

Do item-level Category Fluency scores predict performance in other cognitive tests? A study of MCI informed by Cluster Analysis

Matteo De Marco¹, Akeliah Smith², Daniel J. Blackburn³, Annalena Venneri^{3,4}

¹Department of Psychology, Brunel University of London, UK;

²Department of Computer Science, Brunel University of London, UK

³Department of Neuroscience, University of Sheffield, UK

⁴Dipartimento di Medicina e Chirurgia, Università degli Studi di Parma, Italy

INTRODUCTION

We hypothesised that item-level Category Fluency (CF) scores would predict performance in other neuropsychological tests, and that semantic metrics would be better descriptors of CF word complexity than non-semantic metrics.

MATERIALS AND METHODS

CF words from 266 participants (Table 1) with Mild Cognitive Impairment were scored as a function of 11 semantic (Box 1) and 5 non-semantic (Box 2) metrics (Table 2, Fig. 1).

Partitional k -means Cluster Analyses were carried out on these 16 variables to capture word complexity in a unitary way and clusterise each cohort in a data-driven manner.

$k = 2$ and $k = 3$ were identified as the best k -solutions, as prompted by the within-cluster sum of squares "elbow" point, and by the Silhouette method (Fig. 2). Non-parametric inferential tests (Wilcoxon Signed Rank Test for $k = 2$, and Kruskal Wallis Test for $k = 3$, conservatively thresholded at a $p < 0.001$) were run to compare, across the resulting clusters (Fig. 3), performance in 10 neuropsychological tests. Additional models were run to compare individual metrics.

Table 1. Clinical description of the sample

Variable	Average	SD	Min	Max
<i>Demographic and General Clinical Variables</i>				
Age at Enrolment (years)	64.65	8.32	51	89
Educational Attainment (years)	12.39	2.88	4	24
Sex (F/M frequencies)	117/149			
Mini Mental State Examination	26.68	1.77	24	30
<i>Neuropsychological Test Scores - Not Used as Part of the Analyses</i>				
Confrontation Naming Test	19.04	1.58	3	20
Digit Span Test - Forwards	5.76	1.19	3	9
Digit Span Test - Backwards	4.09	1.07	2	8
Token Test	32.4	2.75	20	36
Pyramids and Palm Trees Test	50.28	1.9	38	52
Visuoconstructive Praxis Test	11.46	2.21	3	14
Stroop Test - Error Interference	2.76	5.41	0	29
<i>Neuropsychological Test Scores - Used as Part of the Analyses</i>				
Category Fluency Test	23.94	7.13	2	46
Digit Cancellation Test	44.83	9.82	18	60
Letter Fluency Test	28.39	13.3	1	79
Stroop Test - Time Interference	34.21	34.96	0	473
Raven Coloured Progressive Matrices	27.68	5.32	12	36
Rey-Osterrieth Complex Figure - Copy	29.16	6.23	5.5	43
Rey-Osterrieth Complex Figure - Recall	10.71	5.93	0	26.5
Paired Associated Learning Test	10.99	4.67	1	22
Logical Memory Test - Immediate Recall	9.33	4.46	0	35
Logical Memory Test - Delayed Recall	11.15	5.6	0	26

Box 1. Description of the 11 semantic item-level metrics included in this study

CFT Item-Level Metric	Definition	Reference
Age of Acquisition	Average estimated age at which people would have understood the meaning of a standardised entry if somebody had used it in front of them.	Kuperman et al., 2012
Concreteness	Degree to which the referent of a standardised entry can be experienced "directly through one of the five senses".	Brysbaert et al., 2014
Frequency	Logarithmic value corresponding to the "frequency per million words" of standardised entries, as per 2008-2014 BBC channel subtitles.	van Heuven et al., 2014
Prevalence	z-transformed proportion of participants that would correctly recognise the standardised entry as a word that they know.	Brysbaert et al., 2019
Recognition Time	z-transformed average response time taken by study participant to identify a standardised entry as a word that they know.	Mandera et al., 2020
Valence	Degree to which a standardised entry makes people feel "happy, pleased, satisfied, contented, hopeful", as opposed to "completely unhappy, annoyed, unsatisfied, melancholic, despaired, or bored".	Warriner et al., 2013
Arousal	Degree to which a standardised entry makes people feel "stimulated, excited, frenzied, jittery, wide-awake, or aroused", as opposed to "relaxed, calm, sluggish, dull, sleepy, or unaroused".	Warriner et al., 2013
Dominance	Degree to which a standardised entry makes people feel "controlled, influenced, cared-for, awed, submissive, or guided", as opposed to "in control, influential, important, dominant, autonomous, or controlling".	Warriner et al., 2013
Body Object Interaction	Degree to which a human body can physically interact with what a standardised entry represents.	Pexman et al., 2019
Contextual Diversity	Log-transformed average of cosines (indicating vector similarity), calculated from pairs of semantic vectors of "1,000-word contexts" in which the standardised entry appears.	Hoffman et al., 2013
Iconicity	Degree to which the standardised entry "sounds like its meaning".	Winter et al., 2024

Table 2. Item-level metrics

CFT Item-Level Metric	Average	SD	Min	Max
<i>Semantic Metrics</i>				
Age of Acquisition	5.21	0.45	4.05	6.69
Concreteness	4.89	0.03	4.8	4.95
Frequency	4.07	0.15	3.69	4.67
Prevalence	2.32	0.08	1.94	2.44
Recognition Time	-0.61	0.03	-0.66	-0.49
Valence	6.19	0.17	5.66	6.8
Arousal	4.09	0.2	3.62	5.34
Dominance	5.35	0.12	5	5.83
Body Object Interaction	5.18	0.2	4.59	6.05
Contextual Diversity	1.46	0.04	1.36	1.55
Iconicity	3.66	0.13	3.05	4.12
<i>Non-Semantic Metrics</i>				
Graphemic Length	5.68	0.38	3.5	6.67
Number of Syllables	1.86	0.17	1	2.36
Consonant-Vowel Ratio	1.44	0.13	1.1	2.5
Consonant Complexity	6.41	0.67	4.58	8.32
Orthographic Neighbourhood	7.22	1.51	3	12.8

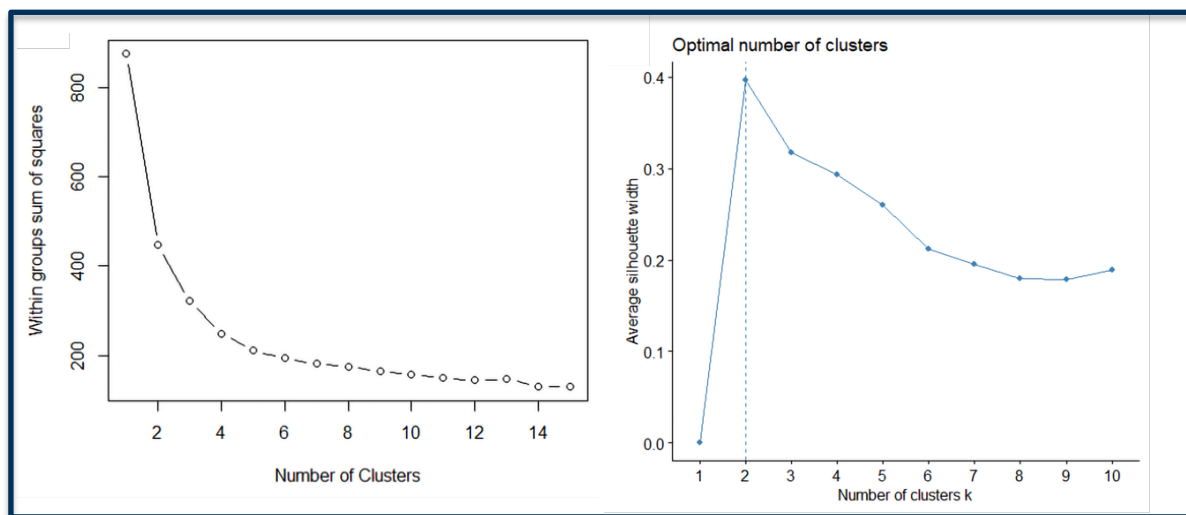


Fig. 2. k -solutions from the Elbow and Silhouette methods

Box 2. Description of the 5 non-semantic item-level metrics included in this study

CFT Item-Level Metric	Definition	Reference
Graphemic Length	Number of graphemes the written form of the standardised entry is based on.	N/A
Number of Syllables	Number of syllables the phonological form of the standardised entry is based on.	N/A
Consonant-Vowel Ratio	Ratio between consonants and vowels the phonological form of the standardised entry is based on.	Dufau et al., 2015
Consonant Complexity	Average complexity of consonant clusters for each syllable of the phonological form of the standardised entry.	Riley et al., 2015
Orthographic Neighbourhood	Number of vocabulary entries that differ from a standardised entry by one grapheme, i.e., either added (e.g., "BEAR" - "BEARD"), removed (e.g., "BEAR" - "BAR"), or replaced (e.g., "BEAR" - "BEAT").	Yarkoni et al., 2008

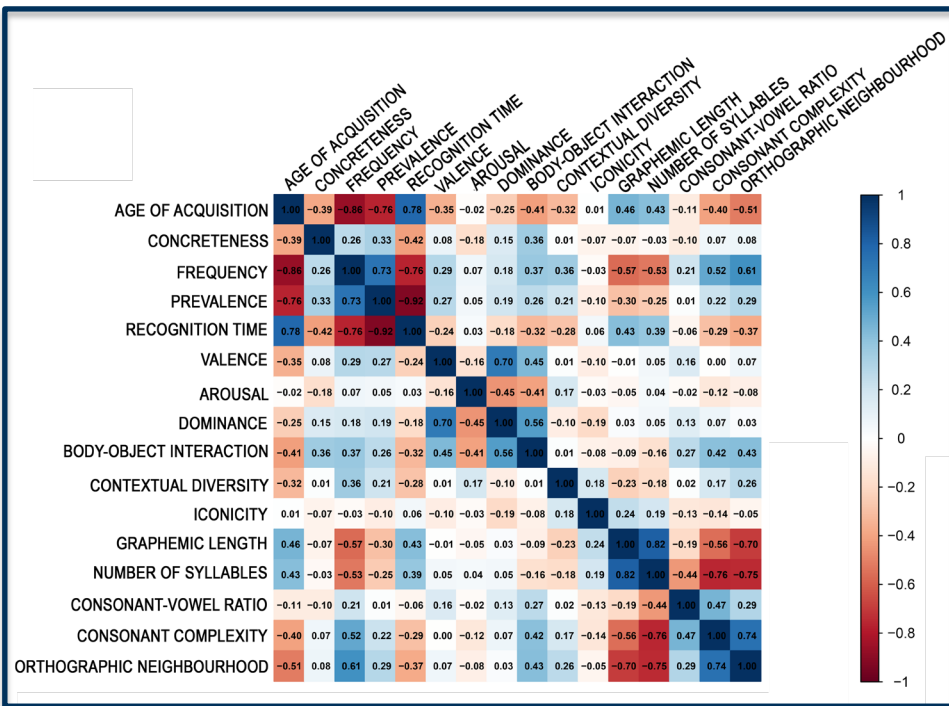


Fig. 1. Metric-to-metric correlations

RESULTS

Table 3. Cluster differences in cognitive and CF item-level metrics

Variable	$k = 2$	$k = 3$	$k = 3$ - post hoc
<i>Neuropsychological Test Performance</i>			
Category Fluency Test	4.61E-08 *	3.61E-08 *	K1 > K3 > K2
Digit Cancellation Test	0.054	0.135	
Letter Fluency Test	7.58E-04 *	6.64E-04 *	K1, K2 > K3
Stroop Test - Time Interference	0.053	0.21	
Raven Coloured Progressive Matrices	0.158	0.294	
Rey-Osterrieth Complex Figure - Copy	0.386	0.618	
Rey-Osterrieth Complex Figure - Recall	0.774	0.178	
Paired Associated Learning Test	0.002	0.002	
Logical Memory - Immediate Recall	0.009	0.003	
Logical Memory - Delayed Recall	0.129	0.022	
<i>CFT Item-Level Metrics</i>			
Age of Acquisition	5.17E-11 *	2.63E-14 *	K1 > K2 > K3
Concreteness	0.408	0.063	
Frequency	2.12E-15 *	2.20E-16 *	K3 > K2 > K1
Prevalence	1.19E-05 *	9.64E-11 *	K3 > K2 > K1
Recognition Time	1.19E-06 *	4.93E-12 *	K1 > K2 > K3
Valence	0.794	0.704	
Arousal	0.236	0.342	
Dominance	0.343	0.775	
Body Object Interaction	2.22E-06 *	6.29E-07 *	K3 > K2 > K1
Contextual Diversity	4.11E-04 *	1.04E-04 *	K2, K3 > K1
Iconicity	0.97	0.225	
Graphemic Length	2.20E-16 *	2.20E-16 *	K1 > K2 > K3
Number of Syllables	2.20E-16 *	2.20E-16 *	K1 > K2 > K3
Consonant-Vowel Ratio	0.006	0.005	
Consonant Complexity	2.20E-16 *	2.20E-16 *	K3 > K2 > K1
Orthographic Neighbourhood	2.20E-16 *	2.20E-16 *	K3 > K2 > K1
<i>General Cognitive Status</i>			
Mini Mental State Examination	0.080	0.044	



Fig. 3. $k = 2$ and $k = 3$ solutions

Between-cluster differences were found, under both k -solutions, in quantitative CF scores and Letter Fluency (LF) performance (Table 3, Fig. 4). Differences across clusters were found in relation to both semantic and non-semantic metrics.

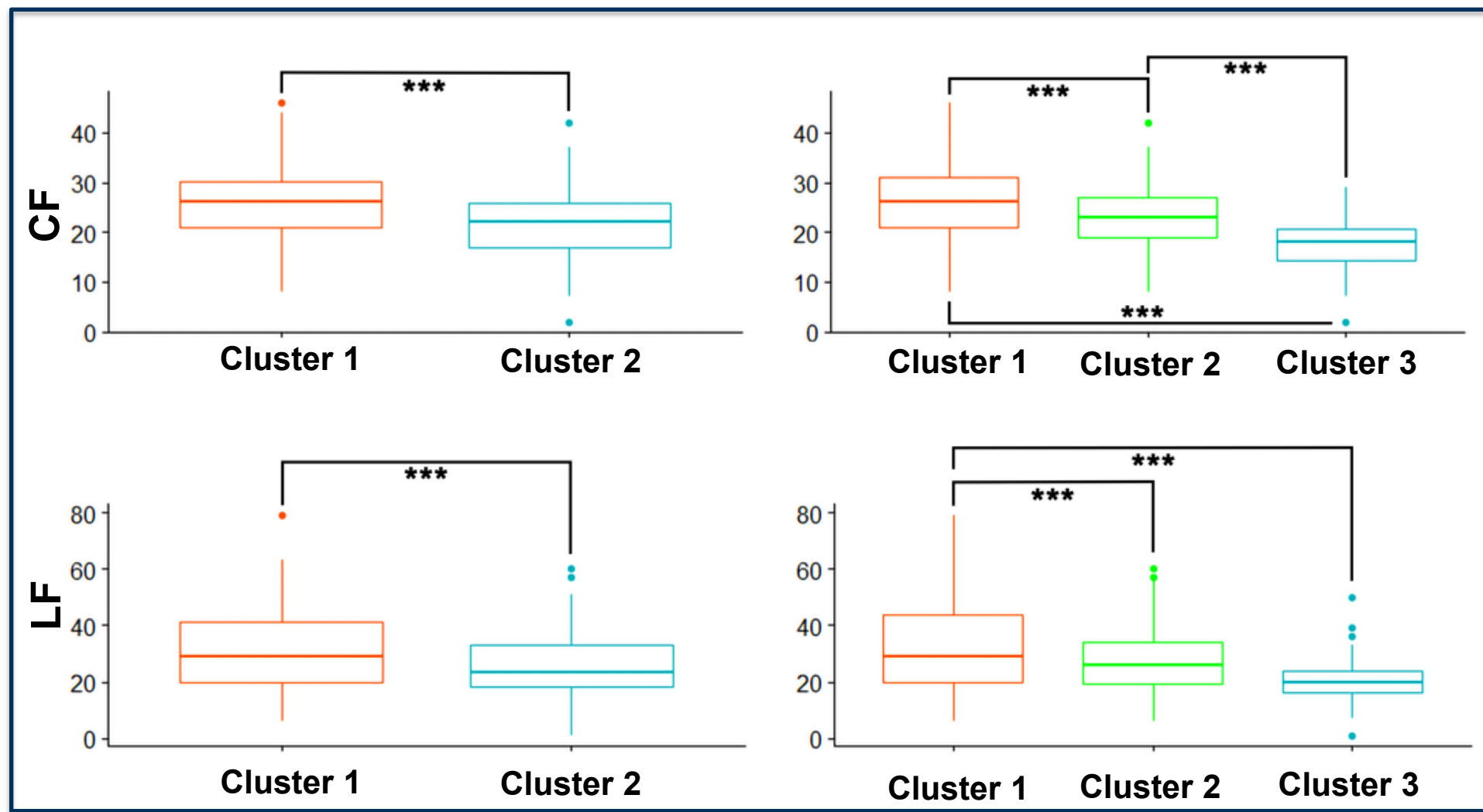


Fig. 4. Between-cluster effects in the two fluency tests

CONCLUSION

A large body of research has reported altered item-level CF scores in people with Mild Cognitive Impairment or dementia. This study shows that item-level scores can also predict quantitative LF performance in people with Mild Cognitive Impairment. These findings also indicate that both semantic and non-semantic metrics play a central role in characterising complexity of CF words.