Psychological Medicine, 2002, **32**, 909–917. © 2002 Cambridge University Press
DOI: 10.1017/S0033291702005597 Printed in the United Kingdom

Semantic clustering in verbal fluency: schizophrenic patients *versus* control participants

B. ELVEVÅG,[1] J. E. FISHER, J. M. GURD AND T. E. GOLDBERG

From the Clinical Brain Disorders Branch, National Institute of Mental Health/National Institutes of Health, Bethesda, MD, USA; and Neuropsychology Unit, University Department of Clinical Neurology, Oxford

ABSTRACT

Background. Schizophrenic patients generate fewer words than healthy controls during verbal fluency tasks. The structure of output may explain why patients generate fewer exemplars.

Methods. Twenty-four healthy controls and 24 patients with schizophrenia participated in six, 3 min semantic fluency tasks. In a subsequent session, participants were given cards, each printed with one of their own words generated from previous fluency tasks. Participants were to sort the cards into categories (e.g. subcategories of 'animals'), thus defining their own semantic subcategories of words, and thereby eliminating experimenter assumptions about word relatedness. These clusters were matched with fluency output of each participant. The time spent searching through semantic networks within clusters and switching to other clusters when locating and producing associated words were measured.

Results. Patients produced fewer words and spent more time switching to words within clusters and to different clusters than controls, but otherwise response profiles were similar. Although controls returned more frequently to clusters and consequently made more switches between these clusters than patients, this group difference disappeared when the total number of words produced was covaried.

Conclusions. Consistent with previous literature, patients produced fewer words and made more errors than controls. The absence of a group difference in number of different clusters or mean number of items per cluster suggests that patients are similar to controls with respect to number of ideas in their semantic network. Patients' longer between-cluster switching times indicate a general slowness that may be attributed to difficulties finding new words within a semantic field.

INTRODUCTION

There is a remarkable amount of structure to output of verbal fluency tasks in which participants are required to recall as many members of a category (e.g. 'animals') as possible. Bousfield & Sedgewick (1944) fitted a negatively accelerated asymptotic exponential function to the output in fluency tasks. Indeed, it has been noted that 'items are not produced uniformly in time but in spurts. The produced words are clustered in time' (Gruenewald & Lockhead, 1980, p. 229). Such free recall may involve searching for meanings, rather than for individual items, such that 'a cluster is those words that become available when a semantic field or meaning is accessed. … A cluster is a collection of produced terms associated or connected with a particular semantic field in memory, and may contain from one to very many words' (Gruenewald & Lockhead, 1980, p. 230). Gruenewald & Lockhead (1980) suggested that the search time to access semantic fields increases as the task duration increases (a function variously modelled as exponential – Bousfield & Sedgewick, 1944; or hyperbolic –

[1] Address for correspondence: Dr B. Elvevåg, Clinical Brain Disorders Branch, National Institute of Mental Health/National Institutes of Health, Building 10, Room 4S235, MSC 1379, Bethesda, MD 20892, USA.

https://doi.org/10.1017/S0033291702005597 Published online by Cambridge University Press

Bousfield *et al.* 1954). Once a field is assessed, and a cluster produced, the rate of production of associated words within clusters is rather constant.

From a series of 15 or 30 min data collection sessions in which participants named as many members of a certain category as possible, Gruenewald & Lockhead (1980) reported that the between-cluster inter-item intervals were greater than the within-cluster inter-item intervals, and that the number of clusters produced, decreased over time. Their work has provided the framework for the current study which examines temporal differences between these two components of the search process, operationalized in the current paper 'as between-cluster inter-item interval' times (search for the semantic field and its member word) and 'within-cluster inter-item interval' times (output of semantic associates). If participants returned to a cluster, it was scored as a repeat visit (to a previously accessed cluster). The assumption is that such measures 'may reflect the dependency of the search process on semantic structure' (Gruenewald & Lockhead, 1980, p. 237).

Tasks requiring search through semantic structure are of particular interest in the neuropsychology of schizophrenia, as the literature is replete with reports of semantic memory disorders in this patient group (Tamlyn *et al.* 1992; Chen *et al.* 1994; McKenna *et al.* 1994; Aloia *et al.* 1996; Goldberg *et al.* 1998). Our previous work has highlighted deficits in the search process through this type of memory (Gurd *et al.* 1997). One method with which to assess the organization of semantic memory employs verbal fluency tasks to explore the way in which words are clustered together, as well as the frequency with which switches to a new subcategory are made (e.g. in the category 'animals', switching from the subcategory 'domestic' to 'farm' etc.). It is well documented that patients with schizophrenia produce significantly fewer words on semantic fluency tasks (Allen *et al.* 1993; Gourovitch *et al.* 1996; Joyce *et al.* 1996; Granholm *et al.* 1998; Elvevåg *et al.* 2001). Examining the nature of clustering and switching deficits on output from verbal fluency tasks has been applied to a wide variety of patient populations (e.g. patients with Parkinson's disease – Gurd, 1993; patients with schizophrenia – Robert *et al.* 1998; and frontal-lobe *versus* temporal-lobe lesions – Troyer *et al.* 1998), as well as younger *versus* older healthy adults (Troyer *et al.* 1997). In the current study on patients with schizophrenia, we had two aims: the first was to have participants delineate their own semantic boundaries, and the second was to examine the temporal aspects of clustering in output from verbal fluency tasks.

A recent study by Robert *et al.* (1998) compared a large group of schizophrenic patients ($N = 78$) with control participants ($N = 62$) on 2 min fluency tasks, employing a method similar to that of Troyer and colleagues. They concluded that patients had significant impairments in switching on phonological fluency tasks (letters 'P' and 'R'), and problems in both switching and clustering during semantic fluency task (categories 'animals' and 'fruits'). We previously employed the method of Troyer *et al.* (1997, 1998) to examine clustering and switching for the category 'animals' in a 3 min fluency task (on 3 consecutive days) in order to compare performance of 15 schizophrenic patients with 15 control participants (Elvevåg *et al.* unpublished data). Since patients with schizophrenia generally produce fewer fluency items than controls, they logically produce either fewer clusters, fewer items per cluster, fewer repeat visits per cluster, or all three. Adoption of this methodology was subjective and the results were not easy to interpret. In terms of number of clusters, the groups were comparable, except that patients produced the same number of clusters in all sessions, whereas controls produced substantially more in the third session. Concerning the number of words per cluster, the groups were also comparable. However, when examined as a function of session, controls' cluster sizes were smaller than patients' in the first session, and patients' larger than controls' in the third session. Thus, when collapsed across testing sessions, the groups were similar (Elvevåg *et al.* unpublished data). Moreover, the way in which these methods (Robert *et al.* 1998) defined 'clustering' and 'switching' is of crucial importance; we discuss this below.

First we address Robert *et al.*'s (1998) method of clustering. The experimenter defined *post hoc* the number of words related by cluster 'only when at least three consecutive words were semantically related' (p. 541; Robert *et al.* 1998). This process contrasts with a similar

https://doi.org/10.1017/S0033291702005597 Published online by Cambridge University Press

method used by Troyer *et al.* (1997, 1998) in which cluster size was determined by any two words co-occurring semantically, where 'cat' and 'dog' would be defined as one cluster with a size of 1 (i.e. number of words minus 1 = cluster size). Our view is that since patients often produce fewer words, they will by default generate fewer clusters if this method is employed. Participants, especially patients, might generate smaller clusters, with the risk that the cluster would not be counted in either method. For example, when scoring clusters in 'animals': Dog, cat, monkey, hen, rooster, goose, whale, fly, cockroach, black beetle, snake, pigeon, gull, nightingale, canary (p. 546 – Robert *et al.* 1998), despite the participant producing 15 words, they are recorded as producing four semantic clusters. Consequently, there are three 'dead words', namely words that are not included in the count of cluster size. Our view is that this exclusion unfairly biases patients who already are producing fewer words overall. Furthermore, the authors do not covary these indices with total verbal output. This is surprising since one would expect that with fewer words being generated, there would by default be less scope for clustering and switching (as indeed Robert *et al.* report for patients with schizophrenia).

Secondly, we address the method employed by Robert *et al.* (1998) to calculate switching. This analysis was accomplished by counting the number of transitions between clusters, including single words. Thus, in the example given above, the authors calculated switching as the total word production (15) minus the related words that formed clusters (12) plus the number of clusters (4) = 7. Again, by having these so-called 'dead-words' excluded from their method of calculating clusters, the number of switches relative to clusters would be inflated considerably.

To summarize, we argue that there is an inherent confound in the methods described above, given that the experimenters' judgement of cluster boundaries is pivotal. To quote Bousfield (1953): '… many of the temporally distinguishable groups of items revealed no obvious basis of relationship. For example, on one instance, *hawk* was followed by *sparrow*. According to our definition this pair might be a cluster since there are several ways in which they could be related. Some hawks prey on sparrows, and sparrowhawk is the name of a bird. *In this situation we cannot rely on the experimenters' subjective judgment, and we would prefer not to rely on the subject's introspections*' (p. 229; our italics). Moreover, one cannot assume *a priori*, either: (*i*) that patients' semantic structures match those of the experimenter's, or; (*ii*) that all patients share the same semantic category structure. On the contrary, it would be reasonable to remain agnostic on the first issue, and rely on the evidence, and secondly, to employ analysis strategies that are tailored to, and incorporate individual differences in semantic category structure in both patients and controls. Thus, in light of first-hand experience with the method adopted by Troyer *et al.* (1997, 1998) (which was quite similar to that employed by Robert *et al.* 1998), and its above mentioned difficulties, we adopted the following strategy for measuring semantic category structure.

The current study was designed to have participants delineate their own semantic category boundaries – in contrast to the methods employed by Troyer *et al.* (1997, 1998) and Robert *et al.* (1998). Our method is based upon that of Gurd (1993), in which participants performed 1 min fluency tasks, and then without warning were asked to sort their responses into semantic categories on a subsequent testing session. This approach of having participants themselves sort the words and therefore each word contributing in the semantic analysis therefore eliminates 'dead words' that may unfairly bias the results against the group which produces fewer words.

Based upon previous findings, we hypothesized that patients with schizophrenia would produce fewer words per category (e.g. 'animals') than control participants. More specifically, we predicted that patients would produce fewer clusters (e.g. fewer subcategories within the category 'animal') and fewer visits per cluster as well as fewer words per initial and repeat visits to a cluster. This prediction was based on the observation that patients with schizophrenia do not seem to search as extensively for category exemplars as do control participants. Consequently, we expected schizophrenic patients to make fewer switches. We also hypothesized that the within-cluster and between-cluster times would be longer for patients with schizophrenia, due to their gen-

https://doi.org/10.1017/S0033291702005597 Published online by Cambridge University Press

erally slow performance. Moreover, we expected these search times to be disproportionately long due to the postulated disorganization of slowed search through semantic networks in patients with schizophrenia.

METHOD

Baseline tests

Three baseline tests indexed intellectual function. The first was a test of reading proficiency, the Wide Range Achievement Test-Reading (WRAT-R) (Jastak & Wilkinson, 1984), which is used as an estimate of pre-morbid functioning. The second was a standard test of current non-verbal or fluid intelligence (g_f) – Cattell's Culture Fair Intelligence Test, Scale 2, Form A (Cattell, 1971; Institute for Personality and Ability Testing, 1973). The test involves timed spatial problem solving tasks. Participants were also evaluated on a short version of the Wechsler Adult Intelligence Scale-Revised (WAIS-R) (Wechsler, 1981), which is also used to estimate current functioning.

Participants

Twenty-four in-patients from the National Institute of Mental Health Neuropsychiatric Research Hospital participated in this study (for participant details, see Table 1). All patients fulfilled DSM-IV criteria for schizophrenia as determined by the Structured Clinical Interview for DSM-IV (SCID). Patients generally had multiple hospital admissions due to incomplete responses to conventional treatments. Twenty-four normal healthy control participants were recruited through the National Institutes of Health volunteer office. No participant, control or patient, with a history of substance abuse or known neurologic condition was included in this study. All participants had normal or corrected-to-normal vision. All control participants were paid for their participation and patients completed the study as part of their protocol for entering the hospital, and after informed consent had been obtained.

Table 1. *Characteristics of patient and control samples*

	Patients $N = 24$ (19M, 5F) Mean (S.D.)	Controls $N = 24$ (13M, 11F) Mean (S.D.)
Age (years)*	36·9 (9·4)	31·2 (9·3)
WRAT-R IQ*	98·9 (15·1)	109·4 (11·5)
WAIS-R IQ*	83·3 (16·3)	112·8 (17·0)
Cattell IQ*	95·5 (13·1)	111·7 (13·2)
Neuroleptic medication	23	
Clozapine/olanzapine	9	
Risperidone	7	
High potency drugs†	7	
Anticholinergics	5	
Adjunctives‡	5	

* $P < 0{\cdot}05$.
† Haloperidol, fluphenazine and loxapine.
‡ Lithium, depakote, sertraline, lorazepam, venlafaxine, clonazepam and buspirone.

Design and procedure

Participants were required to name continuously as many members of each specified semantic category (e.g. 'sports') as they could for a period of 3 min. Their output was recorded on a tape-recorder for subsequent analysis of the time between words. Participants were each given the categories 'sports', 'animals', 'furniture', 'vehicles', 'fruit' and then lastly 'vegetables'. On a subsequent testing session (mean of 13 days later for patients, and a mean of 40 days later for controls), participants were given a stack of cards on which the words that they had previously generated had been written (one per card). They were asked to sort the cards into subcategories of their own choosing. For example, in our study each 'sports' item produced during the verbal fluency task was printed on a separate file-card and given to the participant to sort into subcategories (according to meaning). Participants were told that there was no 'correct way' to execute this task, and there could be as many or as few subcategories, with as few or many items as they deemed appropriate (thus, items were not assessed as correct or incorrect of category membership). The number and size of the categories was not specified by the experimenter. No time-limit was imposed on this sorting task. When the participant had finished sorting, the experimenter recorded the subcategories, and the participant was given the next category to sort. This was done for each of the six categories. Each participant's fluency production was then subdivided into clusters, delineated according to their card-sort designations. In other words, cluster membership was determined by each

https://doi.org/10.1017/S0033291702005597 Published online by Cambridge University Press

participant's sorting task results. Hence, the experimenter did not have to make judgements about the relatedness of others' words. The experimenter counted the number of clusters, and the number of items within each cluster.

The clusters defined by sorting were recorded, and each of the fluency items was labelled according to the order in which the words had been generated. So, if the first item produced to the category 'animals' was 'cat', this was assigned to cluster A, and the first item produced of the next cluster (e.g. 'pig') was assigned to cluster B, and so on. Thus, by timing the verbal fluency tasks, it was possible to analyse the temporal aspects of clustering (Gurd, 1993; cf. Gruenewald & Lockhead, 1980), since fluency output was taped. In order to calculate the times between each word, we created a timing programme that recorded the time that elapsed between each word (via spacebar pressing). As the tape was reviewed, the spacebar was pressed each time a word occurred. This process was employed three times and the average of the three trials was calculated. The inter-item intervals for words within the same cluster were then averaged (within-cluster inter-item intervals), and the between-word times for words when switching between clusters were also averaged (between-cluster inter-item intervals). This is illustrated as follows. Assume that in the verbal fluency task (category 'sports'), the participant produced words in the following order: 'tennis', 'rugby', 'soccer', 'football', 'badminton' and 'squash' and subsequently sorted the cards so that 'tennis', 'badminton' and 'squash' were in one pile and 'football', 'soccer' and 'rugby' were in a separate pile. The resulting cluster sequence would be: 'tennis' (cluster A_1), 'rugby' (cluster B_1), 'soccer' (cluster B_1), 'football' (cluster B_1), 'badminton' (cluster A_2) and 'squash' (cluster A_2). This data can be divided into two different clusters (A and B), with one visit to cluster A, followed by three visits to cluster B, followed by two more visits to cluster A (A_1, B_1, B_1, B_1, A_2, A_2). The times from 'tennis' to 'rugby' and from 'football' to 'badminton' were taken and said to be a 'between-cluster' time, because the cluster being accessed is changing from cluster A to cluster B in the first instance, and from B to A subsequently. The times from 'rugby' to 'soccer', 'soccer' to 'football', as well as 'badminton' to 'squash' were measured and referred to as a 'within-cluster' time.

RESULTS

The data were analysed for each of the six categories (since category sizes vary). A repeated measures analysis of variance (ANOVA) was calculated with group (controls, patients) as the between-group factor and category (sports, animals, furniture, vehicles, fruit, vegetables) as the within-subject factor.

Is there a difference between patients and controls in terms of overall verbal fluency output?

Overall control participants produced significantly more words than schizophrenic patients ($F(1,46) = 37{\cdot}68$, $P < 0{\cdot}001$). There was a significant main effect of category since both groups produced a disproportionately large number of animal words ($F(5,230) = 56{\cdot}01$, $P < 0{\cdot}001$), and there was a significant group × category interaction ($F(5,230) = 8{\cdot}26$, $P < 0{\cdot}001$). An analysis excluding the number of errors made (e.g. repetitions and non-words) showed that the error rate was very low (a mean of correct responses across the six categories of 32·01 for controls and 18·22 for patients; $P < 0{\cdot}001$) (see Table 2).

Given that our patient and control groups differed in age and intelligence (IQ), the latter being an important aspect of the schizophrenia illness (e.g. Weickert *et al*. 2000), we performed covariate analyses (ANCOVAs) using total word

Table 2. *Comparison of patients and controls on measures derived from the verbal fluency task*

	Patients	Controls
Average across categories	Mean (S.D.)	Mean (S.D.)
Total words (mean/3 min)	19·56 (8·76)	32·67 (9·25)
Total words excl. errors	18·22 (8·27)	32·01 (9·05)
Different clusters, *N*	6·10 (3·65)	7·13 (3·17)
Items per cluster	4·18 (3·39)	5·29 (2·39)
Repeat cluster visits, *N*	10·68 (5·63)	16·78 (5·53)
Items per repeat cluster	2·06 (2·63)	2·52 (0·77)
Items in first cluster visit	2·92 (3·37)	2·54 (2·22)
Items in repeat cluster	1·85 (1·23)	2·00 (0·68)
Within-cluster time (s)	9·12 (6·81)	5·53 (5·85)
Between-cluster time (s)	11·59 (5·89)	7·83 (3·29)

https://doi.org/10.1017/S0033291702005597 Published online by Cambridge University Press

production as the dependent measure, with age and IQ scores (as measured by the WRAT-R, WAIS-R and the Cattell tests separately) as covariates. In all cases, the group difference in word production remained significant (all $P < 0{\cdot}001$). These ANCOVAs may be interpreted as suggesting that the total verbal output cannot be accounted for by differences in age or issues such as a general cognitive ability (i.e. differences in IQ). Similar findings, namely that reduced verbal fluency was not an artefact of lower IQ in the patient group, have been reported by Crawford *et al*. (1993).

Is there a group difference in the number and size of clusters produced?

As can be seen in Table 2, controls tended to produce more clusters than patients, but this was not significantly so ($F(1,46) = 1{\cdot}68$, $P = 0{\cdot}2$). There was a significant main effect of category, probably because of the large numbers of animals produced by participants as a whole ($F(5,230) = 5{\cdot}61$, $P < 0{\cdot}001$), but there was not a significant group × category interaction ($F(5,230) = 1{\cdot}78$, $P > 0{\cdot}1$).

Controls produced more items per cluster than patients, but this difference failed to reach significance ($F(1,46) = 2{\cdot}76$, $P = 0{\cdot}1$). There was a trend towards a significant main effect of category ($F(5,230) = 2{\cdot}03$, $P = 0{\cdot}08$), probably because of the large numbers of animals produced by controls especially. There was a trend towards a group × category interaction ($F(5,230) = 1{\cdot}97$, $P = 0{\cdot}08$).

Next we examined the number of visits (i.e. repeats) to a cluster. For example, output labelled as A_1, B_1, A_2, C_1 etc. would be counted as having two visits to cluster A. The controls produced significantly more such cluster visits than did the patients ($F(1,46) = 27{\cdot}11$, $P < 0{\cdot}001$). There was a significant main effect of category ($F(5,230) = 16{\cdot}42$, $P < 0{\cdot}0001$), but there was no significant group × category interaction ($F(5,230) = 2{\cdot}17$, $P > 0{\cdot}05$). (We do not present analyses examining the groups in terms of 'switches' between clusters since this index is clearly simply a function of the number of clusters and visits to these clusters.) In order to control for any unfair bias introduced by controls producing significantly more words in general, we recalculated the above analysis using total number of words as a covariate; importantly the main effect of group disappeared ($P > 0{\cdot}5$). (Since there was no significant main effects of the main measures of number and size of clusters and size of repeat visits to clusters, we did not perform covariate analyses with total output as a covariate. For the number of repeat visits to clusters there was a significant main effect of group, and so we did perform a covariate analysis.)

There was no significant difference between the groups in the size of these repeat visits to clusters ($F(1,46) = 1{\cdot}04$, $P > 0{\cdot}1$). There was a significant main effect of category ($F(5,230) = 2{\cdot}61$, $P < 0{\cdot}05$), but no significant group × category interaction ($F(5,230) = 1{\cdot}74$, $P > 0{\cdot}1$).

Is there a difference between patients and controls in terms of the number of items produced on their first cluster visit *versus* subsequent visits to a cluster?

An ANOVA calculated with group (controls, patients) as the between-group factor and first cluster *versus* the average of subsequent cluster visits (first cluster, subsequent cluster visits) and categories (sports, animals, furniture, vehicles, fruit, vegetables) as the within-subject factors did not find a significant main effect of group ($F(1,46) = 0{\cdot}16$, $P > 0{\cdot}5$), although there were significant main effects of first cluster *versus* subsequent cluster visits ($F(1,46) = 8{\cdot}40$, $P < 0{\cdot}01$), and of category ($F(5,230) = 3{\cdot}54$, $P < 0{\cdot}005$). There was not a significant interaction between group × first cluster *versus* subsequent visits to these clusters ($F(1,46) = 0{\cdot}92$, $P > 0{\cdot}1$). There were no further interactions (group × category, $F(5,230) = 1{\cdot}65$, $P > 0{\cdot}1$; first cluster *versus* subsequent visits to clusters × category, $F(5,230) = 0{\cdot}95$, $P > 0{\cdot}1$; group × first cluster *versus* subsequent visits to clusters × category, $F(5,230) = 1{\cdot}30$, $P > 0{\cdot}1$).

Is there a difference between patients and controls in terms of within-cluster times and between-cluster times?

An ANOVA calculated with group (controls, patients) as the between-group factor and timing measurement (within-subcluster, between-subcluster) and category (sports, animals, furniture, vehicles, fruit, vegetables) as the within-subject factors. There were significant main effects of group ($F(1,46) = 19{\cdot}85$, $P < 0{\cdot}0001$), timing measurement ($F(1,46) = 16{\cdot}13$, $P < 0{\cdot}001$), and

https://doi.org/10.1017/S0033291702005597 Published online by Cambridge University Press

category ($F(5,230) = 4{\cdot}06$, $P < 0{\cdot}001$). Although patients spent more time than controls within-clusters and between-clusters, the two-way interaction between group × timing measurement did not reach significance ($F(1,46) = 0{\cdot}02$, $P > 0{\cdot}5$). Neither was there a significant interaction between group × category ($F(5,230) = 0{\cdot}41$, $P > 0{\cdot}5$), although there was between timing measurement and category ($F(5,230) = 2{\cdot}53$, $P < 0{\cdot}05$). Finally there was not a significant three-way interaction between group × timing measurement × category ($F(5,230) = 0{\cdot}13$, $P > 0{\cdot}5$). This suggests that although patients are both slower within-clusters and between-clusters they are equivalently so. In other words, patients with schizophrenia are not disproportionately slower going from word to word within a cluster than they are switching to a word in another cluster.

DISCUSSION

Consistent with previous literature, we found that patients with schizophrenia produced fewer words and made more errors than healthy controls on a semantic verbal fluency task (Allen *et al.* 1993; Gourovitch *et al.* 1996; Joyce *et al.* 1996; Granholm *et al.* 1998; Elvevåg *et al.* 2001). As we have shown previously, there are large differences between word production depending upon the category selected (e.g. 'animal' *versus* 'furniture'), and thus it is possible to observe category-specific fluency deficits in patients with schizophrenia (e.g. Elvevåg *et al.* 2001). Previous studies have speculated that the mechanism behind this poverty of fluency may be due to problems in the organization and thus retrieval of this information (e.g. Allen *et al.* 1993; Robert *et al.* 1998). Other studies have come to similar conclusions. For example, Iddon *et al.* (1998) found that on a task that required self-generation of effective mnemonic strategies on a verbal task that schizophrenic patients were impaired at semantically classifying words as compared to controls, and thus did not benefit from effective strategies. In a similar vein, the study by Robert *et al.* (1998) reported that schizophrenic patients had problems in switching and clustering during the semantic fluency task. In the current study, we employed a novel method to assess and define clustering and did not find support for clustering problems *per se* in schizophrenia, or that the lexicon structure (i.e. cluster size) is smaller in schizophrenia (cf. Allen *et al.* 1993; Elvevåg *et al.* 2001). Rather, we found that patients' general slowness contributes to the difficulties that they have in finding and thus producing new words within their semantic field.

As stated in the Introduction, since schizophrenic patients generally have been shown to produce fewer fluency items (as in this study also), they logically must either have fewer categories, or fewer items per category, or both. However, although there was a trend for patients to produce fewer clusters and fewer items per cluster, this did not reach significance. However, there was a significant group difference in the number of visits to a cluster, with the patient group showing significantly fewer subsequent visits than controls. Importantly, covarying for the total number of words generated eliminated this group difference, a finding that demonstrates our argument (in the Introduction) that with fewer words being generated, there is by default less scope for clustering and switching, and thus the findings by Robert *et al.* (1998) may be confounded by this factor. We argue that these findings are important for two reasons. First, our index of clustering is clearly not an arbitrary measure, since it is comparable across participants and across groups. Moreover, the potential criticism that it could be affected disproportionately by patients' lack of cooperation, or motivation, is clearly not warranted. Secondly, it is interesting to note that this method whereby a participant creates their own subcategories does not yield a difference between the groups in terms of their category number or size. Thus, it is very probable that Bousfield (1953) was correct in advising against imposing judgements upon individuals' clustering. This may especially be the case for patients who have been observed to produce many 'unusual' associations in word association tasks (e.g. Kent & Rosanoff, 1910). Employing the clustering methods of Troyer *et al.* (1997, 1998) or Robert *et al.* (1998) may simply reify our notions that patients manifest problems in semantic organization, rather than attempts to get participants to individually group their own clusters. Moreover, failing to take into account the significant group difference in the number of words produced may confound the entire issue.

https://doi.org/10.1017/S0033291702005597 Published online by Cambridge University Press

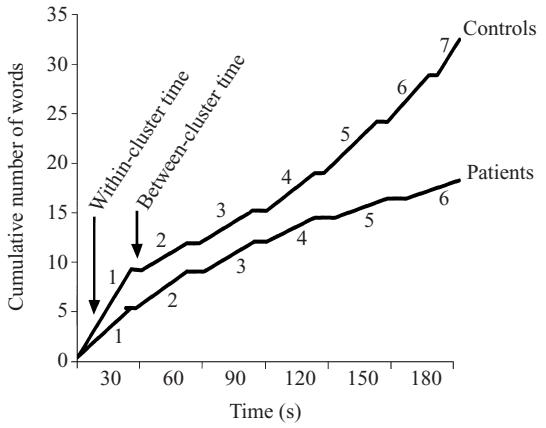


FIG. 1. The mean cumulative number of words and number of clusters produced by control participants and schizophrenic patients over 3 min of fluency tasks. The steep part of the plotted lines reflects the within-cluster production time and the horizontal part the time between-clusters.

Potential criticisms may be directed at our method whereby participants' sort their own clusters. Indeed, it is possible that participants' sorting does not directly mirror their organization at output. However, we contend that it is more likely that participants' own sorting reflects their own associations and clusters better than if the experimenter is adopting his or her own comparatively arbitrary delineation, as characterized by Robert *et al.* (1998). If one is concerned about reliability, it would be possible in future studies to repeat the same card-sort tasks; the participants returned for testing only once in the current study. However, since participants produced a comparable number of clusters and members per cluster, our method is not noticeably lacking in reliability. Thus, our results suggest that despite having access to the same number of ideas (clusters) in their semantic network as control participants, schizophrenic patients do have difficulties generating new words from within their semantic field.

It is our view that this detailed analysis of verbal fluency output will serve to constrain theories about the precise nature of the problems that patients with schizophrenia reliably demonstrate on simple fluency tasks. It is somewhat surprising to find that schizophrenic patients' performance is not qualitatively different. With reference to Fig. 1, we suggest that patients' production of fewer clusters and a slightly smaller cluster size (though not significantly so), contributes to their lower overall word production rate. Indeed, it is striking how similar the groups are on the measures we examined.

The fact that patients produced only slightly fewer clusters than controls, may indicate that patients categorize items in roughly the same manner and have access to as many semantic categories as controls. As indicated by the term within-cluster and between-cluster times, patients' overall diminished verbal output appears to be due to a generalized slowing of the semantic search process which in turn, results in fewer clusters. In other words, patients 'ran out of time'. We believe that if the patients were given more time on the task, their results might even more closely resemble the performance of controls. This quantitative explanation may at first seem somewhat counter-intuitive given the theoretical motivation of our study. The key finding from verbal fluency tasks in schizophrenia is a lower production of the number of words in a given period of time than healthy individuals. A possible reason for this slowness may be because of a slightly sub-optimal semantic organizational strategy. Speculatively, it is possible that this sub-optimal network hinders the semantic activation spreading in the normal fashion. Indeed, the finding of more repeat visits to clusters by control participants may be indicative of an organization strategy that is beneficial for memory retrieval, and one that is not functioning optimally in patients with schizophrenia. This organization may facilitate control participants in their search of another word both within the same cluster and to a different cluster, and in turn speed up their overall word production. Speculatively, these results of patients could be explained by hypo-priming in the sense of semantic activation not spreading in the normal fashion; indeed there is evidence for hypo-priming in schizophrenia (Ober *et al.* 1995; Aloia *et al.* 1996).

In conclusion, the results of this study suggest that patients with schizophrenia access a similar number of ideas (i.e. words), but that they are slower and less able to generate as many ideas in the same amount of time as healthy participants, probably because of a sub-optimal organizational strategy in their semantic networks.

We are grateful for the useful comments of Dr Daniel Weinberger.

https://doi.org/10.1017/S0033291702005597 Published online by Cambridge University Press

REFERENCES

Allen, H. A., Liddle, P. F. & Frith, C. D. (1993). Negative features, retrieval processes and verbal fluency in schizophrenia. *British Journal of Psychiatry* **163**, 769–775.

Aloia, M. S., Gourovitch, M. L., Weinberger, D. R. & Goldberg, T. E. (1996). An investigation of semantic space in patients with schizophrenia. *Journal of the International Neuropsychological Society* **2**, 267–273.

Bousfield, W. A. (1953). The occurrence of clustering in the recall of randomly arranged associates. *Journal of General Psychology* **49**, 229–240.

Bousfield, W. A. & Sedgewick, H. W. (1944). An analysis of sequences of restricted associative responses. *Journal of General Psychology* **30**, 149–165.

Bousfield, W. A., Sedgewick, H. W. & Cohen, B. H. (1954). Certain temporal characteristics of the recall of verbal associates. *American Journal of Psychology* **67**, 111–118.

Cattell, R. B. (1971). *Abilities: Their Structure, Growth and Action*. Houghton-Mifflin: Boston.

Chen, E. Y. H., Wilkins, A. J. & McKenna, P. J. (1994). Semantic memory is both impaired and anomalous in schizophrenia. *Psychological Medicine* **24**, 193–202.

Crawford, J. R., Obonsawin, M. C. & Bremner, M. (1993). Frontal lobe impairment in schizophrenia: relationship to intellectual functioning. *Psychological Medicine* **23**, 787–790.

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.

Goldberg, T. E., Aloia, M. S., Gourovitch, M. L., Missar, D., Pickar, D. & Weinberger, D. R. (1998). Cognitive substrates of thought disorder, I: the semantic system. *American Journal of Psychiatry* **155**, 1671–1676.

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

Granholm, E., Chock, D. & Morris, S. (1998). Pupillary responses evoked during verbal fluency tasks indicate semantic dysfunction in schizophrenia. *Journal of Clinical and Experimental Neuropsychology* **20**, 856–872.

Gruenewald, P. J. & Lockhead, G. R. (1980). The free recall of category examples. *Journal of Experimental Psychology: Human Learning and Memory* **6**, 225–240.

Gurd, J. M. (1993). Studies of Verbal Fluency Deficits in Patients with Parkinson's Disease. D.Phil. thesis, University of Oxford.

Gurd, J. M., Elvevåg, B. & Cortina-Borja, M. (1997). Semantic category word search impairment in schizophrenia. *Cognitive Neuropsychiatry* **2**, 291–302.

Iddon, J. L., McKenna, P. J., Sahakian, B. J. & Robbins, T. W. (1998). Impaired generation and use of strategy in schizophrenia: evidence from visuospatial and verbal tasks. *Psychological Medicine* **28**, 1049–1062.

Institute for Personality and Ability Testing (1973). *Measuring Intelligence with the Culture Fair Test*. The Institute for Personality and Ability Testing: Champaign, Ill.

Jastak, S. & Wilkinson, G. S. (1984). *The Wide Range Achievement Test – Revised Administration Manual* (*Revised Edition*). Jastak Assoc., Inc: Wilmington, DE.

Joyce, E. M., Collinson, S. L. & Crichton, P. (1996). Verbal fluency in schizophrenia: relationship with executive function, semantic memory and clinical alogia. *Psychological Medicine* **26**, 39–49.

Kent, G. H. & Rosanoff, A. J. (1910). A study of association in insanity. *American Journal of Psychiatry* **67**, 317–390.

McKenna, P. J., Mortimer, A. M. & Hodges, J. R. (1994). Semantic memory and schizophrenia. In *The Neuropsychology of Schizophrenia* (ed. A. S. David and J. C. Cutting), pp. 163–178. Lawrence Erlbaum Associates: Hillsdale, NJ.

Ober, A. B., Vinogradov, S. & Shenaut, G. K. (1995). Semantic priming of category relations in schizophrenia. *Neuropsychology* **9**, 220–228.

Robert, P. H., Lafont, V., Medecin, I., Berthet, L., Thauby, S., Baudu, C. & Darcourt, G. (1998). Clustering and switching strategies in verbal fluency tasks: comparison between schizophrenics and healthy adults. *Journal of the International Neuropsychological Society* **4**, 539–546.

Tamyln, D., McKenna, P. J., Mortimer, A. M., Lund, C. E., Hammond, S. & Baddeley, A. D. (1992). Memory impairment in schizophrenia: its extent, affiliations and neuropsychological character. *Psychological Medicine* **22**, 101–115.

Troyer, A. K., Moscovitch, M. & Winocur, G. (1997). Clustering and switching as two components of verbal fluency: evidence from younger and older adults. *Neuropsychology* **11**, 138–146.

Troyer, A. K., Moscovitch, M., Winocur, G., Alexander, M. P. & Stuss, D. (1998). Clustering and switching on verbal fluency: the effects of focal frontal- and temporal-lobe lesions. *Neuropsychologia* **36**, 499–504.

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

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

https://doi.org/10.1017/S0033291702005597 Published online by Cambridge University Press