Journal of the International Neuropsychological Society (2005), **11**, 152–162.
Copyright © 2005 INS. Published by Cambridge University Press. Printed in the USA.
DOI: 10.1017/S1355617705050162

Verbal memory in first-episode schizophrenia: Heterogeneity in performance?

JASPER NUYEN, MARGRIET M. SITSKOORN, WIEPKE CAHN, AND RENÉ S. KAHN
Rudolf Magnus Institute of Neuroscience, Department of Psychiatry, University Medical Center, Utrecht, the Netherlands

(RECEIVED April 25, 2003; REVISED July 2, 2004; ACCEPTED July 20, 2004)

Abstract

It is unclear if the commonly observed "subcortical" verbal memory profile in schizophrenic patients is present at the onset of the disease. Therefore, the performance of 43 first-episode patients with schizophrenia or schizophreniform disorder on the Dutch version of the California Verbal Learning Test (VLGT) was compared to that of 43 normal comparison participants. We hypothesized that the first-episode patients would exhibit a "subcortical" memory profile, that is, they would show a primary retrieval deficit. This hypothesis was not confirmed: the patients displayed a profile suggestive of a prominent storage deficit, that is, a "cortical" memory profile. Subsequently, patients' VLGT performance was cluster analyzed to determine whether subgroups could be identified exhibiting a cortical, subcortical, and normal profile, respectively. Two subgroups ($N = 22$; $N = 13$) exhibited memory impairments, while one subgroup ($N = 8$) was unimpaired. The memory profiles of the two impaired subgroups differed both qualitatively and quantitatively, but did not conform neatly to a cortical and a subcortical profile. Demographic and verbal fluency data provided limited validation of the subgroup classification. Our results may suggest that combining the verbal memory performance of first-episode patients obscures meaningful heterogeneity. Alternatively, the cluster solution could merely reflect a continuum of severity. (*JINS*, 2005, *11*, 152–162.)

Keywords: First-episode schizophrenia, Verbal memory profiles, Heterogeneity, California Verbal Learning Test

INTRODUCTION

Cognitive deficits in patients with schizophrenia appear especially pronounced in the domain of verbal memory (Heinrichs & Zakzanis, 1998; Saykin et al., 1991). As there are several different profiles of verbal memory impairment, it is important to know which memory profile is associated with schizophrenia.

Research has shown that neurological disorders with known pathology in the frontal-striatal system, like Huntington's disease (HD) and Parkinson's disease (PD), present a profile of verbal memory impairment that is distinct from that observed in disorders with predominantly medial temporal lobe involvement, such as Alzheimer's disease (AD) (Butters et al., 1995; Delis et al., 1994; Savage, 1997). Patients with HD, PD, and AD show impairment on measures of immediate free recall and all usually exhibit deficits in the use of organizational strategies during learning. However, HD and PD patients demonstrate normal retention rates between immediate and delayed free recall, whereas AD patients consistently display rapid loss of information after a delay. When provided with a recognition format, HD and PD patients typically show a disproportional improvement in performance, resulting in (nearly) normal performance on measures of delayed recognition. In contrast, the performance of AD patients does not improve on recognition relative to free recall. Finally, providing semantic cues commonly leads to a disproportionately better recall in HD and PD patients, while AD patients are less likely to benefit from this retrieval aid. The so-called "subcortical" profile seen in HD and PD patients is thought to represent primarily a retrieval deficit, whereas the so-called "cortical" pattern observed in AD patients is thought to reflect primarily a storage deficit.

The cortical and subcortical memory profiles appear to be relevant to schizophrenia because, among other brain

Address correspondence and reprint requests to: Margriet M. Sitskoorn, Director, Cognitive Neuroscience Unit, Department of Psychiatry, University Medical Center Utrecht (HPN B.01.206), P.O. Box 85500, Utrecht, 3508 GA, The Netherlands. E-mail: M.Sitskoorn@azu.nl

https://doi.org/10.1017/S1355617705050162 Published online by Cambridge University Press

regions, the frontal lobes (Zakzanis & Heinrichs, 1999), the basal ganglia (Keshavan et al., 1998), and the (medial) temporal lobes (Nelson et al., 1998; Zakzanis et al., 2000) have been implicated in this disorder. Furthermore, since regional abnormalities are rather subtle and not consistently found, investigators have hypothesized dysfunction of frontal-striatal (Buchsbaum et al., 1992; Manoach et al., 2000) and frontal-(medial) temporal (Fletcher et al., 1999; Meyer-Lindenberg et al., 2001) circuitry in schizophrenia. A meta-analysis of studies investigating memory function in schizophrenia showed that patients did not perform differentially on measures of immediate *versus* delayed free recall, but exhibited significantly better performance when retrieval cues and a recognition format were provided (Aleman et al., 1999). These findings suggest that schizophrenic patients in general display a subcortical memory profile.

PART 1

Do First-Episode Patients with Schizophrenia Exhibit a "Subcortical" Verbal Memory Profile?

The suggestion that patients with schizophrenia typically exhibit a subcortical memory profile is based on findings of studies involving patients with varying degrees of chronicity.

Conclusions from such studies may have been influenced by variations regarding duration of neuroleptic treatment, institutionalization, and/or illness. Therefore, it is unclear if a subcortical profile is present at the onset of schizophrenia. Although studies have consistently found impaired verbal memory function among patients experiencing their first-episode of schizophrenia (Albus et al., 1996; Bilder et al., 2000; Binder et al., 1998; Hoff et al., 1992; Saykin et al., 1994; Sitskoorn et al., 2002), they did not examine the nature of the memory deficit. Mohamed et al. (1999) investigated the memory profile in first-episode patients in more detail and found their patients to be as impaired on the immediate recall of a short story as they were on delayed recall. This finding could suggest that deficits in encoding and/or retrieval rather than in storage underlie the verbal memory impairment seen in first-episode schizophrenia. However, Mohamed et al. did not investigate whether the performance improved disproportionately when retrieval aids were provided, which is considered to be a distinctive feature of a primary retrieval deficit

To more comprehensively characterize the verbal memory profile in first-episode schizophrenia, we examined the performance of patients with first-episode schizophrenia or schizophreniform disorder on selected measures of the Dutch version of the California Verbal Learning Test (CVLT; Delis et al., 1987). It was hypothesized that the first-episode patients would exhibit a subcortical memory profile (i.e., they would exhibit a primary retrieval deficit).

Methods

Research participants

Forty-three first-episode patients with schizophrenia (N = 32) or schizophreniform disorder (N = 11) and 43 normal comparison subjects participated in the study. At the time of assessment, all participants were in the age between 16 and 38 years, were physically healthy, and had no history of neurological illnesses or other medical conditions that have a detrimental impact on the central nervous system (as determined by a medical check list or medical records review). No participant had drug or alcohol dependency over a period of 3 months prior to assessment. Furthermore, all participants had either a premorbid IQ estimate or a current IQ estimate of at least 85, as evidenced by scores on the Dutch adaptation of the National Adult Reading Test (Schmand et al., 1992) or the combined scores on the subtests Vocabulary and Block Design of the Wechsler Adult Intelligence Scale (Silverstein, 1982).

The first-episode patients were involved in one of the treatment programs at the Department of Psychiatry of the University Medical Center (UMC) Utrecht. All patients fulfilled DSM–IV (American Psychiatric Association, 1994) criteria for schizophrenia or schizophreniform disorder as determined by the Comprehensive Assessment of Symptoms and History (CASH; Andreasen, 1987). The CASH is a structured interview for documenting the signs, symptoms, and history of participants in research studies of the major psychotic and affective disorders. Trained and experienced psychologists and psychiatrics conducted diagnostic assessment. Patients were excluded when a comorbid Axis I disorder was diagnosed and/or when positive symptomatology (i.e., hallucinations and/or delusions) was present for more than two years. Thirty patients were treated with atypical antipsychotic drugs and five with typical antipsychotics. Eight patients were receiving no antipsychotic medication. Seven patients took also anticholinergic medication.

Normal comparison participants were recruited by advertisements in national newspapers and received compensation for their participation. The CASH was used to screen out psychopathology in these participants. The first-episode patients were comparable to the normal comparison participants regarding baseline characteristics (Table 1). Because schizophrenia may influence educational attainment, parental education was used as matching variable (Resnick, 1992). Note that not all patients and comparison participants were administered both the estimate of premorbid IQ and the current IQ estimate.

Procedure and VLGT

The Dutch version of the CVLT, the "Verbale Leer en geheugen test" (VLGT; Mulder et al., 1996) was administered as part of a comprehensive neuropsychological battery. Trained psychometricians tested participants, supervised by a senior neuropsychologist (M.S.).

https://doi.org/10.1017/S1355617705050162 Published online by Cambridge University Press

Table 1. Subject characteristics

	First-episode patients ($N = 43$)	Normal comparison participants ($N = 43$)	p
Gender (male, %)	35 (81.4%)	33 (76.7%)	.60[e]
Handedness (right, %)	37 (86.0%)	39 (90.7%)	.36[e]
Age (years) ($M \pm SD$)	25.4 ± 5.0	25.0 ± 6.0	.72[f]
Premorbid IQ estimate ($M \pm SD$)[a]	103.9 ± 6.4	106.6 ± 6.7	.74[f]
IQ estimate ($M \pm SD$)[b]	113.2 ± 11.4	114.2 ± 10.5	.10[f]
Parental education (median)[c]	5.0	6.0	.55[g]
Diagnosis			
Schizophrenia (%)	32 (74.4%)		
Schizophreniform disorder (%)	11 (25.6%)		
Antipsychotic medication			
Atypical (%)	30 (69.8%)		
Typical (%)	5 (11.6%)		
No (%)	8 (18.6%)		
Anticholinergic medication (%)[d]	7 (17.1%)		

Note. Parental education is highest level of education attained by the parents of a participant as measured by a Dutch scoring system (Verhage, 1983), ranging from unfinished primary education (level 1) to university education (level 7).
[a]Based on 32 first-episode patients and 30 normal comparison participants.
[b]Based on 20 first-episode patients and 27 normal comparison participants.
[c]Based on 42 normal comparison participants.
[d]Based on 42 first-episode patients.
[e]Fisher's exact test.
[f]Independent-samples *t* test.
[g]Mann-Whitney *U* test.

The VLGT involves the verbal presentation of a "shopping" list (List A) over five immediate recall trials. List A consists of four words from each of four semantic categories (i.e., fruit, spices, clothing, and tools). Examinees are not informed of this latent semantic structure of the list. Following the five consecutive learning trials, a second list of 16 words (List B) is presented for one trial. This interference list also consists of words from four different semantic categories, of which two overlap with the categories of List A. Following recall of List B, short delay free and cued recall of List A are tested. In the cued recall procedure examinees are asked to name items belonging to the four categories, with the names of the categories serving as cues. After an interval of 20 min, recall of List A is tested again in a free and cued recall procedure (long delay). Finally, a recognition trial is presented, in which examinees have to discriminate List A words from 28 distracters.

We were primarily interested in assessing memory components which have been shown to be impaired in cortical and subcortical dementia, including immediate free recall of information, retention of previously learned information over a delay, ability to take advantage of provided semantic cues or of a recognition format, and ability to group words into semantically related categories (Table 2).

Statistical analysis

Analyses were carried out with SPSS (version 9.0 for Windows). Tests were two-sided and significance was accepted at the 5% level. Between-group differences were examined using *t* tests. In case of non-normality and/or inequality of variances the nonparametric Mann-Whitney *U* Test was used. Effect sizes were calculated on the basis of Cohen's (1988) *d* for each comparison, however, only medium ($.5 \leq d < .8$) and large ($d \geq .8$) effects were reported.

Results

Scores on the selected VLGT measures are shown in Table 2.

The first-episode patients were impaired on both immediate free recall [$t(84) = -7.6$, $p < .001$; $d = 1.6$] and use of semantic clustering [$t(84) = -3.8$, $p < .001$; $d = .8$] relative to the normal comparison participants. Additionally, the patients showed impaired retention of information over a delay in comparison with the comparison participants ($Z = -2.6$, $p < .01$; $d = .5$). Furthermore, the effect on performance of semantic cueing and of a recognition format did not differ significantly between patients and comparison participants.

Discussion

The finding of impaired retention, together with the finding that retrieval aids did not improve recall disproportionately, suggests that a primary storage deficit rather than retrieval deficit underlies the verbal memory impairment shown by schizophrenic patients at the onset of the disease. The

https://doi.org/10.1017/S1355617705050162 Published online by Cambridge University Press

Table 2. Group scores on the VLGT measures

Memory function and measure	First-episode patients ($N = 43$)	Normal comparison participants ($N = 43$)	p
Immediate free recall			
Total Recall List A, Trials 1–5 ($M \pm SD$)	45.7 ± 8.7	59.1 ± 7.7	<.001[a]
Retention of information over a delay			
(Long Delay Free Recall − Recall List A, Trial 5)/ (Recall List A, Trial 5) × 100 ($M \pm SD$)	−12.6 ± 21.2	−3.7 ± 11.8	<.01[b]
Response to semantic cueing			
(Long Delay Cued Recall − Long Delay Free Recall)/ (Long Delay Free Recall) × 100 ($M \pm SD$)	7.5 ± 18.1	4.3 ± 10.2	.24[b]
Response to recognition format			
Z-score (Discriminability Index) − *Z*-score (Long Delay Free Recall) ($M \pm SD$)	.1 ± 1.4	.0 ± .7	.46[b]
Use of semantic organizational strategy			
Semantic Clustering ($M \pm SD$)	1.6 ± .7	2.2 ± .8	<.001[a]

Note. To compare performances on Discriminability Index and Long Delay Free Recall, *Z* scores were calculated using mean and standard deviation of the normal comparison group. VLGT = Verbale Leer en Geheugen Test.
[a]Independent-samples *t* test.
[b]Mann-Whitney *U* test.

observed memory profile most resembles the profile seen in cortical dementia, which may suggest primary cortical, specifically temporal lobe, involvement. This suggestion is tentatively made because the first-episode patients overall did not display the same magnitude of memory impairment as seen in patients with AD. For instance, the first-episode patients' scores on immediate recall, long delay free recall, and discriminability were ~1.5 standard deviations below those of the normal comparison participants, while AD patients have been found to score 3.2–4.0 standard deviations below average on these measures (Delis et al., 1991). Additionally, it is yet unknown to what extent the neuroanatomical correlates of memory impairments in (first-episode) schizophrenia are comparable to those supposed to be involved in cortical (and subcortical) dementia.

In contrast to our findings, Mohamed et al. (1999) found no evidence of a storage deficit in first-episode patients. Additionally, the cortical memory profile shown by our first-episode patients differs from the generally observed subcortical profile in schizophrenic patients with mixed chronicity. A possible explanation for these discrepancies concerns the possibility of heterogeneity in verbal memory performance in schizophrenia. Utilizing a discriminant function analysis (Massman et al., 1992) that differentiates HD patients, AD patients, and normal comparison participants on the basis of CVLT performance, Paulsen et al. (1995) found that most schizophrenic patients (50%) in their large sample were classified as having a subcortical memory profile; 15% of the patients were classified as having a cortical memory profile and the remaining 35% presented a normal profile. This finding does not necessarily implicate that three such subgroups are present in a schizophrenia sample, because the discriminant formulas forced the patients into this classification. However, Turetsky et al. (2002) provided additional support for the validity of the cortical-subcortical distinction in schizophrenia. Importantly, they cluster analyzed CVLT performance to categorize patients into subgroups and, thus, used no *a priori* constraints to force the patients into the expected memory profiles. Yet, three subgroups of schizophrenic patients were found that exhibited a subcortical, cortical, and normal memory profile, respectively. Furthermore, cortical patients tended to have structural and functional pathology in the temporal lobe, while temporal lobe pathology seemed relatively absent in subcortical patients. These latter patients exhibited ventricular enlargement and isolated frontal gray matter reduction, which may suggest frontal-striatal pathology.

Although our first-episode patients exhibited a verbal memory profile suggestive of a primary storage deficit, previous research suggests that heterogeneous memory performance within our patient sample might exist. Accordingly, by examining the total group of patients, it is possible that a relatively large subgroup displaying a cortical memory profile might have masked the presence of distinct memory profiles exhibited by other subgroups. The high degree of variability exhibited by the first-episode patients on the measures that are supposed to differentiate between cortical and subcortical memory profiles (i.e., retention of information over a delay, responses to semantic cueing and a recognition format) is consistent with the possibility of heterogeneity.

https://doi.org/10.1017/S1355617705050162 Published online by Cambridge University Press

PART 2

Are there Heterogeneous Profiles of Verbal Memory Performance Among First-Episode Patients with Schizophrenia?

To further elucidate the unexpected findings, we explored whether homogeneous patient subgroups could be identified by cluster analyzing the VLGT scores of the same first-episode patients. It was hypothesized that three meaningful patient subgroups would be identified: a subgroup that displayed a cortical memory profile, a subgroup that exhibited a subcortical memory profile, and a subgroup that showed intact memory performance.

Cluster analysis has been criticized because of its alleged tendency to generate bias due to the neuropsychological test(s) selected (Seaton et al., 2001). Therefore, the subgroups of first-episode patients derived from cluster analyzing verbal memory performance were validated by comparing their verbal fluency performance. The rationale for this was three-fold. Firstly, like verbal memory performance, assessment of verbal fluency performance has been found to be useful in discriminating subcortical dementia from cortical dementia (Monsch et al., 1994; Rosser & Hodges, 1994). That is, AD patients tend to show a disproportionately greater deficit on category fluency in comparison with letter fluency, while HD patients tend to exhibit equally severe impairment on both types of fluency tasks. It has been proposed that the differential deficit in AD reflects abnormalities in the semantic store, while the global deficit seen in HD reflects impaired initiation and retrieval strategies (Rosser & Hodges, 1994). Given the differences in pathology in AD and HD and findings in other neuropsychiatric populations, it has been suggested that deficits in semantic memory reflect temporal neocortical damage and that impaired retrieval mechanisms reflect frontal-striatal dysfunction (Kremen et al., 2003). Neuroimaging studies involving healthy controls are consistent with this suggestion (Gourovitch et al., 2000; Mummery et al., 1996; Paulesu et al., 1997; Pujol et al., 1996). Secondly, verbal fluency is substantially impaired in schizophrenic patients relative to controls (Heinrichs & Zakzanis, 1998). Thirdly, evidence has been found for heterogeneity among schizophrenic patients regarding letter *versus* category fluency performance (Elvevåg et al., 2001; Gourovitch et al., 1996; Kremen et al., 2003). In sum, like verbal memory profiles, patterns of fluency performance have been associated with a retrieval (subcortical/frontal-striatal) deficit–storage (cortical/temporal) deficit distinction and performance heterogeneity in patients with schizophrenia. If cluster analysis would classify our first-episode patients into three subgroups that exhibited a cortical, a subcortical, and a normal memory profile, respectively, we would expect these subgroups to exhibit similar profiles on the basis of their fluency performance.

Methods

Verbal fluency

Verbal fluency tasks were administered as part of the comprehensive neuropsychological battery. Letter fluency was assessed by asking participants to produce as many different words as they could beginning with the letters N and A within one minute. The average number of words for the two letters was used as the score on letter fluency. Category fluency was assessed by asking participants to produce as many exemplars as possible belonging to the "animal" category within one minute.

Statistical analysis

Analyses were carried out with SPSS (version 9.0 for Windows). The five VLGT variables were transformed into standardized *Z* scores before analysis. Ward's method of hierarchical cluster analysis with squared Euclidean distance as the similarity measure was used to determine the clusters. This is a frequently used and recommended cluster analytic method (Aldenderfer & Blashfield, 1984). A three-clusters solution was specified at onset. Importantly, no *a priori* constraints forced the patients into subgroups displaying the expected memory profiles. Rather, the classification of patients into subgroups was empirically determined by a cluster algorithm using the scores on VLGT indices. To evaluate the stability of the cluster solution obtained using Ward's method, the solution was compared to that derived from another algorithm, namely k-means cluster analysis using an iterative procedure. Similarity of the solutions was measured with Cohen's kappa.

To characterize the memory profiles of the clusters, analyses of variance (ANOVA's) comparing the clusters on the VLGT measures were performed. In addition, each cluster's performance was compared with that of the normal comparison group with *t* tests. In case of non-normality and/or inequality of variances, nonparametric equivalents (Kruskal-Wallis *H* or Mann-Whitney *U* Tests) were used.

Differences among the clusters on demographic, clinical (i.e., proportion of schizophreniform patients, proportion of patients taking antipsychotic and anticholinergic medication), and fluency variables were examined to evaluate the external validity of the cluster solution. ANOVA's or Kruskal-Wallis *H* Tests were used for comparisons involving continuous variables and chi-square tests for comparisons concerning categorical variables. The premorbid and current IQ estimates of the three clusters were presented for descriptive purposes only and not statistically compared, because sample sizes were too small for meaningful analysis. Verbal fluency performance was examined with a 3 (cluster: cluster 1, cluster 2, cluster 3) $\times$ 2 (fluency type: letter fluency, category fluency) repeated-measures multivariate analysis of variance (MANOVA). A significant interaction for cluster $\times$ fluency type would indicate

https://doi.org/10.1017/S1355617705050162 Published online by Cambridge University Press

that a least one cluster showed a different pattern of fluency performance.

All tests were two-sided and significance was accepted at the 5% level. *Post-hoc* comparisons were conducted with Bonferroni adjusted level of significance. Since comparisons between the clusters and normal comparisons and between-clusters comparisons were influenced by unequal and small sample sizes, effect sizes were calculated using Cohen's *d* to make understanding of the comparisons involving numerical variables easier. Only medium and large effect sizes were reported.

Results

Cluster analysis using Ward's method yielded three clusters consisting of 13, 22, and 8 first-episode patients, respectively. Only one case that was assigned to Cluster 1 by Ward's method was grouped into Cluster 2 by the *k*-means cluster analysis, resulting in almost perfect agreement between the two cluster solutions (kappa of .9).

Overall VLGT performance of the clusters and comparison group is shown in Figure 1. Figure 2 shows the performance of the clusters on the VLGT measures *Z*-scored relative to the normal comparison group.

Between-cluster analyses revealed that Cluster 3 performed better than either Cluster 1 ($p < .001$; $d = 2.9$) or Cluster 2 ($p < .001$; $d = 2.5$) on immediate free recall. Furthermore, Clusters 1 and 2 were not equally impaired on free recall, with Cluster 2 showing better performance ($p < .05$; $d = 1.0$). Cluster 3 also had a higher score on the semantic clustering index than either Cluster 1 ($p < .001$; $d = 3.2$) or Cluster 2 ($p < .001$; $d = 2.9$). A large effect size (.8) indicated that Cluster 2 had a better semantic clustering score than Cluster 1. Probably, this comparison was not significant ($p = .08$) due to unequal and small sample sizes. Cluster 2 performed worse on retention of information than either Cluster 1 ($p < .05$; $d = 1.2$) or Cluster 3 ($p < .005$; $d = 1.7$). No significant between-cluster differences were found regarding the effect of semantic cueing, although a medium effect size (.6) suggested that Cluster 2 benefited more from this retrieval aid than Cluster 1. Cluster 1 performed worse than either Cluster 2 ($p < .001$; $d = 2.5$) or Cluster 3 ($p < .005$; $d = 2.0$) on recognition relative to delayed free recall. A large effect size (1.0) indicated that Cluster 2 benefited more from the recognition procedure than Cluster 3, although the comparison between Cluster 2 and 3 was not significant. Further analyses (data not shown) revealed that Cluster 1 made more false positive errors during recognition than either Cluster 2 or Cluster 3 ($p < .001$), while there were no significant between-cluster differences regarding number of false negative errors and response bias.

Cluster 1 performed worse than the comparison group on immediate free recall ($p < .001$; $d = 2.6$) and on the semantic clustering index ($p < .001$; $d = 1.7$). The two groups did not differ significantly regarding retention of information and response to semantic cueing. However, Cluster 1's performance deteriorated on the recognition format relative to

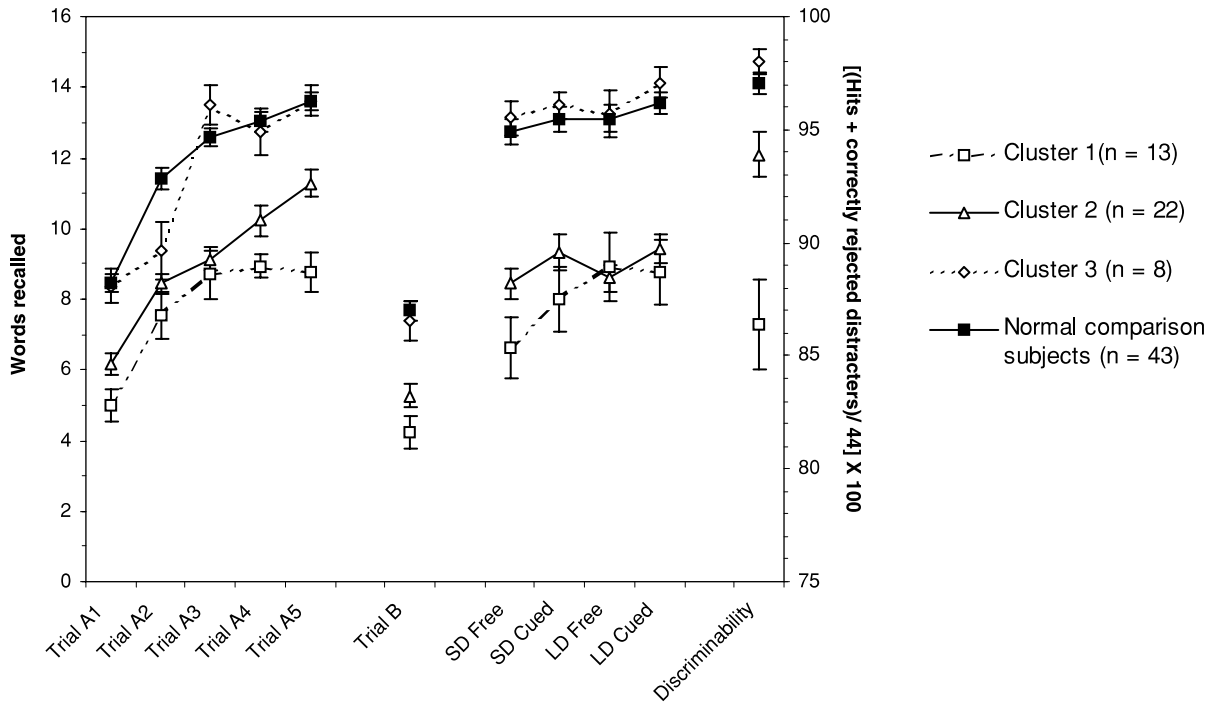


Fig. 1. Mean (± *SEM*) performance on the Dutch version of the California Verbal Learning Test of the three clusters obtained within the first-episode group and the normal comparison group. SD = short delay recall. LD = long delay recall.

https://doi.org/10.1017/S1355617705050162 Published online by Cambridge University Press

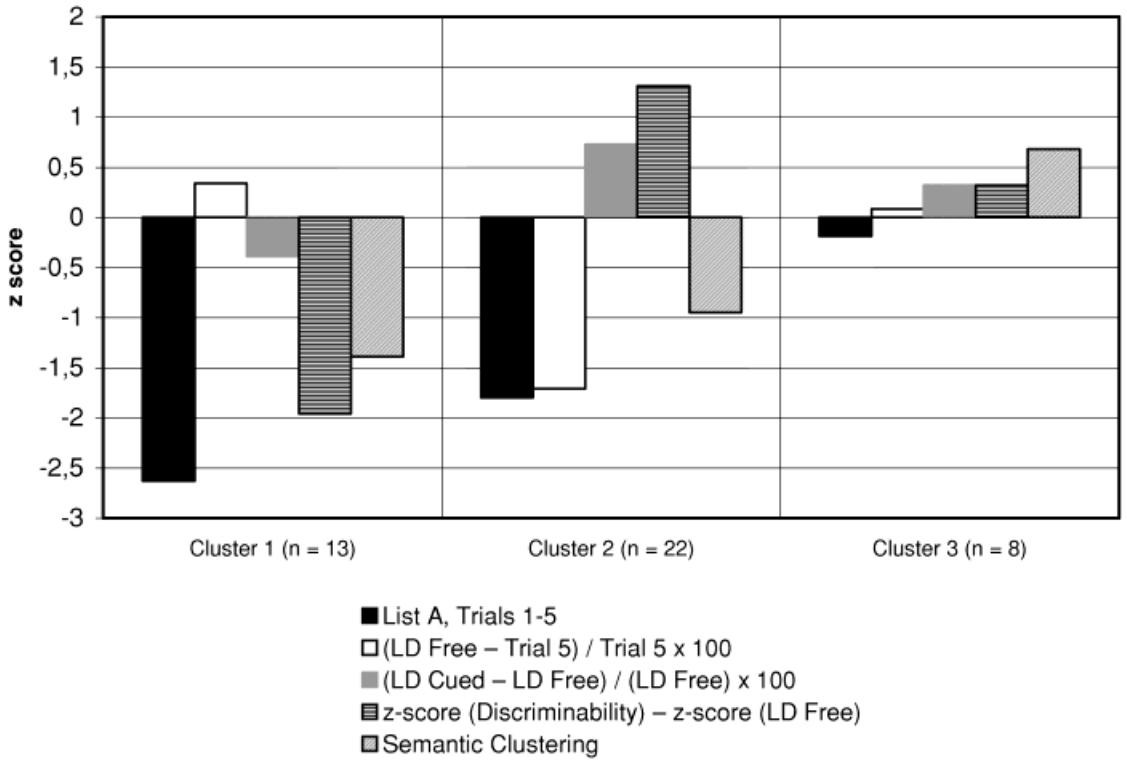


Fig. 2. Mean *Z* scores on measures of the Dutch version of the California Verbal Learning Test for the three clusters obtained within the first-episode group. Positive *Z* scores indicate better performance, and negative *Z* scores indicate poorer performance, relative to the normal comparison group. LD = long delay recall.

the comparison group ($p < .001$; $d = 1.7$). Cluster 2 differed from the comparison group on all five VLGT measures. Cluster 2 performed worse on immediate free recall ($p < .001$; $d = 2.1$), on the semantic clustering index ($p < .001$; $d = 1.3$), and on retention of information ($p < .001$; $d = 1.6$). However, Cluster 2's performance improved more when semantic cues ($p < .05$; $d = .5$) or a recognition format ($p < .001$; $d = 1.2$) were provided. The performance of Cluster 3 did not differ significantly from that of the comparison group on any of the selected VLGT measures. However, a large effect size (.9) suggested that Cluster 3 had a better semantic clustering.

No overall significant differences between the clusters were found for any of the demographic and clinical characteristics (Table 4). Noteworthy, the between-cluster comparisons for educational level and proportion of patients with schizophreniform disorder approached significance.

The MANOVA comparing the three clusters on verbal fluency performance revealed a nearly significant main effect for cluster [$F(2,39) = 3.2$, $p = .052$] (Table 3). *Post-hoc* analyses demonstrated that both Cluster 1 ($d = 1.6$) and Cluster 2 ($d = 1.1$) performed worse than Cluster 3 on letter fluency ($p < .05$). No significant between-cluster differences were found regarding category fluency. This negative finding probably reflected small and unequal sample sizes, because large effect sizes indicated that again both Cluster 1 ($d = .8$) and Cluster 2 ($d = .9$) performed worse than Cluster 3. More importantly, no significant interaction for cluster × fluency type was found [$F(2,39) = 0.1$, $p = .91$]. Performing separate MANOVA's involving each of the clusters and the normal comparison group also revealed no significant interaction effects. In addition, the effect sizes for letter and category fluency were very similar for each cluster-comparison group contrast (for Cluster 1: $d = 1.0$ and .9; for Cluster 2: $d = 1.0$ and 1.1), which suggested that neither letter nor category fluency was differentially impaired in one of the clusters relative to the comparison group. The fluency performance of Cluster 3 was similar to that of the normal comparison participants (letter fluency: $M = 14.2$, $SD = 4.5$; category fluency: $M = 25.8$, $SD = 4.3$).

Discussion

The patient subgroups formed on the basis of cluster analysis of verbal memory performance did not show the expected memory profiles. Although one first-episode subgroup did exhibit unimpaired memory performance, the other two subgroups exhibited profiles of memory impairments that did not conform neatly to the profiles associated with cortical (i.e., AD) and subcortical dementia (i.e., HD). Both memory-impaired subgroups exhibited impairments on immediate free recall and use of semantic clustering, with one subgroup showing more severe impairments relative to the other. The first-episode patients of the more severely impaired subgroup did not benefit from retrieval cues, whereas those of the less severely impaired sub-

https://doi.org/10.1017/S1355617705050162 Published online by Cambridge University Press

Table 3. Cluster scores on the VLGT and Verbal Fluency measures

	Cluster 1 ($N = 13$)	Cluster 2 ($N = 22$)	Cluster 3 ($N = 8$)	P
VLGT				
Immediate free recall				
Total List A, Trials 1–5 ($M \pm SD$)	38.9 ± 7.8	45.3 ± 5.2	57.6 ± 4.7	<.001[b]
Retention of information over a delay				
(Long Delay Free Recall − Recall List A, Trial 5)/ (Recall List A, Trial 5) × 100 ($M \pm SD$)	.3 ± 26.2	−23.9 ± 14.0	−2.7 ± 10.8	<.005[c]
Response to semantic cueing				
(Long Delay Cued Recall − Long Delay Free Recall)/ (Long Delay Free) × 100 ($M \pm SD$)	.3 ± 21.5	11.7 ± 17.4	7.5 ± 10.2	.30[c]
Response to recognition format				
Z-score (Discriminability Index) − *Z*-score (Long Delay Free Recall) ($M \pm SD$)	−1.5 ± 1.0	1.0 ± .9	.2 ± .6	<.001[b]
Use of semantic organizational strategy				
Semantic Clustering ($M \pm SD$)	1.1 ± .5	1.4 ± .4	2.8 ± .5	<.001[b]
Verbal fluency				
Letter fluency (average of N and A) ($M \pm SD$)[a]	10.0 ± 2.9	10.5 ± 4.2	14.4 ± 2.6[a]	<.05[b]
Category fluency (animals) ($M \pm SD$)[a]	20.4 ± 6.8	20.9 ± 5.2	25.6 ± 5.7[a]	.14[b]

Note. VLGT = Verbale Leer en Geheugen Test.
[a]Based on 7 first-episode patients.
[b]ANOVA.
[c]Kruskal-Wallis *H* test.

group did disproportionately. Additionally, the performance of the more impaired subgroup became even worse on recognition testing relative to free recall, a feature also observed in cortical dementia (Delis et al., 1991; Massman et al., 1992). However, poor retention of information over a delay, the distinctive characteristic of a cortical memory profile, was not observed in this subgroup. In contrast, the less impaired subgroup did exhibit impaired

Table 4. Demographic and clinical characteristics of the three clusters

	Cluster 1 ($N = 13$)	Cluster 2 ($N = 22$)	Cluster 3 ($N = 8$)	p
Gender (male, %)	12 (92.3%)	17 (77.3%)	6 (75.0%)	.58[d]
Handedness (right, %)	11 (84.6%)	18 (81.8%)	8 (100%)	.46[d]
Age (years) ($M \pm SD$)	27.2 ± 4.8	24.3 ± 5.1	25.5 ± 5.0	.27[e]
Premorbid IQ estimate ($M \pm SD$)[a]	102.8 ± 5.9	104.7 ± 7.5	103.3 ± 1.5	
IQ estimate ($M \pm SD$)[b]	105.2 ± 11.5	114.4 ± 11.7	119.8 ± 4.2	
Education (median)	5.0	6.0	6.0	.074[f]
Parental education (median)	5.0	5.5	6.0	.18[f]
Diagnosis				
Schizophrenia (%)	12 (92.3%)	16 (72.7%)	4 (50.0%)	.080[d]
Schizophreniform disorder (%)	1 (7.7%)	6 (27.3%)	4 (50.0%)	
Antipsychotic medication				
Atypical (%)	9 (69.2%)	15 (68.2%)	6 (75.0%)	.86[d]
Typical (%)	2 (15.4%)	3 (13.6%)	0 (0.0%)	
No (%)	2 (15.4%)	4 (18.2%)	2 (25.0%)	
Anticholinergic medication (%)	3 (23.1%)	4 (19.0%)[c]	0 (0.0%)	.42[d]

Note. (Parental) education is highest level of education attained by (the parents of) a participant as measured by a Dutch scoring system (Verhage, 1983), ranging from unfinished primary education (level 1) to university education (level 7). VLGT = Verbale Leer en Geheugen Test.
[a]Based on 11 Cluster 1 patients, 17 Cluster 2 patients and 4 Cluster 3 patients; no statistical analysis was performed.
[b]Based on 5 Cluster 1 patients, 11 Cluster 2 patients and 4 Cluster patients; no statistical analysis was performed.
[c]Based on 21 patients.
[d]Chi-square test.
[e]ANOVA.
[f]Kruskal-Wallis *H* test.

https://doi.org/10.1017/S1355617705050162 Published online by Cambridge University Press

retention. Overall, the memory profile of the more impaired subgroup is suggestive of a prominent encoding deficit (as reflected by the deficient use of semantic clustering in combination with the normal retention of information and the impaired recognition performance), while the memory dysfunction of the less impaired subgroup seems to be caused by mild deficits in storage (as reflected by the impaired retention of information), retrieval (as reflected by the disproportionate improvement in performance when retrieval aids were provided), and potentially encoding (as reflected by the deficient use of semantic clustering). Hence, our data suggest that a cortical-subcortical distinction is not valid in first-episode patients with schizophrenia. This suggestion is consistent with the finding that the two memory-impaired subgroups showed the same pattern of letter *versus* category fluency performance. Both subgroups were equally impaired on letter and category fluency relative to controls. This performance pattern is found in patients with HD and has been proposed to reflect impaired initiation and retrieval strategies (Rosser & Hodges, 1994).

Apart from the cortical-subcortical distinction, the question remains whether the cluster-derived subgroups reflect meaningful distinct subtypes of first-episode patients. Supportive of the validity of the cluster solution is that the memory profiles shown by the subgroups appear to be different in nature. However, verbal fluency performance provided limited support for the validity of the cluster-derived subgroups. On a general level of functioning, the fluency performance of the three subgroups mimicked their memory performance. That is, the memory-unimpaired subgroup exhibited normal verbal fluency performance, while the two memory-impaired subgroups were impaired on verbal fluency. The two impaired subgroups did not differ significantly in terms of level and pattern of fluency impairment. This could suggest that the cluster solution merely reflected the specific memory abilities measured by the VLGT, and that the results cannot be generalized to other cognitive abilities. On the other hand, it is possible that verbal fluency is a less sensitive measure than the VLGT in detecting subtle cognitive differences between subgroups of first-episode patients. Further external validity analyses suggested that clinical and demographic characteristics were not associated with the subgroup classification. However, the relatively small subgroup sizes may have prevented meaningful effects from obtaining statistical significance in some cases. In addition, inspection of the subgroups' mean scores on the current IQ estimate may suggest a possible relationship between current intellectual functioning and subgroup assignment. That is, the more severely memory-impaired subgroup exhibited a lower current IQ estimate than the less impaired subgroup, which in turn showed a lower current IQ estimate than the unimpaired subgroup. Although the suggestion of an influence of current intellectual function is speculative, as it is based on mean scores of subsets of small subgroups, it may point to the possibility that the categorization into subgroups was mainly determined by level of general impairment (Seaton et al., 2001). If so, then the subgroups could simply reflect a continuum of memory impairment, rather than reflecting meaningful subtypes of first-episode patients. The mean scores at least suggest that it may have been important to adjust for current IQ.

Our negative findings regarding first-episode subgroups showing cortical and subcortical memory profiles contrast with the results of Turetsky et al. (2002), who did find that cluster analyzing verbal memory performance formed two schizophrenic subgroups that exhibited memory profiles and brain abnormalities consistent with a cortical (temporal)–subcortical (frontal-striatal) distinction. Additionally, these negative findings preclude the unexpected cortical memory profile shown by our total group of first-episode patients to be explained in terms of a dominant influence of a relatively large cortical subgroup. It is unclear what causes the inconsistency with previous research. A possibility is that there is a "moderating" role of assessing schizophrenic patients during the first stage of illness when they are recovering from the acute period (Rund, 1998). Longitudinal studies have suggested that significant aspects of the cognitive impairment shown by schizophrenic patients during the early course of illness tend to improve over time (Hoff et al., 1999; Rund, 1998). Such cognitive improvements have been linked to improvements in symptomatology associated with long-term antipsychotic treatment (Gold et al., 1999; Hoff et al., 1999). As far as this cognitive recovery relates to specific aspects of memory dysfunction, as has been found in some studies (Gold et al., 1999; Sweeney et al., 1991), it is possible that cluster analyzing verbal memory performance will give different results depending on whether one examines scores of first-episode schizophrenic patients or patients later in the course of illness (as in the study by Turetsky et al., i.e., the mean illness duration of their patients was approximately 7 years). Likewise, possible improvements in memory functions after the initial onset of illness may provide an explanation for the discrepancy between the cortical memory profile shown by our first-episode patients and the generally observed subcortical profile in chronic schizophrenic patients. Longitudinal follow-up investigation of the memory performance of large samples of first-episode patients is needed to determine whether memory profiles change during the course of illness. In this regard, it is interesting that Harvey et al. (2002) found little temporal stability for the memory-based subcortical classification of elderly schizophrenic patients, thereby casting doubt on the assumption that cortical *versus* subcortical memory profiles in schizophrenia reflect stable brain abnormalities.

Our findings should be considered in the context of some methodological limitations. Firstly, the analyses of the between-subgroup differences were hampered by small and unequal sample sizes and therefore the results should be considered preliminary. Secondly, our study sample included first-episode patients with schizophreniform disorder. It is possible that some of these patients may not go on to develop full-blown schizophrenia. If so, between-subgroup comparisons could have been confounded by the nonequal distri-

https://doi.org/10.1017/S1355617705050162 Published online by Cambridge University Press

bution of schizophreniform patients among the subgroups. Thirdly, examination of the external validity of the cluster solution was limited. Data about symptom severity and symptom profiles as well as neuroanatomical and neurophysiological measures at the time of neuropsychological assessment were lacking. Also, a restricted set of two neuropsychological tests was examined. For instance, it is possible that between-group differences with respect to attentional ability correlated with the memory-based categorization.

In sum, the results of our study suggest that first-episode patients with schizophrenia or schizophreniform disorder show a profile of verbal memory impairments suggestive of a primary storage deficit. However, the apparently distinctive memory profiles shown by cluster-derived subgroups could indicate that combining the memory performance of first-episode patients may obscure meaningful heterogeneity in verbal memory performance. Further longitudinal research with larger samples of first-episode patients and extensive cluster validation is needed.

REFERENCES

Albus, M., Hubmann, W., Ehrenberg, Ch., Forcht, U., Mohr, F., Sobizack, N., Wahlheim, Ch., & Hecht, S. (1996). Neuropsychological impairment in first-episode and chronic schizophrenic patients. *European Archives of Psychiatry and Clinical Neuroscience*, *246*, 249–255.

Aldenderfer, M.S. & Blashfield, R.K. (1984). *Cluster analysis*. Newberry Park, California: Sage Publications.

Aleman, A., Hijman, R., De Haan, E.H.F., & Kahn, R.S. (1999). Memory impairment in schizophrenia: A meta-analysis. *American Journal of Psychiatry*, *156*, 1358–1366.

American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders*, (4th ed.). Washington, DC: American Psychiatric Association Press.

Andreasen, N.C. (1987). *Comprehensive Assessment of Symptoms and History (CASH)*. Iowa City, IA: University of Iowa.

Bilder, R.M., Goldman, R.S., Robinson, D., Reiter, G., Bell, L., Bates, J.A., Pappadopulos, E., Willson, D.F., Alvir, J.M.J., Woerner, M.G., Geisler, S., Kane, J.M., & Lieberman, J.A. (2000). Neuropsychology of first-episode schizophrenia: Initial characterization and clinical correlates. *American Journal of Psychiatry*, *157*, 549–559.

Binder, J., Albus, M., Hubmann, W., Scherer, J., Sobizack, N., Franz, U., Mohr, F., & Hecht, S. (1998). Neuropsychological impairment and psychopathology in first-episode schizophrenic patients related to the early course of illness. *European Archives of Psychiatry and Clinical Neuroscience*, *248*, 70–77.

Buchsbaum, M.S., Haier, R.J., Potkin, S.G., Nuechterlein, K., Bracha, H.S., Katz, M., Lohr, J., Wu, J., Lottenberg, S., Jerabek, P.A., Trenary, M., Tafalla, R., Reynolds, C., & Bunney, W.E. (1992). Frontostriatal disorder of cerebral metabolism in never-medicated schizophrenics. *Archives of General Psychiatry*, *49*, 935–942.

Butters, N., Delis, D.C., & Lucas, J.A. (1995). Clinical assessment of memory disorders in amnesia and dementia. *Annual Review of Psychology*, *46*, 493–523.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum.

Delis, D.C., Filoteo, J.V., Massman, P.J., Kaplan, E., & Kramer, J.L. (1994). The clinical assessment of memory disorders. In L.S. Cermark (Ed.), *Neuropsychological explorations of memory and cognition: Essays in honor of Nelson Butters* (pp. 223–238). New York: Plenum Press.

Delis, D.C., Kramer, J.H., Kaplan, E., & Ober, B.A. (1987). *California Verbal Learning Test: Manual*. San Antonio, TX: Psychological Corporation.

Delis, D.C., Massman, P.J., Butters, N., Salmon, D.P., Cermak, L.S., & Kramer, J.H. (1991). Profiles of demented and amnesic patients on the California Verbal Learning Test: Implications for the assessment of memory disorders. *Psychological Assessment*, *3*, 19–26.

Elvevåg, B., Weinstock, D.M., Akil, M., Kleinman, J.E., & Goldberg, T.E. (2001). A comparison of verbal fluency tasks in schizophrenic patients and normal controls. *Schizophrenia Research*, *51*, 119–126.

Fletcher, P., McKenna, P.J., Friston, K.J., Frith, C.D., & Dolan, R.J. (1999). Abnormal cingulate modulation of fronto-temporal connectivity in schizophrenia. *Neuroimage*, *9*, 337–342.

Gold, S., Arndt, S., Nopoulos, P., O'Leary, D.S., & Andreasen, N.C. (1999). Longitudinal study of cognitive function in first-episode and recent-onset schizophrenia. *American Journal of Psychiatry*, *156*, 1342–1348.

Gourovitch, M.L., Goldberg, T.E., & Weinberger, D.R. (1996). Verbal fluency deficits in patients with schizophrenia: Semantic fluency is differentially impaired as compared with phonologic fluency. *Neuropsychology*, *10*, 573–577.

Gourovitch, M.L., Kirkby, B.S., Goldberg, T.E., Weinberger, D.R., Gold, J.M., Esposito, G., Van Horn, J.D., & Berman, K.F. (2000). A comparison of rCBF patterns during letter and semantic fluency. *Neuropsychology*, *14*, 353–360.

Harvey, P.D., Moriarty, P.J., Bowie, C., Friedman, J.I., Parrella, M., White, L., & Davis, K.L. (2002). Cortical and subcortical cognitive deficits in schizophrenia: convergence of classifications based on language and memory skill areas. *Journal of Clinical and Experimental Neuropsychology*, *24*, 55–66.

Heinrichs, R.W. & Zakzanis, K.K. (1998). Neurocognitive deficit in schizophrenia: A quantitative review of the evidence. *Neuropsychology*, *12*, 426–445.

Hoff, A.L., Riordan, H., O'Donnell, D.W., Morris, L., & DeLisi, L.E. (1992). Neuropsychological functioning of first-episode schizophreniform patients. *American Journal of Psychiatry*, *149*, 898–903.

Hoff, A.L., Sakuma, M., Wieneke, M., Horon, R., Kushner, M., & DeLisi, L.E. (1999). Longitudinal neuropsychological follow-up study of patients with first-episode schizophrenia. *American Journal of Psychiatry*, *156*, 1336–1341.

Keshavan, M.S., Rosenberg, D., Sweeney, J.A., & Pettegrew, J.W. (1998). Decreased caudate volume in neuroleptic-naive psychotic patients. *American Journal of Psychiatry*, *155*, 774–778.

Kremen, W.S., Seidman, L.J., Faraone, S.V., & Tsuang, M.T. (2003). Is there disproportionate impairment in semantic or phonemic fluency in schizophrenia? *Journal of the International Neuropsychological Society*, *9*, 79–88.

Manoach, D.S., Gollub, R.L., Benson, E.S., Searl, M.M., Goff, D.C., Halpern, E., Saper, C.B., & Rauch, S.L. (2000). Schizophrenic subjects show aberrant fMRI activation of dorsolateral prefrontal cortex and basal ganglia during working memory performance. *Biological Psychiatry*, *48*, 99–109.

Massman, P.J., Delis, D.C., Butters, N., Dupont, R.M., & Gillin, J.C. (1992). The subcortical dysfunction hypothesis of mem-

https://doi.org/10.1017/S1355617705050162 Published online by Cambridge University Press

ory deficits in depression: Neuropsychological validation in a subgroup of patients. *Journal of Clinical and Experimental Neuropsychology*, *14*, 687–706.

Meyer-Lindenberg, A., Poline, J.-B., Kohn, P.D., Holt, J.L., Egan, M.F., Weinberger, D.R., & Berman, K.F. (2001). Evidence for abnormal cortical functional connectivity during working memory in schizophrenia. *American Journal of Psychiatry*, *158*, 1809–1817.

Mohamed, S., Paulsen, J.S., O'Leary, D., Arndt, S., & Andreasen, N. (1999). Generalized cognitive deficits in schizophrenia. *Archives of General Psychiatry*, *56*, 749–754.

Monsch, A.U., Bondi, M.W., Butters, N., Paulsen, J.S., Salmon, D.P., Brugger, P., & Swenson, M.R. (1994). A comparison of category and letter fluency in Alzheimer's disease and Huntington's disease. *Neuropsychology*, *8*, 25–30.

Mulder, J.L., Dekker, R., & Dekker, P.H. (1996). *Verbale Leer en Geheugen Test: Handleiding*. Lisse: Swets & Zeitlinger.

Mummery, C.J., Patterson, K., Hodges, J.R., & Wise, R.J. (1996). Generating 'tiger' as an animal name or a word beginning with T: Differences in brain activation. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *263*, 989–995.

Nelson, M.D., Saykin, A.J., Flashman, L.A., & Riordan, H.J. (1998). Hippocampal volume reduction in schizophrenia as assessed by magnetic resonance imaging: A meta-analytic study. *Archives of General Psychiatry*, *55*, 433–440.

Paulesu, E., Goldacre, B., Scifo, P., Cappa, S.F., Gilardi, M.C., Castiglioni, I., Perani, D., & Fazio, F. (1997). Functional heterogeneity of left inferior frontal cortex as revealed by fMRI. *Neuroreport*, *8*, 2011–2017.

Paulsen, J.S., Heaton, R.K., Sadek, J.R., Perry, W., Delis, D.C., Braff, D., Kuck, J., Zisook, S., & Jeste, D.V. (1995). The nature of learning and memory impairments in schizophrenia. *Journal of the International Neuropsychological Society*, *1*, 88–99.

Pujol, J., Vendrell, P., Deus, J., Kulisevsky, J., Marti-Vilalta, J.L., Garcia, C., Junque, C., & Capdevila, A. (1996). Frontal lobe activation during word generation studied by functional MRI. *Acta Neurologica Scandinavica*, *63*, 109–121.

Resnick, S.M. (1992). Matching for education in studies of schizophrenia [letter; comment]. *Archives of General Psychiatry*, *49*, 246.

Rosser, A. & Hodges, J.R. (1994). Initial letter and semantic category fluency in Alzheimer's disease, Huntington's disease, and progressive supranuclear palsy. *Journal of Neurology, Neurosurgery and Psychiatry*, *57*, 1389–1394.

Rund, B.R. (1998). A review of longitudinal studies of cognitive functions in schizophrenia patients. *Schizophrenia Bulletin*, *24*, 425–435.

Savage, C.R. (1997). Neuropsychology of subcortical dementias. *Psychiatric Clinics of North America*, *20*, 911–931.

Saykin, A.J., Gur, R.C., Gur, R.E., Mozley, P.D., Mozley, L.H., Resnick, S.M., Kester, D.B., & Stafiniak, P. (1991). Neuropsychological function in schizophrenia. Selective impairment in memory and learning. *Archives of General Psychiatry*, *48*, 618–624.

Saykin, A.J., Shtasel, D.L., Gur, R.E., Kester, D.B., Mozley, L.H., Stafiniak, P., & Gur, R.C. (1994). Neuropsychological deficits in neuroleptic naive patients with first-episode schizophrenia. *Archives of General Psychiatry*, *51*, 124–131.

Schmand, B., Lindeboom, J., & Van Harskamp, F. (1992). *Nederlandse Leestest voor Volwassenen*. Lisse: Swets & Zeitlinger.

Seaton, B.E., Goldstein, G., & Allen, D.N. (2001). Sources of heterogeneity in schizophrenia: The role of neuropsychological functioning. *Neuropsychology Review*, *11*, 45–67.

Silverstein, A.B. (1982). Two- and four-subtest short forms of the Wechsler Adult Intelligence Scale–Revised. *Journal of Consulting and Clinical Psychology*, *50*, 415–418.

Sitskoorn, M.M., Nuyen, J., Appels, M.C.M., Van der Wee, N.J.A., & Kahn, R.S. (2002). Release from proactive inhibition in schizophrenia and its potential as a genotypic marker. *Journal of Clinical and Experimental Neuropsychology*, *24*, 67–81.

Sweeney, J.A., Haas, G.L., Keilp, J.G., & Long, M. (1991). Evaluation of the stability of neuropsychological functioning after acute episodes of schizophrenia: One-year followup study. *Psychiatry Research*, *38*, 63–76

Turetsky, B.I., Moberg, P.J., Mozley, L.H., Moelter, S.T., Agrin, R.N., Gur, R.C., & Gur, R.E. (2002). Memory-delineated subtypes of schizophrenia: Relationship to clinical, neuroanatomical, and neurophysiological measures. *Neuropsychology*, *16*, 481–490.

Verhage. (1983). *Revised scoring method for educational level*. University Hospital Groningen: Department of Neuropsychology.

Zakzanis, K.K. & Heinrichs, R.W. (1999). Schizophrenia and the frontal brain: A quantitative review. *Journal of the International Neuropsychological Society*, *5*, 556–566.

Zakzanis, K.K., Poulin, P., Hansen, K.T., & Jolic, D. (2000). Searching the schizophrenic brain for temporal lobe deficits: A systematic review and meta-analysis. *Psychological Medicine*, *30*, 491–504.

https://doi.org/10.1017/S1355617705050162 Published online by Cambridge University Press