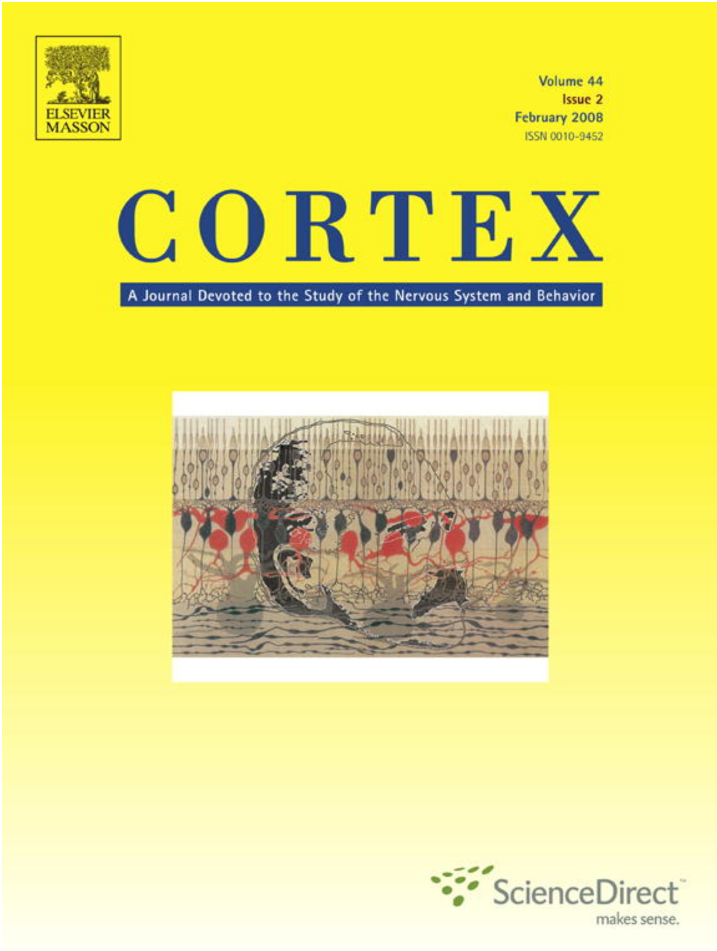
ELSEVIER
MASSON
Volume 44
Issue 2
February 2008
ISSN 0010-9452
CORTEX
A Journal Devoted to the Study of the Nervous System and Behavior
ScienceDirect
makes sense.

cortex 44 (2008) 161–169

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/cortex

Research report

Cut the coda: Early fluency intervals predict diagnoses

Sven-Erik Fernaeus[a,*], Per Östberg[a], Åke Hellström[b] and Lars-Olof Wahlund[a]

[a]Section for Clinical Geriatrics, Department of Neurobiology, Caring Sciences and Society, Karolinska Institutet, Stockholm, Sweden
[b]Department of Psychology, Stockholm University, Sweden

ARTICLE INFO

Article history:
Received 1 March 2005
Reviewed 24 June 2005
Revised 30 August 2005
Accepted 18 April 2006
Action editor Robin Morris
Published online 17 November 2007

Keywords:
Alzheimer's disease
Category fluency
Dementia
Factor analysis
Letter fluency
Mild cognitive impairment
Semantic memory
Word fluency

ABSTRACT

The aim of this study was threefold: (i) to clarify whether letter and category fluency tap different cognitive abilities; (ii) to make diagnostic comparisons and predictions using temporally resolved fluency data; (iii) to challenge and test the widely made assumption that 1-min sum scores are the fluency test measure of choice in the diagnosis of dementia. Scores from six 10-sec intervals of letter and category fluency tests were obtained from 240 participants including cognitive levels ranging from mild subjective cognitive complaints to Alzheimer's disease. Factor analysis revealed clearly separate factors corresponding to letter and category fluency. Category fluency was markedly impaired in Alzheimer's disease but not in Mild Cognitive Impairment. Only scores from relatively early intervals predicted Alzheimer's disease and Mild Cognitive Impairment. The conclusions are (i) letter and category fluency are different tests, category fluency being the best diagnostic predictor; (ii) it would be possible to administer category fluency tests only for 30 sec, because after this point the necessary differential diagnostic information about the patient's word fluency capacity has already been gathered.

1. Introduction

There are two major types of word fluency tests, both with a standard time limit of 1 min. Letter or "phonemic" fluency requires participants to produce words that begin with a given letter. In category or semantic fluency the task is to enumerate nouns from a taxonomy, for example, animals, or from a semantic field like supermarket items. Although word fluency tests are widely used, there has been relatively little work on the internal structure of the tasks or on processes that lead up to the participant's response. One-minute sum scores are normally employed as measures of fluency performance. This procedure, however, makes it impossible to analyze the temporal distribution of word production. Hence differences of interest between diagnostic groups and different fluency tests may be hidden in temporally lumped fluency data.

On the other hand, the use of lumped 1-min scores implicitly assumes that word production from all time intervals are relevant to diagnosis. This has never been demonstrated.

Studies that go beyond lumped scores include the assessment of "clustering" and "switching" based on relations between the words produced (Troyer et al., 1997; cf. Mayr, 2002) and studies of the temporal distribution of words produced (Crowe, 1998; Fernaeus and Almkvist, 1998; Fernaeus

* Corresponding author. Section for Clinical Geriatrics, Department of Neurobiology, Caring Sciences and Society, Karolinska Institutet, Novum, Floor 5, Karolinska University Hospital, Huddinge SE-141 86 Stockholm, Sweden.
E-mail address: sven-erik.fernaeus@swipnet.se (S.-E. Fernaeus).
0010-9452/$ – see front matter
doi:10.1016/j.cortex.2006.04.002

et al., 2001). Fernaeus and Almkvist (1998) found, using temporally resolved fluency data, that word production in letter fluency follows a negatively accelerated curve, approaching an asymptotic level after about 30 sec. They also performed an exploratory factor analysis based on word production scores from six 10-sec intervals of letter fluency. This analysis resulted in two factors. The first factor loaded on the initial intervals of letter fluency and was associated with rapid semi-automatic word retrieval. The second factor loaded on the later intervals and was interpreted as reflecting non-automatic, effortful word retrieval. Clearly, then, these results cannot be reconciled with notions of a homogeneous structure behind lumped fluency scores. Another observation that casts doubt on the homogeneity of lumped scores is that frontal white matter lesions selectively impair the initiation of letter fluency (Fernaeus et al., 2001).

Similarity of retrieval modes across tasks appears to be a parsimonious assumption about word production in fluency tasks in general (Crowe, 1998). In reality, however, there exists considerable evidence in favor of different retrieval modes in category and letter fluency. This evidence includes task-dependent vulnerability to demands from extra tasks (Martin et al., 1994), task-dependent sizes of word clusters (Troyer et al., 1997), and task-dependent impact of antimuscarinic medication (scopolamine; Pompeia et al., 2002). Furthermore, letter fluency is disproportionately impaired in minor Broca's aphasia (Coslett et al., 1991), nonfluent progressive aphasia (Hodges and Patterson, 1996), and cerebellar lesions (Leggio et al., 2000). Conversely, category fluency is disproportionately impaired in normal aging (Tomer and Levin, 1993; Kozora and Cullum, 1995; Ravdin et al., 2003), semantic dementia (Hodges and Patterson, 1996), Alzheimer's disease (Monsch et al., 1992, 1994, 1997), Parkinson's disease (Henry and Crawford, 2004b), schizophrenia (Kremen et al., 2003), and depression (Fossati et al., 2003; cf. Ravdin et al., 2003). To this may be added task-dependent brain activation patterns (Mummery et al., 1996; Keilp et al., 1999) and shorter duration of dementia illness in patients with relatively worse letter fluency performance (Coen et al., 1996).

The dissociations mentioned above indicate that different brain lesions affect letter and category fluency differentially. Apparently the two types of tests involve different cognitive processes – partly, at least – and hence differential neural activation patterns. Letter fluency may be characterized as a form-oriented word association test in which visual orthographic word patterns are evoked initially. These word-initial orthographic patterns are then ''read'' aloud, occasionally producing some semantic activation that must be inhibited in order not to violate continued form-appropriate morphological or phonological encoding (Indefrey and Levelt, 2000). Category fluency, in contrast, may be characterized as a content-oriented word association test. Clinical tradition provides that category fluency is focused solely on the retrieval of nouns denoting physical objects, although retrieval of activity verbs has been investigated recently (Piatt et al., 1999a, 1999b, 2004; Östberg et al., 2005, 2007). The initiation process probably varies depending on the semantic category and on strategies adopted by the subject. For instance, coordinate word association (''cat–dog'') may be a low-level strategy used initially, and mentally touring a zoo or natural habitats may be a high-level strategy used later in the test (Indefrey and Levelt, 2000).

The specific neural correlates of letter and category fluency impairments have been much debated. Differential letter fluency deficits are associated with frontal lobe lesions (Henry and Crawford, 2004a). Disregarding focal neocortical temporal lobe lesions, a common denominator of disproportionate category fluency deficits appears to be pathological involvement of the limbic loop. Severe pathology of higher-order components of the limbic loop indeed characterizes not only Alzheimer's disease but also Parkinson's disease and other neurodegenerative processes (Braak et al., 2000). These components, including the perirhinal, entorhinal, and parahippocampal cortices, receive massive input from distributed neural networks in the high association cortices and constitute a major gateway for neocortical input to the hippocampal formation (Braak et al., 2000). Semantic neuronal networks are held to be widely distributed over higher sensory and motor association cortices in both cerebral hemispheres (Pulvermüller, 2003). Lesions of the higher-order limbic loop components are therefore likely to impede the extraction of neocortical information vital for flexible semantic processing on which category fluency depends.

As far as we know, letter versus category fluency has not been comprehensively studied before in a clinical group associated with prominent limbic loop pathology apart from Alzheimer's disease. The clinical diagnosis of Alzheimer's disease, however, is usually not made until pathoanatomical changes have definitely spread into the neocortex (Braak and Braak, 1991). At this stage dementia becomes apparent, and neocortical semantic networks themselves are degraded or devastated. Earlier stages of cognitive decline must therefore be investigated to find out whether limbic-stage pathology in itself affects fluency performance. In looking for such a stage of cognitive decline, we may note that Alzheimer-type changes systematically appear first in the transentorhinal subregion of the perirhinal cortex (Braak and Braak, 1991). These highly selective changes predate the clinical diagnosis of Alzheimer's disease with years or even decades, that is, Alzheimer's disease has a long preclinical phase characterized by a progressive and systematic spread of pathoanatomical changes. When objective signs of cognitive impairment appear, they are often not severe enough to warrant a diagnosis of dementia. This stage of objective cognitive decline is often referred to as mild cognitive impairment (e.g., Davis and Rockwood, 2004). It is probably characterized by pathology largely confined to the limbic loop (limbic stages III and IV of Braak and Braak, 1991; Mesulam, 2000). A corollary of this reasoning is that diagnostic prediction of this stage of cognitive decline should be based on category rather than letter fluency measures.

2. Aims

The first and most basic objective of the study was to establish, on the basis of temporally resolved data from a memory clinic clientele, whether letter and category fluency tap different cognitive abilities, as may be postulated on psycholinguistic grounds (Indefrey and Levelt, 2000). This aim was also motivated on the one hand by previous claims of similar retrieval modes in letter and category fluency (Crowe, 1998), and, on the other hand, by the recent finding of separate letter, noun (category) fluency, and verb fluency factors (Östberg et al., 2005). A second objective was to make specific diagnostic comparisons using temporally resolved fluency data. Thirdly, we sought to challenge and test the widespread practice of using 1-min sum scores by analyzing the diagnostic power of temporally resolved letter and category fluency data. More specifically, we attempted to find out which part(s) of 1-min performance that are the best predictors of cognitive decline. Participants were drawn from a memory clinic.

3. Method

3.1. Participants

Two hundred and forty consecutive patients (119 women and 121 men) aged between 32 and 87 years (mean = 65.7, standard deviation – SD = 10.3) participated in the study. Testing was conducted as part of an examination for cognitive problems at the Department of Geriatrics, Karolinska University Hospital, Huddinge. None of these patients participated in the studies of Fernaeus and Almkvist (1998) or Fernaeus et al. (2001). The participants had completed at least one letter fluency test and one category fluency test. Group comparisons were made for the three most common diagnostic categories, which represent postulated stages of preclinical and clinical dementia: Alzheimer's disease (AD): $N = 58$; Mild Cognitive Impairment – MCI: $N = 82$; Subjective Cognitive Impairment (without objective impairment) (SCI): $N = 45$, see Table 1. AD was diagnosed according to the International Classification of Diseases (ICD-10). MCI patients had subjective cognitive complaints, performed at least 1.5 SD unit below average on one or more neuropsychological test (usually memory), but did not fulfill clinical diagnostic criteria for dementia according to DSM-IV (Wahlund et al., 2003). SCI was diagnosed in cases that were referred because of subjective cognitive complaints, but in which no objective signs of cognitive impairment or cerebral pathology were found despite extensive examination. Other diagnoses included unspecified dementia ($N = 18$), vascular dementia ($N = 16$), frontotemporal dementia ($N = 11$), semantic dementia ($N = 7$), progressive nonfluent aphasia ($N = 1$), alcohol dementia ($N = 1$), and focal brain lesion ($N = 1$). Neither temporal patterns in letter fluency or category fluency nor parameters derived from these patterns were used as markers for specific diagnoses. The Mini-Mental State Examination (MMSE) was used to assess the level of global cognitive impairment (Folstein et al., 1975). MMSE scores ranged between 7 and 30. All SCI subjects had MMSE > 27.

There were significantly more women than men in the AD group and more men in the SCI group, $\chi^2_2 = 7.94$, $p < .05$. There was no significant main effect of sex on MMSE score, $F_{1,176} < 1$, and no interaction between sex and diagnostic group, $F_{1,2} < 1$. There was a significant main effect of sex on educational level (years of study), $F_{2,179} = 4.67$, $p < .05$, but a post hoc comparison between sexes did not confirm this (Scheffé, $p = .14$). There was no significant sex × education interaction, $F_{1,2} < 1$.

3.2. Procedure

The participants were tested by the second author (P.Ö.). In letter fluency, the initial letters used were F, A, and S. [The proportions of words beginning with these letters in Swedish newspaper text (Allén, 1972) are about 8% for F, 4% for A, and 10% for S]. Proper nouns and numerals were not allowed. In category fluency, the categories used were animals, fruits, and clothes. All participants took all the letter fluency tests as well as the Animals test. Some participants also took either the Fruits or the Clothes test. Category fluency was performed first, thereafter letter fluency with an unrelated intervening task (reading single words and pseudowords aloud) in order to minimize stuck-in-set perseveration. In both letter fluency and category fluency the time allowed for word generation was 60 sec; words were transcribed and counted for each of six 10-sec intervals. If a word was started in one interval and finished in the next, it was scored within the first interval. Repetitions were omitted from the score. Mean scores for each interval in letter and category fluency were used for analysis.

Table 1 – Mean age, education, and Mini-Mental State Examination (MMSE) scores in the three major diagnostic groups

Sex	Diagnosis	Age		Education			MMSE		MMSE adjusted for age and education
		Mean	SD		Mean	SD	Mean	SD	
Female	AD, N = 22	66.6	10.2				22.59	4.53	
Male	AD, N = 36	69.5	8.6				21.50	4.22	
	AD total, N = 58	68.3	9.4	9.7	10.0	3.3	22.05	4.34	22.27
Female	MCI, N = 46	65.9	9.8	N.A.	Age adj.		27.17	2.41	
Male	MCI, N = 36	65.4	9.6				27.20	2.96	
	MCI total, N = 82	65.8	9.7	10.9	11.0	3.9	27.19	2.64	27.25
Female	SCI, N = 29	60.8	9.0	N.A.	Age adj.		29.10	1.08	
Male	SCI, N = 16	54.8	13.0				29.06	1.12	
	SCI total, N = 45	58.6	10.8	12.6	12.3	3.3	29.09	1.1	28.69
Total mean and SD in age and MMSE		64.8	10.5		10.9	3.7	26.04	4.15	

N.A. = not age-adjusted.

3.3. Statistical analysis

Two hundred and forty cases were included to satisfy a stringent rule of thumb for the number of cases in factor analysis, namely, the ''Rule of 200'' (Gorsuch, 1983). This rule requires at least 200 cases regardless of subject-to-variable ratio. Because of the unequal number of participants in the categories with dementia (e.g., more than three times as many with AD as with vascular dementia), only the AD category was selected for statistical comparisons with SCI and MCI. Analyses were based on mean scores, which by definition are ratio variables which permit the use of discriminant analysis and factor analysis. Factor analysis was used after checking for Kaiser–Meyer–Olkin's measure of adequacy (KMO; Kaiser, 1974) in order to analyze the multivariate structure of the fluency tasks.

In the heterogeneous kind of sample we studied it is necessary to assume that the factors may be correlated. Therefore, rotational techniques like Direct Oblimin were used in addition to orthogonal rotations like Varimax. We used Maximum likelihood factor analysis to provide statistical guidance about the precise number of factors (see Östberg et al., 2005). Thus besides inspection of the scree plot of eigenvalues, we selected the solution with the highest number of factors that yielded a significantly lower χ^2 value than the solution with one factor less.

Predictions of diagnostic group membership from mean word fluency scores at different intervals and demographic variables (age, educational level) were made using logistic regression analyses. Discriminant analyses were performed to corroborate the results of the logistic regression analyses.

Exponential and piecewise linear breakpoint analyses were conducted in order to find out whether the main diagnostic groups (SCI, MCI, AD) were characterized by different performance functions using fluency interval as the main independent variable. If a breakpoint function satisfies the performance pattern, then it is reasonable to assume that at least two temporally different processes are involved in the performance. Function parameters resulting from these analyses would therefore describe concisely the production curves typical for each diagnostic category.

Statistical analyses were performed using Statistica for Macintosh, Statistica 6.0 for Windows (Statsoft, Tulsa, TX, USA) and SPSS for Macintosh (Statistical Package for the Social Sciences, SPSS Inc., Chicago, IL, USA).

4. Results

The results are presented in the following order: first, factor analyses based on the entire sample including all diagnostic groups; secondly, comparisons between specific diagnostic groups (SCI, MCI, and AD) and tests (letter vs. category fluency); thirdly, predictions of group membership based on test performance and background variables (age and education). Because all participants did not perform each and every fluency subtest, there were minor differences in the number of cases forming the basis for the statistical analyses.

4.1. Factor analysis of interval scores in letter and category fluency

The temporally resolved fluency data consisted of six 10-sec interval variables from each fluency task (letters F, A, and S in letter fluency; animals, clothes, and fruits in category fluency). Average category (animals/clothes/fruits) interval scores and average letter (F/A/S) interval scores were used as input for factor analyses.

After checking for the KMO's measure of sampling adequacy (which was .94), two factor analyses of the scores from all 240 participants were conducted. In order to test a model of orthogonal factors a Varimax rotation was made, which showed a pattern of factor loadings roughly reflecting the two types of test, see Table 2. The two factors explained 36.4% and 30.5%, respectively.

As seen in Table 2, all loadings are positive and are thus to be found in the upper right quadrant formed by the orthogonal axes representing the two factors. However, this always happens if the factors explaining the variance are actually correlated.

In order to study further the relationship between factors a second factor analysis using the principal axis factors method was made. Two factors with eigenvalues above 1 emerged, where the first factor had an eigenvalue of 6.79 and was responsible for 56.6% of the total variance. The second factor had an eigenvalue of 1.24, which explained 10.4% of the total variance. An oblique solution produced a pattern of loadings shown in Table 3.

The Maximum Likelihood method was applied in order to determine more precisely the number of factors involved. Two factors with eigenvalues above 1 emerged again with a cumulative variance at 60.4% based on initial eigenvalues of 6.79 and 1.24, and with a cumulative variance at 66.9%. This last solution did not depart from significance, $\chi^2_{43} = 65.16$, $p = .02$, but was significantly better than a one-factor solution according to the χ^2 comparisons, $\chi^2_{11} = 171.19$, $p < .001$. Although a three-factor solution departed from

Table 2 – Factors and loadings based on principal factors extraction and Varimax rotation

Variables	Factors	
	I	II
FAS interval 1	**.735**	.392
FAS interval 2	**.834**	.294
FAS interval 3	**.825**	.341
FAS interval 4	**.807**	.297
FAS interval 5	**.805**	.280
FAS interval 6	**.789**	.299
CF interval 1	.377	.657
CF interval 2	.346	.687
CF interval 3	.272	**.732**
CF interval 4	.205	**.786**
CF interval 5	.258	**.740**
CF interval 6	.297	.666
Variance	.364	.305

Note: interval scores are sorted and loadings > .700 are boldface.

Table 3 – Factor loadings based on principal axis factors and Oblimin rotation with Kaiser normalization

Variables	Factors	
	I	II
FAS interval 2	**.894**	
FAS interval 3	**.866**	
FAS interval 4	**.830**	
FAS interval 5	**.824**	
FAS interval 6	**.790**	
FAS interval 1	.676	.162
CF interval 4	−.118	**.838**
CF interval 5		**.732**
CF interval 3		**.724**
CF interval 2		.647
CF interval 6		.604
CF interval 1		.603

Note that factor loadings are sorted, and that loadings > .700 are boldface. Loadings < .05 are not shown.

significance, it did not meet the criteria for inclusion in the final solution, because the last factor was based only on two items and explained only 2.1% of the total variance. The factor analysis based on the Maximum Likelihood extraction method, which resulted in two correlated factors, $r = .74$, was therefore chosen as the one that best fitted the data. Because the pattern of loadings of these factors almost completely coincided with intervals of the two different test types, they were labeled letter fluency and category fluency. In letter fluency all loadings except for the first interval were above .70. In category fluency the highest loadings were found for intervals 3–5, see Table 4.

One final aim was to determine the effect of age and education on the interval scores of category and letter fluency, so a principal axis factor analysis with Oblimin rotation that included these variables was conducted. Three factors emerged with eigenvalues 6.99, 1.38, and 1.03, explaining 49.2%, 9.9%, and 7.4%, respectively, of the total variance. The two largest factors were similar to the ones found earlier, namely, letter fluency and category fluency, whereas age and education loaded on the third factor. Because effects of age and education are presented later in the article and because only these two variables had substantial loadings on the third factor, the table of loadings from this last factor analysis is not presented.

Table 4 – Factor loadings based on Maximum Likelihood extraction and Oblimin rotation with Kaiser normalization

Variables	Factors	
	Letter fluency	Category fluency
FAS interval 2	**.899**	−.047
FAS interval 3	**.870**	.015
FAS interval 4	**.830**	.007
FAS interval 5	**.815**	.007
FAS interval 6	**.790**	.021
FAS interval 1	.677	.160
CF interval 4	−.121	**.843**
CF interval 3	.002	**.725**
CF interval 5	.001	**.724**
CF interval 2	.106	.641
CF interval 6	.136	.611
CF interval 1	.074	.608

Note that factor loadings are sorted and loadings > .700 are boldface.

4.2. *Diagnostic differences in category and letter fluency*

In order to study the difference between the three major diagnostic groups (AD, MCI, and SCI) in category versus letter fluency, MANCOVAs were conducted with diagnostic group as between-groups variable, test type as the first within-groups factor, and interval as the second within-groups factor. Age and education were included as covariates.

There were significant main effects of (i) diagnostic group, $F_{2,180} = 58.19$, $p < .001$, indicating significant differences between all three groups in the expected direction (Scheffé, $p < .001$); (ii) test type, $F_{1,182} = 160.08$, $p < .001$, indicating higher scores in category fluency than in letter fluency; and (iii) interval, $F_{5,910} = 443.61$, $p < .001$, indicating higher scores initially. There was also a significant interaction between diagnostic group and test type, $F_{2,182} = 14.93$, $p < .01$, indicating non-significant differences between category fluency and letter fluency scores in AD but significant effects of test type in SCI and MCI, with superior category fluency performance (Scheffé, $p < .001$). There was also a significant interaction between diagnostic group and interval, $F_{10,910} = 7.00$, $p < .001$, indicating significant differences between groups in all intervals (Scheffé, $p < .01$). Finally, there was a significant test type × interval interaction, indicating higher initial scores in category fluency than in letter fluency, Rao $R_{5,178} = 36.67$, $p < .001$, see Fig. 1.

There were no significant within-cells effects of the covariates age and education, Pillai–Bartlett$_{24,340} = 1.36$, n.s. There were, however, significant effects of education in the first three intervals (Scheffé, $p < .05$) and the last three intervals (Scheffé, $p < .001$) of letter fluency.

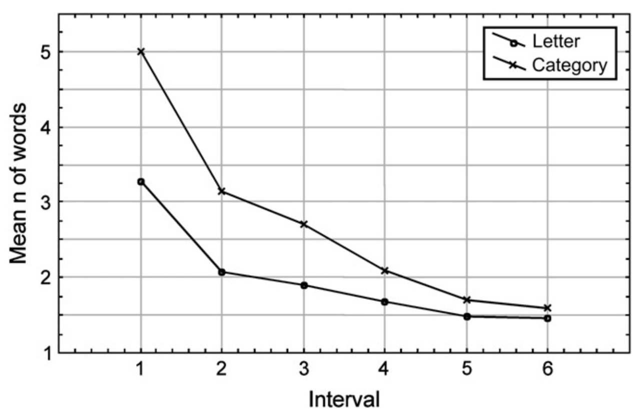

Fig. 1 – Interaction between type of test (letter fluency vs. category fluency) and interval (1–6).

4.2.1. *Prediction of diagnostic group membership*

Logistic regression and discriminant analyses were conducted with mean combined interval scores from category fluency (animals/clothes/fruits) and mean combined interval scores from letter fluency (F/A/S), education and age as predictor variables, and clinical diagnosis as the dependent variable. Diagnostic groups (AD, MCI, SCI) were selected as dependent variables at the two hypothesized transgressions, that is, SCI versus MCI and MCI versus AD. These analyses were performed in order to elucidate the interactions found in the MANCOVAs and to evaluate the potential of category and letter fluency to predict objective cognitive impairment (SCI vs. MCI) and dementia (MCI vs. AD). Logistic regression analyses were followed by discriminant analyses to determine the number of cases classified according to expectation.

4.2.2. *Prediction of MCI versus AD*

A logistic regression analysis was performed to evaluate the prediction of MCI versus AD from mean fluency interval scores. There was a significant prediction of diagnostic group, Nagelkerke $R^2 = .59$, $p < .001$. Three word fluency parameters were included in the model. These were category fluency intervals 1–3. Category fluency interval 2 was the first variable included in the stepwise model, $\chi^2 = 49.82$, $p < .001$. Category fluency interval 3 improved the prediction, $\chi^2 = 22.60$, $p < .001$. Category fluency interval 1 added further improvement to the model, $\chi^2 = 7.52$, $p < .01$. Neither letter fluency intervals nor age or education contributed significantly to the prediction.

The discriminant analysis of AD versus MCI resulted in an 86.6% classification of MCI patients as such, and 74.1% correct classification of AD patients, with a total correct classification of 81.4%; this classification was based on the same predictors as the logistic regression model.

4.2.3. *Prediction of SCI versus MCI*

In order to further investigate the differential diagnostic validity of letter and category fluency in identifying mild but objective cognitive impairment, the prediction of MCI versus SCI from the same variables as above was analyzed using logistic regression analysis. There was a moderate significant prediction of MCI versus SCI, Nagelkerke $R^2 = .29$, $p < .001$. Two word fluency parameters were included in the model. These were category fluency interval 3 and letter fluency interval 3. Category fluency interval 3 was the first variable included in the stepwise model, $\chi^2 = 15.02$, $p < .001$. Age improved the prediction, $\chi^2 = 10.04$, $p < .01$. Letter fluency interval 3 added further improvement to the model, $\chi^2 = 4.55$, $p < .05$.

The discriminant analysis of SCI versus MCI resulted in a 70.7% classification of MCI patients as such, and a 31.1% correct classification of SCI patients, with a total correct classification of 68.9%. The classification was based on the same predictors as the logistic regression model.

4.2.4. *Word fluency function estimates in category and letter fluency*

Exponential and piecewise linear breakpoint functions were applied to the data in order to define functions for the three major diagnostic groups (AD, MCI, SCI) in category fluency. The exponential function was the standard, $y = c + \exp(b0 + b1 \times x1 + b2 \times x2\ldots)$, using the Statistica package. The fitted breakpoint model: no. of words $= B0 - B1 \times$ interval (if interval $<$ breakpoint); no. of words $= B0 - B1 \times$ breakpoint $- B2$ (interval $-$ breakpoint) (if interval $>$ breakpoint).

In SCI, piecewise linear regression with breakpoint had a fit of $R = .87$ (for intercepts and estimates for slopes and a breakpoint at 3.59 words, see Table 5). An exponential function had a fit of $R = .73$ and a linear regression $r = -.70$.

In MCI the fit for a breakpoint function was $R = .86$ with a breakpoint at 2.95 words, whereas the fit for an exponential function was $R = .72$ and a linear regression $r = -.67$. In AD there was also a fit of $R = .86$ for a piecewise linear function with a breakpoint at 1.61 words and a fit of $R = .64$ for an exponential function, while a linear regression was $r = -.56$.

For letter fluency the piecewise linear function in SCI had a fit of $R = .82$ with a breakpoint at 3.3, while an exponential function had a fit of $R = .61$, and a linear regression $r = -.54$. In MCI the piecewise linear functions with a breakpoint at 2.12 words had a fit of $R = .84$, whereas an exponential function had a fit of $R = .53$, and a linear regression $r = -.48$. In AD the fit for piecewise linear functions was $R = .81$ with a breakpoint at 1.31 words, and an exponential function had a fit of $R = .46$, while a linear regression was $r = -.37$.

The function analyses revealed that based on group data linear piecewise linear functions had fits about $R = .86$ for category fluency and between .81 and .84 for letter fluency (see Table 5).

The exponential functions were less fit to the data in all three groups with R ranging between .64 and .78 in category fluency and between .46 and .61 for letter fluency. The linear functions ranged between $-.56$ and $-.70$ for category fluency and between $-.37$ and $-.54$ for letter fluency.

As can be seen in Table 5, a comparison between SCI and other groups in the ratio $B1/B3$, that is, the relationship between the initial and later slopes, show that the objectively impaired groups (i.e., AD and MCI) had category fluency ratios ranging between 4 and 5, whereas SCI had a ratio of 2. The later intercepts in category fluency were 1.13 for AD, 2.12 for MCI, and 3.41 for SCI. In letter fluency the later intercepts were .56 for AD, 1.58 for MCI, and 2.58 for SCI.

Table 5 – Breakpoint parameters for the two fluency tests in the diagnostic groups AD, MCI and SCI

Test type/R	B0	B1	B2	B3	Breakpoint	B1/B3
(a) AD, $N = 58$						
LF: $R = .81$	2.74	−.02	.56	−.001	1.31	20
CF: $R = .86$	4.17	−.04	1.13	−.01	1.61	4
(b) MCI, $N = 82$						
LF: $R = .84$	3.87	−.02	1.58	−.01	2.12	2
CF: $R = .86$	5.71	−.05	2.12	−.01	2.95	5
(c) SCI, $N = 45$						
LF: $R = .83$	4.54	−.02	2.58	−.01	3.30	2
CF: $R = .87$	6.28	−.04	3.41	−.02	3.59	2

B0 = initial intercept, B1 = initial slope, B2 = later intercept, B3 = later slope, CF = category fluency, LF = letter fluency.

5. Discussion

The present study used temporally resolved category and letter fluency scores in order to test and challenge the use of 1-min lumped scores from category and letter fluency. Factor analysis was first applied to the temporally resolved data from each type of test. Two factors were found corresponding to letter and category fluency. There was an intercorrelation between the factors, which may have been elevated by the very large variation in general cognitive level in our sample. Nevertheless, the two factors obtained provide evidence that letter and category fluency reflect different psycholinguistic abilities. This finding is consistent with task-specific lead-in processes as well as task-differential psycholinguistic components during the tasks. Morphophonological encoding appears to be the core component in letter fluency whereas category fluency puts major demands on conceptual preparation (Indefrey and Levelt, 2000). Consequently, there is no strong basis for postulating similar retrieval modes in letter and category fluency (cf. Crowe, 1998).

Comparisons between average interval scores in the three main diagnostic groups (AD, MCI, SCI) showed a significant interaction between fluency interval and diagnostic group. This implies that these groups did not just differ in their levels of performance, but that fluency intervals may also be diagnostically relevant. However, because these findings were obtained on the group level they are not immediately applicable clinically. To corroborate the group comparisons we performed logistic regression and discriminant analyses. These analyses showed that specific fluency interval scores have significant predictive diagnostic power. This holds good in particular for intervals 1, 2, and 3 of category fluency in the diagnosis of AD and MCI. Moreover, function analyses showed distinct differences of breakpoint parameters between the three main diagnostic groups. The ratios between the slopes for early and late fluency performance differed strongly between individuals with only subjective complaints (SCI) and those patients who showed objective signs of cognitive decline (MCI) or dementia (AD). This further underscores that a simple one-dimensional analysis of lumped scores may obscure important underlying differences between these groups.

Logistic regression analyses showed that the first three intervals in category fluency add to each other in explaining the differences between MCI and AD. The third intervals of category and letter fluency added to each other in the prediction of objective cognitive decline, that is, in distinguishing MCI from SCI. Certain intervals during fluency performance may therefore reflect different capacities that are hidden by lumped scores.

How can these observations be explained in the context of diagnostics and psycholinguistics? Our results indicate that category fluency is a test that becomes increasingly difficult during performance. The targeted word set in letter fluency (for example, words that begin with <f>) may be more than 10 times larger than that in category fluency (for example, animals). Initial performance in category fluency, nonetheless, exceeds that of letter fluency but also decreases more rapidly. This difference is clearly reflected by the *B* values in our breakpoint function analyses.

Because intervals 1, 2, and 3 in category fluency predicted AD versus MCI, and interval 3 was predictive of MCI versus SCI, the most salient signs of dementia should be sought for in the first half of this test.

We cannot be very specific about the underlying nature of these predictions. However, it is possible that the first and second intervals in category fluency reflect a very basic process in conceptual preparation, for instance, full access to generic-level concepts in the most salient or familiar categories (like domestic animals). Such access would be spared in MCI but impaired in AD, where even firmly established corticocortical semantic networks have become atrophied. Because performance in the third interval of category fluency adds to the prediction of AD versus MCI and MCI versus SCI, it may be viewed as reflecting a separate ability. One hypothesis may be that it concerns the scope of conceptual search which is impaired in AD and in some MCI patients. This deficit is more likely to reflect disconnection of the limbic loop which impedes flexible semantic processing.

A more general neuropsychological interpretation would be that two basic levels of memory are relatively preserved in AD, namely, item memory and associative memory (Murdock, 1974). These memory functions demand less effort than serial memory. Serial memory becomes severely impaired in AD, resulting in a disability to search through non-prototypical categories in a systematic manner. In MCI this effortful systematic search may be impaired owing to specific limbic dysfunction. The term ''implicit memory'' is a near-synonym for ''associative memory'' in the current literature based on the memory systems paradigm (Tulving, 1987; Squire and Zola, 1998). The term ''implicit'' memory, however, seems less appropriate in this context because even in AD the memory search involved in category fluency would be better regarded as a conscious and effortful mental act.

Function estimates clearly showed that piecewise linear breakpoint functions fit the word fluency data better than exponential functions, especially so in letter fluency. Linear functions were even less fit to the data than exponential functions. In AD the breakpoint function explains twice as much of the variance as the linear function in category fluency and 4.8 times as much of the variance in letter fluency. It may be noted that the intercept, *B*2, for the later part of the prediction curve is more than three times greater for SCI than AD. Further, the ratio between the initial slope, *B*1, and the later slope, *B*3, is about twice that of MCI compared to SCI. The function estimates arrived at in this study may consequently be used as norms for the major groups at hand. More specifically, clinicians may be aided in their diagnostic decisions by the individual parameters used in the breakpoint functions, particularly the ratio between initial slope and later slope and the level of the breakpoint.

The last three intervals of category fluency have substantial loadings in the factor analyses. They may therefore contribute to the reliability of these types of test, but they add neither to the prediction of MCI nor to the prediction of AD. This implies that category fluency tests could be substantially shortened without reducing their diagnostic power. One advice to clinicians would then be to stop testing after 30 or perhaps 40 sec, but to keep a good record of each 10-sec interval.

Some limitations to the present study may be pointed out. Owing to the relatively small number of participants with other types of dementia than Alzheimer's, the study did not allow predictions concerning dementia in general or other specific forms of dementia, such as vascular dementia or frontotemporal dementia. We also did not address the actual lexical content in fluency performance. Further studies should explore both these issues.

6. Conclusion

Differences between category fluency and letter fluency were common in this study. Most important of these was the difference in predicting diagnosis along the continuum SCI, MCI, and AD. If a choice has to be made between them, category fluency should obviously be chosen. It would be advisable to pay attention to performance in category fluency tests during the first 30 sec, because after this point predictive information about the patient's word fluency capacity has already been gathered. Before such a procedure is implemented, however, a broader set for normative data would be required for comparative purposes.

Acknowledgements

This research was supported by the Swedish Medical Research Council, Alzheimerfonden, Karolinska Institutet, and Stockholm University. The tests were carried out at the Department of Geriatrics, Karolinska University Hospital, Huddinge, and supervised by Professor Bengt Winblad. We thank Associate Professor Graham Budd for linguistic correction.

REFERENCES

Allén S. *Tiotusen i topp* [Top ten thousand]. Stockholm: Almqvist and Wiksell, 1972.

Braak H and Braak E. Neuropathological stageing of Alzheimer-related changes. *Acta Neuropathologica (Berlin)*, 82: 239–259, 1991.

Braak H, Del Tredici K, Bohl J, Bratzke H, and Braak E. Pathological changes in the parahippocampal region in select non-Alzheimer's dementias. *Annals of the New York Academy of Sciences*, 911: 221–239, 2000.

Coen RF, Maguire C, Swanwick GR, Kirby M, Burke T, Lawlor BA, Walsh JB, and Coakley D. Letter and category fluency in Alzheimer's disease: a prognostic indicator of progression? *Dementia*, 7: 246–250, 1996.

Coslett HB, Bowers D, Verfaellie M, and Heilman KM. Frontal verbal amnesia – phonological amnesia. *Archives of Neurology*, 48: 949–955, 1991.

Crowe SF. Decrease in performance on the verbal fluency test as a function of time: evaluation in a young healthy sample. *Journal of Clinical and Experimental Neuropsychology*, 20: 391–401, 1998.

Davis HS and Rockwood K. Conceptualization of mild cognitive impairment: a review. *International Journal of Geriatric Psychiatry*, 19: 313–319, 2004.

Fernaeus SE and Almkvist O. Word production: dissociation of two retrieval modes of semantic memory across time. *Journal of Clinical and Experimental Neuropsychology*, 20: 137–143, 1998.

Fernaeus SE, Almkvist O, Bronge L, Östberg P, Hellström Å, Winblad B, and Wahlund LO. White matter lesions impair initiation of FAS flow. *Dementia and Geriatric Cognitive Disorders*, 12: 52–56, 2001.

Folstein MF, Folstein SE, and McHugh PR. "Mini-mental state." A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12: 189–198, 1975.

Fossati P, Guillaume Le B, Ergis AM, and Allilaire JF. Qualitative analysis of verbal fluency in depression. *Psychiatry Research*, 117: 17–24, 2003.

Gorsuch RL. *Factor Analysis*. Hillsdale, NJ: Lawrence Erlbaum, 1983.

Henry JD and Crawford JR. A meta-analytic review of verbal fluency performance following focal cortical lesions. *Neuropsychology*, 18: 284–295, 2004a.

Henry JD and Crawford JR. Verbal fluency deficits in Parkinson's disease: a meta-analysis. *Journal of the International Neuropsychological Society*, 10: 608–622, 2004b.

Hodges JR and Patterson K. Nonfluent progressive aphasia and semantic dementia: a comparative neuropsychological study. *Journal of the International Neuropsychological Society*, 2: 511–524, 1996.

Indefrey P and Levelt WJM. The neural correlates of language production. In Gazzaniga MS (Ed), *The New Cognitive Neurosciences*. Cambridge, MA: Cambridge University Press, 2000: 845–865 [chapter 59].

Kaiser HF. An index of factorial simplicity. *Psychometrika*, 39: 31–36, 1974.

Keilp JG, Gorlyn M, Alexander GE, Stern Y, and Prohovnik I. Cerebral blood flow patterns underlying the differential impairment in category vs letter fluency in Alzheimer's disease. *Neuropsychologia*, 37: 1251–1261, 1999.

Kozora E and Cullum CM. Generative naming in normal aging: total output and qualitative changes using phonemic and semantic constraints. *The Clinical Neuropsychologist*, 9: 313–325, 1995.

Kremen WS, Seidman LJ, Faraone SV, and Tsuang MT. Is there disproportionate impairment in semantic or phonemic fluency in schizophrenia? *Journal of the International Neuropsychological Society*, 9: 79–88, 2003.

Leggio MG, Silveri MC, Petrosini L, and Molinari M. Phonological grouping is specifically affected in cerebellar patients: a verbal fluency study. *Journal of Neurology, Neurosurgery and Psychiatry*, 69: 102–106, 2000.

Martin A, Wiggs CL, Lalonde F, and Mack C. Word retrieval to letter and semantic cues: a double dissociation in normal subjects using interference tasks. *Neuropsychologia*, 32: 1487–1494, 1994.

Mayr U. On the dissociation between clustering and switching in verbal fluency: comment on Troyer, Moscovitch, Winocur, Alexander and Stuss. *Neuropsychologia*, 40: 562–566, 2002.

Mesulam MM. Aging, Alzheimer's disease, and dementia. Clinical and neurobiological perspectives. In Mesulam MM (Ed), *Principles of Behavioral and Cognitive Neurology*. 2nd ed. Oxford: Oxford University Press, 2000: 439–522.

Monsch AU, Bondi MW, Butters N, Salmon DP, Katzman R, and Thal LJ. Comparisons of verbal fluency tasks in the detection of dementia of the Alzheimer type. *Archives of Neurology*, 49: 1253–1258, 1992.

Monsch AU, Bondi MW, Butters N, Paulsen JS, Salmon DP, Brugger P, and Swenson MR. A comparison of category and letter fluency in Alzheimer's disease and Huntington's disease. *Neuropsychology*, 8: 25–30, 1994.

Monsch AU, Seifritz E, Taylor KI, Ermini-Fünfschilling D, Stähelin R, and Spiegel R. Category fluency is also predominantly affected in Swiss Alzheimer's disease patients. *Acta Neurologica Scandinavica*, 95: 81–84, 1997.

Mummery CJ, Patterson K, Hodges JR, and Wise RJ. 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*, 22: 989–995, 1996.

Murdock BB. *Human Memory: Theory and Data*. Potomac, Maryland: Lawrence Erlbaum Associates, 1974.

Östberg P, Fernaeus SE, Hellström Å, Bogdanović N, and Wahlund LO. Impaired verb fluency: a sign of mild cognitive impairment. *Brain and Language*, 95: 273–279, 2005.

Östberg P, Crinelli RM, Danielsson R, Wahlund LO, Bogdanović N, and Fernaeus SE. A temporal lobe factor in verb fluency. *Cortex*, 43: 607–615, 2007.

Piatt AL, Fields JA, Paolo AM, Koller WC, and Tröster AI. Lexical, semantic, and action verbal fluency in Parkinson's disease with and without dementia. *Journal of Clinical and Experimental Neuropsychology*, 21: 435–443, 1999a.

Piatt A, Fields JA, Paolo AM, and Tröster AI. Action (verb naming) fluency as an executive function measure: convergent and divergent evidence of validity. *Neuropsychologia*, 37: 1499–1503, 1999b.

Piatt A, Fields JA, Paolo AM, and Tröster AI. Action verbal fluency normative data for the elderly. *Brain and Language*, 89: 580–583, 2004.

Pompeia S, Rusted JM, and Curran HV. Verbal fluency facilitated by the cholinergic blocker, scopolamine. *Human Psychopharmacology*, 17: 51–59, 2002.

Pulvermüller F. *The Neuroscience of Language*. Cambridge: Cambridge University Press, 2003.

Ravdin LD, Katzen HL, Agrawal P, and Relkin NR. Letter and semantic fluency in older adults: effects of mild depressive symptoms and age-stratified normative data. *The Clinical Neuropsychologist*, 17: 195–202, 2003.

Squire LR and Zola SM. Episodic memory, semantic memory, and amnesia. *Hippocampus*, 8: 205–211, 1998.

Tomer R and Levin BE. Differential effects of aging on two verbal fluency tasks. *Perceptual and Motor Skills*, 76: 465–466, 1993.

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

Tulving E. Multiple memory systems and consciousness. *Human Neurobiology*, 6: 67–80, 1987.

Wahlund LO, Pihlstrand E, and Jönhagen ME. Mild cognitive impairment: experience from a memory clinic. *Acta Neurologica Scandinavica*, 107(Suppl. 179): 21–24, 2003.