest that heterogeneity is produced by the combined effects of differing subtypes and variability in severity of the schizophrenia.

To our knowledge, this is the first investigation that has examined cognitive heterogeneity in schizophrenia using two markedly different sets of cognitive measures for classification of the same sample of patients. It has been unclear

https://doi.org/10.1017/S1355617798003531 Published online by Cambridge University Press

whether the clusters previously described reflected the specific psychometric and content characteristics of the tests used, unique demographic and clinical characteristics of the sample studied, or actual cognitive subtypes that transcended particular tests. By applying the two sets of tests to the same sample we effectively controlled for possible variability resulting from sampling characteristics. The comparison between the combined WAIS and abstraction battery with the abstraction battery alone indicated similar solutions. That is, addition of WAIS variables to our previously reported solution involving only the abstraction tests did not substantially alter the cluster pattern.

With regard to the existence of cognitive subtypes, both the Wechsler scale and cognitive battery results provide evidence of two sharply contrasting groups within the same sample; a group with minimal psychometric evidence of cognitive dysfunction and a group with severe, generalized impairment comparable to what is seen in more elderly individuals with progressive dementia. The present study indicates that there is substantial association between the Wechsler Scale and abstraction battery analysis with regard to identification of patients in these groups. That is, both sets of tests, used separately or combined, produced a cluster characterized by near-average to above-average cognitive function and another cluster containing patients with evidence of severe global impairment consistent with dementia. Clusters 1 and 4 of the abstraction battery cluster analysis had similar profiles except for Part B of the Trail Making Test, on which the Cluster 1 participants performed substantially more slowly than did the Cluster 4 participants. In the case of the WAIS-based analysis, Cluster 1 did poorly on the performance tests relative to the verbal tests, while Cluster 3 did relatively poorly on both verbal and performance tests. These findings might suggest the presence of a subtype in which psychomotor function is particularly impaired, and another group in which psychomotor function is impaired to a lesser extent, or in the context of other equally impaired cognitive deficits. This latter group might lie on a continuum with the clusters found in both analyses to have severe generalized impairment.

Having made these observations, it is important to note that the two solutions did not produce a complete set of corresponding clusters. While membership in the high-functioning WAIS based cluster (Cluster 2) was highly associated with membership in the high-functioning abstraction battery cluster (Cluster 2), membership in WAIS-based Cluster 3, reflecting low average intelligence, was associated with relatively equal numbers of cases in three of the four abstraction battery clusters. Some of the discrepancy between solutions may be attributed to differing psychometric characteristics of the two testing procedures. The abstraction battery is likely to have a more pronounced ceiling effect than the Wechsler scales, since most of the tests it contains were designed to detect impairment in brain-damaged patients, but such effects were not apparent in this study. It is nevertheless quite possible to be unimpaired on such neuropsychologically oriented cognitive tests with intellectual function (IQ) levels spanning a broad range. The Wechsler scales were not specifically designed to detect brain damage, but rather, to assess intellectual function in both abnormal and normal populations. What we appear to have demonstrated is the presence of heterogeneity in the case of both batteries, and some degree of association between cluster memberships in the two batteries, but we did not find a cluster for cluster correspondence.

The combined findings of this and our related studies also indicate that the cognitive disorder in schizophrenia does not uniformly involve some particular deficit or pattern of deficits. While a great deal of attention has been called to the presence of perseverative rigidity, particularly as assessed with the Wisconsin Card Sorting Test, we have noted elsewhere (Goldstein & Shemansky, 1996) that the presence of this deficit is typically accompanied by a large number of other cognitive deficits. Furthermore, a substantial proportion of patients with well-diagnosed schizophrenia do not demonstrate perseverative rigidity on the Wisconsin Card Sorting Test.

Studies of cognitive heterogeneity indicate that the thinking disorder in schizophrenia may be characterized by a variety of cognitive profiles and severity levels. The neurobiological correlates of this variability are not at all well understood, but the present findings might suggest that there may not be a single pathophysiology for schizophrenia. Such factors as age, education, and a number of clinical variables such as length of illness account for some, but clearly not all of the discrepancies. For example, it is noteworthy that in the case of the abstraction-battery-based clusters, there is only about a 5-year discrepancy in mean age between the cluster with near-normal cognitive function (Cluster 2) and the one with severely impaired function (Cluster 3). Neither mean age is in the progressive dementia of the elderly range, but the abstraction battery scores of Cluster 3 are indistinguishable from those typically obtained by patients with these disorders (Heaton et al., 1994). Thus, while age differences among the clusters are statistically significant, they do not account for the magnitude of the functional discrepancy. The available literature indicates that iatrogenic variables including medication status and length of institutionalization may make some contribution, but are far from accounting for the observed variability (Braff et al., 1990; Goldstein, 1994).

Another way of comparing the two-cluster solutions is through examination of their external validities. In both solutions, intercluster differences were significantly associated with age, years of education, and general intelligence. Somewhat surprisingly, there was not an association with age of onset of illness. However, it is difficult to study age of onset in a veterans population, because individuals with early onset typically never become members of the armed forces. The other clinical variables showed discrepancies between the two solutions. In the case of the abstraction battery solution, only months of hospitalization showed a significant intercluster difference. In the case of the WAIS-based solution, there were significant intercluster differences for length of illness, times hospitalized, and number

https://doi.org/10.1017/S1355617798003531 Published online by Cambridge University Press

of antipsychotic drugs, but not for length of hospitalization. Thus, the WAIS-based solution may be more sensitive to clinical phenomena than the abstraction-battery-based solution. That may be the case, in part, because some of the Wechsler scales assess abilities that may be more highly associated with premorbid function than with the consequences of the schizophrenia (Barber et al., 1996), while the abstraction battery would appear to be more specifically associated with schizophrenic thought disorder. The significant WRAT Reading scores as external validity criteria for the cluster analyses suggests that there is an influence of premorbid ability on cluster assignment. The clusters characterized by near-normal cognitive function had mean WRAT Reading scores ranging from close to the 12th grade to the 15th grade level. The means for the more impaired clusters were in the 8th to 10th grade level range. Perhaps the patients in the cognitively intact clusters have sufficient "brain reserve" (Satz, 1993) to forestall the appearance of cognitive dysfunction sufficiently severe to appear as deficits on neuropsychological tests.

In summary, cognitive heterogeneity in schizophrenia may be made manifest through a variety of assessment procedures. A comparison between the WAIS and a specialized battery of tests of abstract reasoning and problem-solving both permitted a cluster-analytic four-cluster solution with satisfactory internal validity. Both solutions identified a subgroup of schizophrenic patients with no apparent evidence of cognitive dysfunction and another subgroup with levels of cognitive abilities comparable to those of patients with progressive dementia. More tentatively, there appear to be two additional subgroups with moderate cognitive dysfunction, one with and the other without a significant psychomotor dysfunction component. Future research may productively address itself to the problem of further understanding of these two subgroups, perhaps utilizing more refined tests of such areas as memory and speed of information processing in combination with relevant neurochemical and neuroimaging studies. Clinical and demographic external validity analyses indicate that some, but clearly not all, of the differences among clusters may be associated with age, education, and illness-related considerations. Differences in external validity findings between the WAIS and abstraction test based analyses suggested a possible relationship with premorbid function. An analysis of this consideration using the Reading subtest of the WRAT indicated that the subgroup in all cluster analyses with near-normal cognitive abilities may have had a higher level of premorbid function than did members of the other clusters. One could argue that the entire proposed subtype structure might simply reflect a wide range of continuous variability in cognitive function that emerges as the brain develops. Results of this and related studies discussed above (Goldstein, 1990; Goldstein & Shemansky, 1995, 1996) suggest that a combination of continuity and subtypal factors are involved, and that cognitive heterogeneity in schizophrenia may be associated with a combination of as yet not fully articulated subtypes and a continuum of severity of impairment.

ACKNOWLEDGMENT

Indebtedness is expressed to the Department of Veterans Affairs for support of this research.

REFERENCES

Aldenderfer, M.S. & Blashfield, R.K. (1984). *Cluster analysis*. Beverly Hills, CA: Sage Publications.

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, DC: Author.

Barber, F., Pantelis, C., Bodger, S., & Nelson, H.E. (1996). Intellectual functioning in schizophrenia: A natural history. In C. Pantelis, H.E. Nelson, & T.R.E. Barnes (Eds.), *Schizophrenia: A neuropsychological perspective* (pp. 49–70). Chichester, U.K.: John Wiley & Sons.

Braff, D.L., Heaton, R.K., Kuck, J., Cullum, M., Moranville, J., Grant, I., & Zisook, S. (1991). The generalized pattern of neuropsychological deficits in outpatients with chronic schizophrenia with heterogeneous Wisconsin Card Sorting Test results. *Archives of General Psychiatry*, *48*, 891–898.

DeLisi, L.E. & Nasrallah, H.A. (1995). Current controversies in schizophrenia research. I. Is schizophrenia a heterogeneous disorder? *Schizophrenia Research*, *17*, 133.

Goldberg, T.E. & Weinberger, D.R. (1988). Probing prefrontal function in schizophrenia with neuropsychological paradigms. *Schizophrenia Bulletin*, *14*, 160–185.

Goldstein, G. (1990). Neuropsychological heterogeneity in schizophrenia: A consideration of abstraction and problem solving abilities. *Archives of Clinical Neuropsychology*, *5*, 251–264.

Goldstein, G. (1994). Cognitive heterogeneity in psychopathology: The case of schizophrenia. In P. Vernon (Ed.), *The neuropsychology of individual differences* (pp. 209–233). New York: Academic Press.

Goldstein, G. & Shemansky, W.J. (1995). Influences on cognitive heterogeneity in schizophrenia. *Schizophrenia Research*, *18*, 59–69.

Goldstein, G. & Shemansky, W.J. (1996). Neuropsychological differences between schizophrenic patients with heterogeneous Wisconsin Card Sorting Test Performance. *Schizophrenia Research*, *21*, 13–18.

Heaton, R.K. (1981). *A manual for the Wisconsin Card Sorting Test*. Odessa, FL: Psychological Assessment Resources.

Heaton, R.K., Paulsen, J.S., McAdams, L.A., Kuck, J., Zisook, S., Braff, D., Harris, J., & Jeste, D.V. (1994). Neuropsychological deficits in schizophrenia: Relationship to age, chronicity, and dementia. *Archives of General Psychiatry*, *51*, 469–476.

Jastak, J.F. & Jastak, S.R. (1965). *The Wide Range Achievement Test manual*. Wilmington, DE: Guidance Associates.

Jastak, S. & Wilkinson, G.S. (1984). *Wide-Range Achievement Test–Revised*. Wilmington, DE: Jastak Assessment Systems.

Kaufman, A.S. (1990). *Assessing adolescent and adult intelligence*. Boston: Allyn & Bacon.

Kremen, W.S., Seidman, L.J., Faraone, S.V., Pepple, J.R., Lyons, M.J., & Tsuang, M.T. (1996). The "3 Rs" and neuropsychological function in schizophrenia: An empirical test of the matching fallacy. *Neuropsychology*, *10*, 22–31.

Reitan, R.M. & Wolfson, D. (1993). *The Halstead–Reitan Neuropsychological Test Battery: Theory and clinical interpretation* (2nd ed.). Tucson, AZ: Neuropsychology Press.

https://doi.org/10.1017/S1355617798003531 Published online by Cambridge University Press

Satz, P. (1993). Brain reserve capacity on symptom onset after brain injury: A formulation and review of evidence for threshold theory. *Neuropsychology*, *7*, 273–295.

Spitzer, R.L., Williams, J.B., Gibbon, M., & First, M.B. (1989). *Structured Clinical Interview for DSM–III–R* (Patient ed., 9/1/89 Version). New York: Biometrics Research Department, New York State Psychiatric Institute.

SPSS (1986). *SPSS^x users guide* (2nd ed.). Chicago: SPSS, Inc.

Wechsler, D. (1955). *Manual for the Wechsler Adult Intelligence Scale*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1981). *Manual for the Wechsler Adult Intelligence Scale–Revised*. San Antonio, TX: The Psychological Corporation.

https://doi.org/10.1017/S1355617798003531 Published online by Cambridge University Press