

The verbal fluency task in the Greek population: Normative data, and clustering and switching strategies

MARY H. KOSMIDIS, CHRISTINA H. VLAHOU, PANAGIOTA PANAGIOTAKI,
AND GRIGORIOS KIOSSEOGLU

Department of Psychology, Aristotle University of Thessaloniki, Greece

(RECEIVED January 7, 2002; REVISED May 8, 2003; ACCEPTED May 26, 2003)

Abstract

We assessed 300 healthy adults in Greece on measures of semantic and phonemic verbal fluency in order to develop norms for the Greek population. We also evaluated the strategies that the participants used spontaneously in order to maximize word production, namely clustering and switching techniques. Our tasks comprised three semantic and three phonemic categories. Consistent with previous investigations of English-speaking samples, we found a contribution of demographic variables to word fluency. Specifically, level of education contributed to total word production, number of switches, and number of repetitive responses on both semantic and phonemic tasks, and the average cluster size only on the phonemic task. Age contributed to total word production and cluster size on the semantic task, and to number of switches on both semantic and phonemic tasks. Sex contributed only to total word production on the semantic task. In our sample, clustering and switching strategies were related to total word production on both tasks, suggesting that these strategies were used effectively. We present tables of normative data stratified by age and level of education. We have also included detailed guidelines for scoring clusters relevant to the Greek population. (*JINS*, 2004, *10*, 164–172.)

Keywords: Language, Greek, Verbal fluency, Semantic fluency, Phonemic fluency, Neuropsychological

INTRODUCTION

Verbal fluency tests are used extensively in clinical neuropsychological assessments, as well as in research protocols. Given their widespread use, it is important that appropriate norms for each version are available. Even in English, various letters may differ in their associative value; thus, norms for one set of letters may be inappropriate for another set (Tombaugh et al., 1999). Several groups of investigators have already developed norms for languages other than English: Spanish (Acevedo et al., 2000), Indian (Ratcliff et al., 1998), Flemish (Lannoo & Vingerhoets, 1997) and others, yielding varying results.

Differences in verbal fluency scores among various languages can be attributed to a multitude of factors. In a comparative study of bilingual individuals in New York, investigators found that Spanish speakers produced the smallest number of words compared to Chinese and English speak-

ers, while Vietnamese speakers generated the most words (Dick et al., 2002). The authors attributed this finding to the difference in word length of animal names in each language. Similarly, a study comparing French- and English-speaking Canadian patients on the “FAS” test reported significantly lower scores in the former compared to the latter group (Steenhuis & Ostbye, 1995). Other factors that may influence language differences in word production include culture-specific characteristics such as the degree of familiarity with testing procedures, the salience of test items, and behavioral expectations (Ardila, 1995). Investigators involved with cross-cultural comparisons point to potential differences in the types of experiences and environmental exposure that examinees may have had in different cultures and from which they tend to derive their responses (e.g., natural environment, mass media, etc.) (Acevedo et al., 2000).

Performance on fluency tests is influenced by demographic characteristics. Most studies confirm the contribution of age and education to word production (Cohen & Stanczak, 2000; Crossley et al., 1997; Kempler et al., 1998; Tombaugh et al., 1999; Tomer & Levin, 1993). Moreover,

Reprint requests to: Mary H. Kosmidis, Ph.D., Department of Psychology, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece.
E-mail: kosmidis@psy.auth.gr

semantic and phonemic fluency appear to be affected differentially by these variables: in one report, age accounted for more of the variance than education on semantic fluency, whereas education accounted for more of the variance than age on phonemic fluency (Tombaugh et al., 1999). Some investigators have reported sex differences in word production favoring women relative to men (Acevedo et al., 2000; Bolla et al., 1990), while other studies have failed to find such a difference regardless of task type (Cohen & Stanczak, 2000; Kempler et al., 1998; Tombaugh et al., 1999).

When investigating the cognitive mechanisms involved in word fluency, a common procedure is to evaluate patterns of performance on two different tests or the cognitive strategies used to maximize word generation on each test, rather than merely the total output. One approach to interpreting verbal fluency output is to compare performance on a semantic fluency test to performance on a phonemic test. Despite some commonalities, these two tasks require different cognitive processes. Adequate semantic fluency requires intact semantic memory stores and effective search processes. In contrast, phonemic fluency is less dependent on memory stores, and more dependent on effective initiation and shifting skills.

Another approach to understanding the mechanisms involved in optimal word generation is to examine the cognitive strategies used to complete the task successfully (Troyer, 2000; Troyer et al., 1998). Qualitative analyses of the process of producing words have shown that words are generated in spurts over time rather than at a consistent rate throughout the duration of the task (Gruenewald & Lockhead, 1980; Wixted & Rohrer, 1994). Successful examinees tend to search mentally for subcategories (semantic or phonemic, depending on the nature of the task), and, once identified, produce words within this subcategory. The process of organizing words into semantically or phonemically related subcategories is referred to as clustering (Troyer et al., 1997). Once a subcategory is exhausted, it is most efficient to quickly move to another subcategory or cluster (Bousefield & Sedgewick, 1994; Gruenewald & Lockhead, 1980; Troyer et al., 1997; Wixted & Rohrer, 1994), a tactic referred to as switching. As expected, both of these strategies are positively correlated with the total number of words produced (Robert et al., 1997; Troyer et al., 1997).

Each of the strategies used to maximize word production is mediated by separate brain mechanisms. The strategy of clustering words that are related to a subgroup depends on processes such as verbal memory and word storage. Switching requires the ability to engage in strategic search processes, such as initiation, cognitive flexibility, and mental shifting (Troyer et al., 1997). Both clustering and switching appear to play an important role in semantic fluency, whereas switching appears to be more important than clustering in phonemic fluency.

Our aim in undertaking the present study was to create culture- and language-specific norms for the verbal fluency test for the Greek population. In pursuing this goal, we

sought to develop norms for the entire adult age range and all educational levels. An additional goal was to adjust previous cluster scoring guidelines (Robert et al., 1998) to the types of responses most prevalent in a Greek sample. We expected the output of our sample to be reduced in comparison with those from other countries, since most of our participants would be unfamiliar with such testing procedures, and the words generated would most likely be polysyllabic. We also expected that the choice of words would reflect the types of stimuli (e.g., *animals*, *fruit*) that are more prevalent in the natural and social environment, and, thus, different from those reported previously. To our knowledge, there have been no attempts to date to develop normative data for a Greek verbal fluency test.

METHODS

Research Participants

We assessed 312 community dwelling adults with a brief neuropsychological test battery. The experimenters approached potential healthy participants in a large metropolitan area (sample of convenience), with the goal of including a broad range of adult ages and education levels. Screening consisted of a brief interview in order to exclude from our sample all those that reported a history of a neurological or psychiatric diagnosis, a closed head injury, or any conditions that might indicate cognitive impairment. We were left with a total of 300 healthy participants (140 men). Men and women did not differ significantly in age [$t(298) = -.08, p = .937$; men: $M = 46.4$ years ($SD = 18.7$); women: $M = 46.6$ years ($SD = 16.3$)], or in the highest level of education achieved [$t(298) = .32, p = .748$; men: $M = 11.3$ years ($SD = 4.6$); women: $M = 11.1$ years ($SD = 4.2$)]. All participants reported that Greek was their dominant language and gave their written informed consent to participate in the study.

Procedure

The testing was conducted in Greek. We administered a word fluency test comprising two parts. On the semantic test, we asked participants to generate as many different words as possible belonging to each of the following three semantic categories: animals, fruit, and objects. On the phonemic test, we asked participants to generate as many different words as possible beginning with each of the following three Greek letters: χ (Chi), Σ (Sigma), and α (Alpha). The letters were selected based on the ratio of words in the Greek language starting with these letters relative to the total number of words in a Greek dictionary, which corresponds to the ratio of words in the English language beginning with the letters *F*, *A*, and *S* relative to the total number of words in an English dictionary.

We instructed participants to begin generating items verbally as soon as the researcher announced the category or letter, and to avoid repetitions, variations of the same word,

and proper nouns (on the phonemic test). Examiners allowed 60 s for each trial. We gave no guidelines regarding how the participants were to organize their word search and production, to ensure that any cognitive strategies they used would be spontaneous. The semantic test was administered prior to the phonemic test, and categories and letters were administered in the aforementioned order for all participants.

In scoring test performance, we considered any identical or variations of a previously given word (e.g., act-acting) repetitions. Other types of errors were proper nouns or items irrelevant to the designated category or letter (e.g., a vegetable instead of a fruit, a word beginning with a letter other than that designated), which we considered rule infractions, and did not count in the total number of words generated.

Two of the authors (C.H.V. and P.P.) scored the tests for cluster size and number of switches, achieving an inter-rater reliability score of $r = .91$. We generally followed the cluster scoring guidelines reported by Robert and colleagues (1998), but observed that cultural factors influenced the types of clusters most frequently given by our sample and adjusted our scoring criteria accordingly. We excluded repetitions and intrusions when calculating clusters and switches, according to the rationale and scoring procedure proposed by Robert and colleagues (1998). A detailed description of the categories used and the scoring procedure is provided in the appendix. We calculated average semantic cluster size and number of semantic switches only for the semantic fluency test, and average phonemic cluster size and number of phonemic switches only for the phonemic fluency test.

RESULTS

We explored the following variables in the statistical analyses: total number of words produced on the three semantic categories (*animals*, *fruit*, and *objects*); total number of words produced on the three phonemic categories (*X*, Σ , and *A*); total number of switches on all three semantic categories, as well as on all three phonemic categories; average cluster size for each set of three categories (semantic and phonemic); and total number of repetitive responses and total number of rule infractions for both semantic and phonemic tasks combined.

We explored potential correlations (Pearson coefficients) among test variables, which are presented in Table 1. As expected, the total number of words produced to semantic categories correlated positively with the total number of words produced to phonemic categories [$r = .64, p < .001$]; the more words generated during the semantic task, the more words generated to the phonemic task, as well. On the semantic fluency test, the total number of words generated correlated positively with the number of switches [$r = .73, p < .001$] and average cluster size [$r = .15, p < .01$]. As the number of words produced increased, so did the number of switches and the size of the clusters they made. Switches and average cluster size showed a negative correlation with

Table 1. Correlations of performance variables

Variable	Semantic		Phonemic			
	Total words	Cluster size	Switches	Total words	Cluster size	Switches
Semantic						
Total words	—	—		.64**		
Cluster size	.15*					
Switches	.73**	-.35**				
Phonological						
Total words				—	—	
Cluster size				.32**	—	
Switches				.90**	.09	

* $p < .01$.

** $p < .001$.

each other [$r = -.35, p < .001$]; larger clusters were associated with fewer switches.

On the phonemic test, the total number of words produced correlated positively with the number of switches [$r = .90, p < .001$], as well as with average cluster size [$r = .32, p < .001$]. As participants generated more words, they also produced more switches and larger clusters. On this task, switches and average cluster size were not correlated with each other [$r = -.09, p = .118$].

In order to determine the potential contribution of the factors education, age, and sex on test performance, we performed a series of stepwise linear regression analyses. The results of these analyses are presented in Table 2. On the semantic test, we found a significant effect of the factors education, age, and sex on total word production [$F(3,296) = 45.48, p < .001$]. Performance was increased in participants with more years of education, those who were younger, and those who were women. Switches were

Table 2. Contribution of age, education, and sex to verbal fluency

Variable	Factor	Standardized beta	<i>t</i>	<i>p</i>	<i>R</i> ²
Semantic fluency					
Total words	Education	.48	9.29	.001	
	Age	-.15	-2.88	.004	
	Sex	.14	2.85	.005	.31
Switches	Education	.36	6.79	.001	
	Age	-.29	-5.43	.001	.28
Cluster size	Age	.12	2.08	.038	.01
Phonemic fluency					
Total words	Education	.50	9.88	.001	.25
Switches	Education	.41	7.39	.001	
	Age	-.12	-2.25	.025	.22
Cluster size	Education	.14	2.46	.014	.02
Overall errors					
Repetitions	Education	.16	2.74	.006	.03
Errors	—				

affected by education and age [$F(2,296) = 58.31, p < .001$] (the more educated and the younger the participants, the better their performance). Age alone contributed to the average size of clusters [$F(1,297) = 4.34, p < .05$], with older participants forming larger clusters. On the phonemic test, we found an effect of education on total word production [$F(1,298) = 97.60, p < .001$] and average size of clusters [$F(1,296) = 6.07, p < .05$], suggesting better performance among those with a higher level of education. Switches were influenced by education and age [$F(2,295) = 40.87, p < .001$], with the more educated and younger participants achieving better performance. Finally, there was an effect of education on the overall number of repetitions [$F(1,297) = 7.53, p < .01$], wherein the more educated participants made more repetitions. None of the demographic variables contributed to the number of rule infractions made. Education accounted for 28% of the variance on total words generated on the semantic task, while age accounted for only 2% and sex for 1%. On the phonemic test, education accounted for 25% of the variance on total words generated, whereas neither age nor sex contributed to this test variable.

Given previous reports regarding sex differences in verbal fluency favoring women (Acevedo et al., 2000; Bolla et al., 1990), we compared men and women on the number of words produced on each semantic and phonemic category. Women showed an advantage over men only on the number of fruit generated [$t(297) = -3.30, p < .001$; men: $M = 12.3 (SD = 3.3)$; women: $M = 13.4 (SD = 2.6)$]. Accordingly, women made larger clusters than men in this category [$t(298) = -2.64, p < .01$; men: $M = 3.2 (SD = 1.2)$; women: $M = 3.6 (SD = 1.3)$], while the sexes did not

differ on the number of switches [$t(297) = -.72, p = .474$; men: $M = 6.9 (SD = 3.1)$; women: $M = 7.2 (SD = 3.0)$]. Participants did not differ on any other semantic or phonemic category based on sex.

We have listed normative data for both semantic and phonemic tests in Tables 3–5. We stratified our sample based on age according to graphs illustrating changes over the age range, yielding three groups: 18–39, 40–59, and 60–79 years of age. We also stratified our sample based on education so as to reflect actual school requirements and landmarks (compulsory education in Greece is 9 years): 1–9, 10–12, and 13 or more years. Due to the small number of participants over 70 years of age with a university education ($n = 3$), we did not include them in the normative tables. We ranked participants' total word production on semantic and phonemic fluency tasks by percentiles and presented the results stratified by age and education in Table 3. We did not stratify by sex, however, because its contribution was based only on one category and would have resulted needlessly in small cell sizes for some groups without providing useful information in return.

We also ranked the average cluster size and number of switches for both fluency tasks stratified by age and education, and presented data for these strategies in Tables 4 (semantic fluency test) and 5 (phonemic fluency test).

Finally, we conducted paired samples *t* tests to compare participants' performance on the semantic test to their corresponding performance on the phonemic test. Our sample produced more words overall [$t(299) = 30.50, p < .001$; semantic: $M = 49.26 (SD = 11.45)$; phonemic: $M = 32.43 (SD = 11.06)$], more switches [$t(297) = 4.65, p < .001$; semantic: $M = 29.71 (SD = 10.03)$; phonemic: $M = 27.00$]

Table 3. Normative verbal fluency data stratified by age and education: Total number of words on semantic and phonemic fluency tasks

%ile	Age 18–39 years						Age 40–59 years						Age 60–79 years					
	Education (years)						Education (years)						Education (years)					
	1–9 (<i>n</i> = 21)		10–12 (<i>n</i> = 28)		13–21 (<i>n</i> = 66)		1–9 (<i>n</i> = 34)		10–12 (<i>n</i> = 20)		13–21 (<i>n</i> = 29)		1–9 (<i>n</i> = 55)		10–12 (<i>n</i> = 26)		13–21 (<i>n</i> = 19)*	
	Sem.	Phon.	Sem.	Phon.	Sem.	Phon.	Sem.	Phon.	Sem.	Phon.	Sem.	Phon.	Sem.	Phon.	Sem.	Phon.	Sem.	Phon.
90	58	38	60	42	71	53	58	44	62	42	64	58	51	39	63	39	67	50
80	52	36	59	39	67	47	54	36	61	36	61	56	47	35	56	33	63	47
70	47	34	54	35	64	43	52	32	59	33	58	47	44	31	54	32	62	44
60	43	32	51	34	60	42	47	29	52	32	56	43	41	28	50	31	57	42
50	42	25	49	31	56	39	45	27	50	30	54	37	40	25	47	30	52	40
40	41	21	48	30	55	37	41	25	48	28	53	35	39	22	44	28	51	38
30	39	18	46	29	52	36	39	21	47	27	50	34	37	19	39	24	47	31
20	32	16	43	27	48	34	36	20	42	25	48	31	33	17	37	21	46	29
10	26	15	39	23	44	26	30	16	36	24	41	28	29	13	33	17	42	27
<i>M</i>	42.7	26.5	49.6	32.0	57.8	39.5	44.9	27.9	51.6	30.9	54.3	41.0	40.0	25.5	46.7	28.8	54.7	37.8
<i>SD</i>	10.9	9.6	8.0	7.3	10.1	9.6	10.2	10.5	10.3	7.2	8.8	11.0	7.9	10.2	10.6	7.9	10.1	8.3

Note. Sem. = semantic; Phon. = phonemic; *M* = mean; *SD* = standard deviation.

*Normative data for elderly individuals with a university education were based only on 60–69-year-old participants.

Table 4. Normative data for average semantic cluster size and number of switches stratified by age and education

%ile	Age 18–39 years						Age 40–59 years						Age 60–79 years					
	Education (years)						Education (years)						Education (years)					
	1–9 (n = 21)		10–12 (n = 28)		13–21 (n = 66)		1–9 (n = 34)		10–12 (n = 20)		13–21 (n = 29)		1–9 (n = 55)		10–12 (n = 26)		13–21 (n = 19)*	
	CL	SW	CL	SW	CL	SW	CL	SW	CL	SW	CL	SW	CL	SW	CL	SW	CL	SW
90	4.9	31	4.1	41	4.0	54	4.4	37	4.6	44	4.4	47	4.5	33	4.8	36	4.5	47
80	4.2	30	4.0	40	3.8	47	3.9	32	4.1	36	4.1	42	4.0	31	4.4	32	4.2	34
70	4.0	29	3.6	39	3.7	43	3.7	30	4.0	28	3.6	39	3.8	26	4.0	26	4.0	31
60	3.8	25	3.4	36	3.5	39	3.6	27	3.6	27	3.4	33	3.5	25	3.8	24	3.9	29
50	3.6	22	3.2	35	3.3	36	3.4	24	3.5	26	3.2	32	3.3	22	3.5	23	3.8	27
40	3.5	21	3.0	33	3.0	35	3.3	23	3.2	25	3.1	30	3.1	21	3.1	22	3.7	25
30	3.4	20	2.9	29	2.9	32	3.0	22	3.1	24	3.0	27	3.0	19	3.0	21	3.6	23
20	3.3	17	2.7	27	2.6	29	2.9	18	3.0	23	2.8	23	2.8	17	2.7	20	3.4	22
10	3.0	15	2.4	26	2.5	25	2.3	16	2.7	—	2.7	21	2.5	13	2.6	19	3.3	20
<i>M</i>	3.7	23.7	3.3	33.9	3.3	38.0	3.3	25.9	3.6	29.7	3.4	32.7	3.4	23.0	3.6	25.0	3.9	28.7
<i>SD</i>	.7	7.4	.6	6.3	.6	10.4	.8	8.2	.6	8.8	.7	9.8	.7	6.8	.9	6.7	.7	8.9

Note. CL = average cluster size; SW = switches; *M* = mean; *SD* = standard deviation.

*Normative data for elderly individuals with a university education were based only on 60–69-year-old participants.

(*SD* = 10.13)], and larger clusters [$t(297) = 2.68, p < .01$; semantic: $M = 3.43$ (*SD* = .71); phonemic: $M = 3.13$ (*SD* = 1.85)] on the semantic as compared with the phonemic task.

DISCUSSION

We collected normative data for a verbal fluency test in a Greek sample of healthy adults. Our sample covered a broad range of ages and education levels, so as to maximize the representativeness of our norms. We also evaluated the cog-

nitive strategies that participants in our study used to optimize their word production, namely, clustering and switching tactics.

Our data are generally consistent with previous findings regarding the influence primarily of education and age on verbal fluency scores (Bolla et al., 1990; Cohen & Stanczak, 2000; Kempler et al., 1998; Tombaugh et al., 1999; Troyer, 2000), although we did not find the differential effect of these factors on semantic and phonemic tasks reported by others (Tombaugh et al., 1999). In the present

Table 5. Normative data for phonemic cluster size and number of switches stratified by age and education

%ile	Age 18–39 years						Age 40–59 years						Age 60–79 years					
	Education (years)						Education (years)						Education (years)					
	1–9 (n = 21)		10–12 (n = 28)		13–21 (n = 66)		1–9 (n = 34)		10–12 (n = 20)		13–21 (n = 29)		1–9 (n = 55)		10–12 (n = 26)		13–21 (n = 19)*	
	CL	SW	CL	SW	CL	SW	CL	SW	CL	SW	CL	SW	CL	SW	CL	SW	CL	SW
90	5.7	36	6.8	34	4.1	47	4.3	38	6.0	33	5.0	51	4.1	37	4.8	28	7.2	43
80	5.0	30	4.3	32	4.0	41	4.0	32	5.0	30	4.4	47	3.7	30	4.3	27	5.0	40
70	4.4	27	4.0	31	3.7	38	3.4	27	4.7	26	4.0	33	3.0	27	4.0	25	4.5	36
60	4.0	23	3.7	29	3.3	36	3.2	25	4.0	23	3.8	31	2.8	23	3.9	24	3.9	31
50	3.0	21	3.3	28	3.1	34	3.0	23	3.7	20	3.5	30	2.6	21	3.7	23	3.7	29
40	2.8	16	3.0	27	3.0	32	1.5	20	3.0	19	3.0	29	2.4	18	3.5	22	3.6	28
30	2.6	15	2.0	24	2.0	30	0	19	2.9	18	2.9	28	2.2	15	3.0	20	3.4	26
20	0	13	1.0	21	1.0	26	—	18	2.7	16	2.7	27	2.0	13	1.2	18	3.2	22
10	—	—	0	15	0	21	—	12	2.5	13	0	24	0	10	0	13	2.4	21
<i>M</i>	3.1	21.6	3.5	26.7	3.1	34.1	2.4	23.7	3.9	22.5	3.3	33.7	2.7	22.0	3.2	22.0	4.6	30.7
<i>SD</i>	2.0	8.2	1.8	6.9	1.3	9.4	1.8	9.2	1.6	7.8	1.6	9.9	1.6	9.9	1.8	5.6	3.5	8.2

Note. CL = average cluster size; SW = switches; *M* = mean; *SD* = standard deviation.

*Normative data for elderly individuals with a university education were based only on 60–69-year-old participants.

study, education appeared to be the most influential demographic factor as it contributed to most test variables, and, to a greater extent than age. A higher level of education was associated with increased total word production, number of switches, and repetitions on semantic and phonemic tests, and cluster size on the phonemic test. Age was also an important factor, as it contributed, albeit to a considerably lesser extent, to word production, switches, and cluster size on the semantic test, as well as to switches on the phonemic test; scores on these variables decreased with increasing age, with the exception of cluster size on the semantic test, which increased. Finally, in our sample, sex contributed to a very small extent to overall word production on the semantic, but not on the phonemic test. Upon more direct investigation, we found an advantage of women relative to men in the production of words only in the fruit category. This difference may reflect sociocultural factors, such as increased involvement in food procurement among women in Greek society. It is possible that women's familiarity with the seasons in which various fruit are available provided them with an effective clustering strategy, yielding increased output relative to men. This sex differentiation highlights the importance of the specific category used with respect to interpretation of word fluency results. Despite this isolated sex difference, we stratified our normative data only by age and education in tables with percentile equivalents for clinical use, in order to avoid creating very small cell sizes.

As reported in previous studies (Robert et al., 1998; Troyer et al., 1997), we found that the number of words produced both on the semantic and the phonemic tasks were related to clustering and switching strategies, and that these strategies were negatively related with each other on the semantic test, and unrelated to each other on the phonemic task. This pattern suggests that efficient use of clustering and switching strategies enhanced overall word production. Given the effectiveness of these cognitive strategies in test performance, it is important to evaluate them in addition to total word production scores when attempting to elucidate the reasons for poor fluency performance (i.e., mental initiation, organization skills, access to lexical memory stores). While these correlations may appear to indicate a potential confound of total word generation in estimating clustering and switching strategies, in fact, to correct either variable for the total number of words generated would be equivalent to correcting a cause by its effect, and, as such, inappropriate (Troyer, 2000; Troyer et al., 1997). Additionally, the raw number of switches and cluster size, rather than corrected scores, can be more informative clinically as they have been shown to be reduced in various patient groups (Tröster et al., 1998; Troyer, 1997).

We also found the expected task difference favoring semantic over phonemic verbal fluency. Our sample of healthy adults produced significantly more words (including more switches and greater cluster size) when given semantic categories than when given letters of the alphabet. The latter task may be more challenging in that it provides less struc-

ture to the individual conducting the word search than when given semantic categories, which restrict the range of potential words.

Our results are comparable to some of the norms published for other languages, but lower than others. Kempler and his colleagues (1998) reported a mean of 14 animals generated by an elderly sample with little education and 16.5 among the more highly educated. In our sample, the equivalent subgroup achieved 14.7 (low education), 16.8 (medium education) and 18.8 (high education). Relative to other studies, however, our sample generated fewer words than English (Acevedo et al., 2000; Tombaugh et al., 1999) and Spanish (Acevedo et al., 2000) samples of the same age and educational level. This could be attributed to the higher prevalence of polysyllabic words in Greek, as well as a decreased familiarity with such testing procedures. This finding emphasizes the importance of using norms specific to the task and the population being assessed.

The clinical utility of norms for the current version of the verbal fluency test is that they provide a reference point for neuropsychologists assessing verbal skills in Greek patients presenting with cognitive problems, rather than inappropriately relying on norms for English-speaking populations. Moreover, by presenting data for both semantic and phonemic tasks, as well as for clustering and switching strategies, we hope to provide useful information to assist in making differential diagnoses based on performance pattern rather than on individual scores.

Of course, the current norms are appropriate only for the categories and letters used in the present study, as other categories and letters may yield a different number of responses (Hart et al., 1988; Monsch et al., 1992). Caution should be used when applying these norms to individuals who are not native speakers of Greek (e.g., recent economic immigrants and political refugees to Greece, native Greek-speakers living abroad), as they may underestimate their abilities. Also, clinicians applying these norms should note that we excluded from our sample individuals who had no formal education and were illiterate, because there is evidence that suggests that illiterate individuals process verbal information in a qualitatively different manner (Kosmidis et al., 2003; Reis & Castro-Caldas, 1997).

There are several limitations of the present study that have to do with the selection of the sample. This study was based on a sample of convenience and exclusion criteria were based solely on self-report rather than on medical record review or structured interview. As in similar studies, this type of research runs the risk of sampling bias, regardless of recruitment method, since research volunteers may differ from the population at large in that those who are willing to participate may be more motivated than the average individual to do well on such a challenge or more curious regarding the procedure. Unfortunately, we did not record the demographic characteristics of those who were approached but refused to participate.

Another caveat is the relatively broad age range of the elderly subgroup, considering the large variability often en-

countered in their cognitive skills. We chose not to split this group further (i.e., 60–69 years and 70–79 years), due to the difficulties in recruiting participants over 60 with a university education. We did, however, exclude from the normative data participants over 70 years of age with a university education since there were only 2. Given the importance of establishing valid and reliable normative data for cognitive tests for the elderly, this is an issue that deserves to be addressed in a separate study with a larger sample of participants 60–90 years of age.

As the field and practice of clinical neuropsychology grows in Greece, more extensive normative studies will be needed to provide data that are valid for the population being assessed. Cross-cultural comparisons of test performance may also be useful in elucidating universal language-related cognitive mechanisms. The present investigation is a first step in this direction.

ACKNOWLEDGMENTS

This study was supported by a European Commission 5th Framework Programme awarded to the first author. We would like to thank Eleni Aslanidou, Kyriaki Dimitrakopoulou, Anastasia Emmanouilidou, Artemis Kaldeli, Dimitra Paliakoudi, Ekaterini Patsali, Chrysoula Tsakmakidou, and Kleopatra Tsiro for their contribution to the data collection. We would also like to thank the anonymous reviewers and Dr. Pagona Roussi for their helpful comments.

REFERENCES

- Acevedo, A., Loewenstein, D.A., Barker, W.W., Harwood, D.G., Luis, C., Bravo, M., Hurwitz, D.A., Aguero, H., Greenfield, L., & Duara, R. (2000). Category fluency test: Normative data for English- and Spanish-speaking elderly. *Journal of the International Neuropsychological Society*, *6*, 760–769.
- Ardila, A. (1995). Directions of research in cross-cultural neuropsychology. *Journal of Clinical and Experimental Neuropsychology*, *17*, 143–150.
- Bolla, K.I., Lindgren, K.N., Bonaccorsy, C., & Bleecker, M.L. (1990). Predictors of verbal fluency (FAS) in the healthy elderly. *Journal of Clinical Psychology*, *46*, 623–628.
- Bousefield, W.A. & Sedgewick, C.H.W. (1994). An analysis of restricted associative responses. *Journal of General Psychology*, *30*, 149–165.
- Cohen, M.J. & Stanczak, D.E. (2000). On the reliability, validity, and cognitive structure of the Thurstone Word Fluency Test. *Archives of Clinical Neuropsychology*, *15*, 267–279.
- Crossley, M., D'Arcy, C., & Rawson, N.S.B. (1997). Letter and category fluency in community-dwelling Canadian seniors: A comparison of normal participants to those with dementia of the Alzheimer or vascular type. *Journal of Clinical and Experimental Neuropsychology*, *19*, 52–62.
- Dick, M.B., Teng, E.L., Kempler, D., Davis, D.S., & Taussig, I.M. (2002). The Cross-Cultural Neuropsychological Test Battery (CCNB): Effects of age, education, ethnicity, and cognitive status on performance. In F.R. Ferraro (Ed.), *Minority and cross-cultural aspects of neuropsychological assessment* (pp. 17–41). Lisse, The Netherlands: Swets & Zeitlinger.
- Gruenewald, P.J. & Lockhead, G.R. (1980). The free recall of category examples. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 225–240.
- Hart, S., Smith, C.M., & Swash, M. (1988). Word fluency in patients with elderly dementia of Alzheimer type. *British Journal of Clinical Psychology*, *27*, 115–124.
- Kempler, D., Teng, E.L., Dick, M., Taussig, I.M., & Davis, D.S. (1998). The effects of age, education, and ethnicity on verbal fluency. *Journal of the International Neuropsychological Society*, *4*, 531–538.
- Kosmidis, M.H., Folia, V., Vlahou, C.H., & Kiosseoglou, G. (2003). *Semantic and phonological processing in illiteracy: A dichotomous or continuous process?* Manuscript submitted for publication.
- Lannoo, E. & Vingerhoets, G. (1997). Flemish normative data on common neuropsychological tests: Influence of age, education, and gender. *Psychologica Belgica*, *37*, 141–155.
- Monsch, A.U., Bondi, M.W., Butters, N., Salmon, D.P., Katzman, R., & Thal, L.J. (1992). Comparisons of the verbal fluency tasks in the detection of dementia of the Alzheimer type. *Archives of Neurology*, *49*, 1253–1258.
- Ratcliff, G., Ganguli, M., Chandra, V., Sharma, S., Belle, S., Seaberg, E., & Pandav, R. (1998). Effects of literacy and education on measures of word fluency. *Brain and Language*, *61*, 115–122.
- Reis, A. & Castro-Caldas, A. (1997). Illiteracy: A cause for biased cognitive development. *Journal of the International Neuropsychological Society*, *3*, 444–450.
- 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.
- Robert, P.H., Migneco, V., Marmod, D., Chaix-Medicin, I., Thauby, S., Benoit, M., Beau, C.H., & Darcourt, G. (1997). Verbal fluency in schizophrenia: The role of semantic clustering in category instance generation. *European Psychiatry*, *12*, 124–129.
- Steenhuis, R.E. & Ostbye, T. (1995). Neuropsychological test performance of specific diagnostic groups in the Canadian Study of Health and Aging (CSHA). *Journal of Clinical and Experimental Neuropsychology*, *17*, 773–785.
- Tombaugh, T.N., Kozak, J., & Rees, L. (1999). Normative data stratified by age and education for two measures of verbal fluency: FAS and Animal Naming. *Archives of Clinical Neuropsychology*, *14*, 167–177.
- Tomer, R. & Levin, B.E. (1993). Differential effects of aging on two verbal fluency tasks. *Perceptual and Motor Skills*, *76*, 465–466.
- Tröster, A., Fields, J.A., Testa, J. A., Paul, R.H., Blanco, C.R., Hames, K.A., Salmon, D.P., & Beatty, W.W. (1998). Cortical and subcortical influences on clustering and switching in the performance of verbal fluency tasks. *Neuropsychologia*, *36*, 295–304.
- Troyer, A.K. (2000). Normative data for clustering and switching on verbal fluency tasks. *Journal of Clinical and Experimental Neuropsychology*, *33*, 370–378.
- Troyer, A.K., Moscovitch, M., & Winocur, G. (1997). Clustering and switching as two components of verbal fluency: Evidence from younger and older healthy adults. *Neuropsychology*, *11*, 138–146.
- Troyer, A.K., Moscovitch, M., Winocur, G., Alexander, M., &

Stuss, D. (1998). Clustering and switching on verbal fluency: The effects of focal frontal and temporal-lobe lesions. *Neuropsychologia*, 36, 499–504.

Wixted, J.T. & Rohrer, D. (1994). Analyzing the dynamics of free recall: An integrative review of the empirical literature. *Psychonomic Bulletin & Review*, 1, 89–106.

Appendix

SCORING RULES FOR CLUSTERING AND SWITCHING

Semantic Clusters and Switches

We considered three or more consecutive words belonging to the same semantic subcategory a semantic cluster. We calculated semantic switches (SW; number of transitions between clusters, including single words) by subtracting the total number of related words (RW; all words belonging to a semantic cluster) from the total word production (WP) and adding that to the number of semantic clusters (SC): $(WP - RW) + SC = SW$.

Two of the authors (C.H.V. & P.P.) determined subcategory groups based on naturally occurring clusters in the participants' protocols and familiarity of Greek individuals with items. For example, most Greeks will be familiar with a variety of farm animals, while they may not be as familiar with animals of Africa or the Arctic/Far North. Accordingly, Greeks tend to group fruit based on the time of the year at which they are ripe. When two consecutive words with a strong association in the Greek language were mentioned, they, too, were considered a cluster. We created the following list as a guide in determining strong pairs of words, as well as semantic subcategories.

Animals

1. Strong pairs:
 - Cat–dog, bunny–rabbit, turtle–rabbit, cat–mouse, donkey–horse–mule* (two out of three), *mouse–rat, lion–tiger, hawk–eagle, fox–wolf, fox–chicken, wolf–lamb, and elephant–mouse*.
2. Subcategories:
 - (a) Farm animals: *cow, ox, goat, lamb, billy-goat, bull, chicken, rooster, dog, horse, donkey, mule, rabbit, duck, goose*, etc.
 - (b) Animals of the Greek forest: *wolf, bear, fox, squirrel, raccoon, skunk, wild boar, porcupine, deer, weasel, beaver, badger*, etc.; birds of this category such as *owl, eagle, hawk, cuckoo, crow*, etc.; and snakes of this category such as *viper*, etc.
 - (c) Tropical animals, animals of the steppe, animals of the jungle and safari animals: *crocodile, elephant, hippopotamus, rhinoceros, tiger, lion, puma, antelope, zebra, giraffe, buffalo, camel, kangaroo, koala*; primates such as *monkey, gorilla*, etc.; birds of this category such as *flamingo, parrot, vulture*, etc.; and snakes of this category such as *cobra, python*, etc.

- (d) Reptiles: *crocodile*, all types of *snakes, turtle, iguana, lizard, frog, chameleon*, etc.
- (e) Birds
- (f) Fish, including anything living underwater such as mammals (e.g., *dolphin, whale*) or shells
- (g) Insects

Fruit

1. Strong pairs:
 - Apple–orange, orange–tangerine, cherry–sour-cherry, apple–pear, melon–watermelon*.
2. Categories:
 - (a) Winter fruit:
 - i. Citrus fruits: *orange, tangerine, lemon, bitter orange, grapefruit*
 - ii. *Apple, pear, kiwi, quince*, etc.
 - (b) Spring and summer fruit:
 - i. Tropical fruit: *avocado, mango, pineapple, coconut, banana, papaya*, etc.
 - ii. *Peach, apricot, nectarine, cherry, strawberry, watermelon, fig, grape, berry, melon, pomegranate, cranberry, plum*, etc.
 - (c) Dry fruit: *fig, plum, walnut, peanut, hazelnut, almond*, etc.

Objects

1. Strong pairs:
 - Broom–dustpan, table–chair, cigarette–lighter–ashtray* (two out of three), *pencil–eraser, pencil–workbook, and hammer–nail*.
2. Categories:
 - (a) Furniture
 - (b) Appliances
 - (c) Clothes
 - (d) Linens–rugs: *curtain, pillowcase, bed-sheet, pillow, carpet, towel, doily, doormat, embroidery, tablecloth*, etc.
 - (e) Household items: *door, window, doorknob, blinds, chimney, fireplace, staircase, toiletries, gate, doorbell, lock, radiator, shutters*, etc.
 - (f) Kitchen items: *cookware, Tupperware®*, *utensils, cup, glass, kitchen appliances* (such as *oven, refrigerator*, etc.)
 - (g) Office items: *desk, chair, paper, computer, stationery, pen, pencil, eraser, notebook*, etc.
 - (h) Decorative items: *vase, chandelier, painting, poster, crystal-ware, porcelains, religious icons, ashtray, lamp*, etc.

- (i) Tools: *hammer, nail, drill, screwdriver, saw, rake, hose, shovel, pincer, spade, pick*, etc.
- (j) Vehicles
- (k) Jewelry
- (l) Cosmetics and accessories: *perfume, shampoo, hair band, comb, watch, wallet, after-shave, razor, sponge, nail polish*, etc.

Phonemic Clusters and Switches

We considered three or more consecutive words beginning with the same two letters and having the same sound (e.g., *gallant-gap-gas*), or two consecutive words that differed

only in a vowel sound (e.g., *rule-role*), or words that were homophones (e.g., *route-root*) as a phonemic cluster. We estimated phonemic switches (SW) by subtracting the total number of words related to each phonemic cluster (RW) from the total phonemic word production (WP) and adding that to the number of phonemic clusters (PC): $(WP - RW) + PC = SW$.

If two or more successive words stemmed from the same root (such as *act-action-acting*), we considered them repetitions, and thus did not calculate a cluster based on them. If the words only shared a part/suffix but had a different meaning (e.g., *superman-supermarket-supercilious*), however, we considered them a cluster.