Category fluency test: normative data for English- and Spanish-speaking elderly

Article *in* Journal of the International Neuropsychological Society · December 2000

Source: PubMed

10 authors, including:

Amarilis M. Acevedo
Nova Southeastern University
80 PUBLICATIONS **5,115** CITATIONS

David Andrew Loewenstein
University of Miami
421 PUBLICATIONS **15,300** CITATIONS

Warren Barker
Mount Sinai Medical Center
238 PUBLICATIONS **8,031** CITATIONS

Ranjan Duara
Mount Sinai Medical Center, Miami Beach
164 PUBLICATIONS **8,557** CITATIONS

Journal of the International Neuropsychological Society (2000), **6**, 760–769.

Category Fluency Test: Normative data for English- and Spanish-speaking elderly

AMARILIS ACEVEDO,[1,3] DAVID A. LOEWENSTEIN,[1,2,3] WARREN W. BARKER,[2]
DYLAN G. HARWOOD,[2] CHERYL LUIS,[2] MARINA BRAVO,[2] DEBORAH A. HURWITZ,[2]
HILDA AGUERO,[2] LYNDA GREENFIELD,[2] AND RANJAN DUARA [2,3,4]

[1]Cognitive Rehabilitation Program, Horizon Health Corporation, Mount Sinai Medical Center, Miami Beach, Florida
[2]The Wien Center for Alzheimer's Disease and Memory Disorders, Mount Sinai Medical Center, Miami Beach, Florida
[3]Department of Psychiatry and Behavioral Sciences, University of Miami School of Medicine, Miami, Florida
[4]Department of Medicine, University of Miami School of Medicine, Miami, Florida

(Received August 16, 1999; Revised November 10, 1999; Accepted November 11, 1999)

Abstract

Category fluency tasks are an important component of neuropsychological assessment, especially when evaluating for dementia syndromes. The growth in the number of Spanish-speaking elderly in the United States has increased the need for appropriate neuropsychological measures and normative data for this population. This study provides norms for English and Spanish speakers, over the age of 50, on 3 frequently used measures of category fluency: *animals*, *vegetables*, and *fruits*. In addition, it examines the impact of age, education, gender, language, and depressed mood on total fluency scores and on scores on each of these fluency measures. A sample of 702 cognitively intact elderly, 424 English speakers, and 278 Spanish speakers, participated in the study. Normative data are provided stratified by language, age, education, and gender. Results evidence that regardless of the primary language of the examinee, age, education, and gender are the strongest predictors of total category fluency scores, with gender being the best predictor of performance after adjusting for age and education. English and Spanish speakers obtained similar scores on animal and fruit fluency, but English speakers generated more vegetable exemplars than Spanish speakers. Results also indicate that different fluency measures are affected by various factors to different degrees. (*JINS*, 2000, *6*, 760–769.)

Keywords: Neuropsychological tests, Spanish speakers, Category fluency test, Normative data

INTRODUCTION

In the past few decades, a number of category fluency tasks have been developed to assess the ability of individuals to retrieve exemplars of a given semantic category under time limits. Category fluency tasks have been found to be helpful in assessing aphasia (Goodglass & Kaplan, 1983), in the detection (Crossley et al., 1997; Monsch et al., 1992; Ober et al., 1986), staging (Welsh et al., 1992), and differential diagnosis of dementia syndromes (Pasquier et al., 1995; Stern et al., 1993), and in the qualitative analysis of organizational strategies when speeded access to semantic information is required (Binetti et al., 1995; Carew et al., 1997; Massman et al., 1992; Tröster et al., 1998). Performance on categorical verbal fluency tasks has been found to be affected by a number of factors, including age (Crossley et al., 1997; Tomer & Levin, 1993), gender (Bolla et al., 1998), educational attainment (Crossley et al., 1997), verbal intellectual ability (Bolla et al.,1998), depression (see King & Caine, 1990), and race or ethnicity (Johnson-Selfridge et al., 1998).

Comparison across studies that have assessed category fluency in different populations remains difficult given the wide variability of the parameters used in the administration of the task. One of these parameters is the number of categories that are sampled. Reports in the literature vary from one (Crossley et al., 1997, Johnson-Selfridge et al., 1998), two (Ober et al., 1986, Tomer & Levin, 1993), three (Monsch et al., 1992; Stern et al., 1993) and four (Huff et al., 1986) categories. The number of categories is important, given that a larger sampling procedure usually increases the reliability of the measure and may increase a test's sensi-

Reprint requests to: Ranjan Duara, M.D., Wien Center, Mount Sinai Medical Center, 4300 Alton Road, Miami Beach, FL 33140. E-Mail: billy@msmc.com

tivity and specificity (see Monsch et al., 1992). Duration of the trial also varies across investigations, with most studies using 60 s, whereas others use 30 s (Monsch et al., 1997), 90 s (Ober et al., 1986) or 120 s (Pasquier et al., 1995). The specific categories sampled may also differ, with *animals* being the most popular category used either alone or in combination with other categories. Different research groups use various combinations of categories in conjunction with *animals*, including *fruits* and *vegetables* (Bayles et al., 1989; Bolla et al., 1998; Hodges et al., 1990; Kempler et al., 1998; Lucas et al., 1998; Massman et al., 1992; Monsch et al., 1992) or foods and clothing (Stern et al., 1993; Stricks et al., 1998). Less frequently used categories include *vehicles* (Huff et al., 1986; Rosser & Hodges, 1994), *tools* (Huff et al., 1986), *things in a supermarket* (Mattis, 1988, Randolph et al., 1993), *things people drink* (Randolph et al., 1993), *musical instruments* (Rosser & Hodges, 1994), *first names* (Monsch et al., 1992), and *occupations and furniture* (Gurd & Ward, 1989).

Another factor that limits comparisons across studies is the variability in instructional sets. For example, some instructions include a specific exemplar of the category (e.g., providing the word "dog" when asking respondents to name animals; see Ardila et al., 1994; Goodglass & Kaplan, 1983). Other instructional sets further provide the subject with subcategories (e.g., *animals from the farm or jungle*, *house pets*, *birds*, and *fish*; see Goodglass & Kaplan, 1983; Jacobs et al., 1997; Unverzagt et al., 1999), potentially altering the difficulty of the task. In fact, it has been shown that, at least in certain neurological conditions, provision of within-category cues significantly improves retrieval of exemplars (Randolph et al., 1993).

Studies also vary in a number of other important parameters, including groups composition by age, education, or intellectual characteristics, as well as the use of different exclusionary criteria in participant selection (see Mitrushina et al., 1999, for a thorough discussion on this topic). Finally, although the fluency score is typically based on the number of exemplars produced during the total duration of the trial, a few paradigms are based on the number of exemplars generated during the most productive time segments within the trial (see Goodglass & Kaplan, 1983).

By their very nature, verbal fluency tasks are heavily dependent on the examinee's language skills. About 46% of all foreign-born persons residing in the United States are of Hispanic origin (U.S. Bureau of the Census, 1997), and a high proportion of these do not speak English fluently. Individuals of Hispanic background represent 11% of the total United States population at present (Reed & Ramirez, 1998). In the next decades, a larger proportion of Hispanics than Asians or Blacks residing in the United States will be moving into the elderly age range (Day, 1996). Thus, neuropsychologists will be increasingly faced with the dilemma of evaluating patients whose primary language is Spanish, the most frequent non-English language in which neuropsychological services are provided in the United States (Echemendia et al., 1997). For a neuropsychologist bilingual in English and Spanish, the paucity of standardized neuropsychological tests for Spanish speakers and the dearth of norms for this population limits his or her ability to effectively interpret test findings (Loewenstein et al., 1993). Current limitations extend beyond language, as cultural factors affect the relevance, familiarity, and salience of test stimuli as well as social perceptions and behavioral expectations that may affect testing results (see Ardila, 1995; Echemendia et al., 1997, Loewenstein et al., 1994; Mahurin et al., 1992).

Few pioneer studies have examined verbal fluency in Spanish speakers. Rey and Benton (1991), as part of the Multilingual Aphasia Examination in Spanish, published norms on a letter fluency task for the letters PTM. They used a sample of 234 Spanish speakers, ages 18 to 70, who resided in Texas and Puerto Rico, and found that education, but not age, affected performance. Ponton et al. (1996), as part of their normative study for the Neuropsychological Screening Battery for Hispanics, assessed 300 Spanish speakers, most of whom were from Mexico. They found that phonological fluency, as assessed with the letters *F*, *A*, and *S*, was significantly affected by education, but not age or gender. More recently, Stricks et al. (1998) published norms for the letters *P*, *S*, and *V* in a normative study that included 416 elderly Spanish speakers residing in New York, most of whom were from the Dominican Republic, Cuba, and Puerto Rico (see Jacobs et al., 1997). Results indicated that age and education affected performance significantly.

Regarding category fluency, Ardila et al. (1994) reported fluency norms for the categories *animals* and *fruits* in 346 normal elderly residing in Colombia, South America. They found that, regardless of education, performance was 50% lower in participants in their late 70s as compared to those in their late 50s. Stricks et al. (1998) published norms for the categories *animals*, *foods*, and *clothing* on the Spanish-speaking sample described above, and found a significant interaction of Age × Education on total category fluency score. Ostrosky-Solis et al. (1999) recently published normative data for NEUROPSI, a neuropsychological battery normed on a sample of 800 monolingual Spanish speakers residing in Mexico, with participants aged 16 to 85 years and education from zero to 24 years. A significant effect of education and age, but not gender, was reported for category fluency for *animals* and for phonologic fluency for the letter *F*.

The present study extends existing age norms for English speakers on the three categories that have been most frequently used in neuropsychological studies: *animals*, *vegetables*, and *fruits*. Normative data using a more restricted age range have been recently published for these categories in elderly English speakers who are cognitively intact (see Bolla et al., 1998; Lucas et al., 1998). Fluency data on these categories have also been published in patients with Huntington's disease (Hodges et al., 1990; Massman et al., 1992) and probable Alzheimer's disease (Massman et al., 1992; Monsch et al., 1992). To our knowledge, this is the first normative study of these categories in Spanish speakers. As mentioned above, Stricks et al. (1998) published norms for other categories in Spanish speakers, but their study only offers norms for total fluency scores and for individuals over the age of 65. The present study provides norms for each individual category, which is important, given the suggestion by Bolla et al. (1998) that performance on different category tasks is differentially affected by various demographic variables. The present norms are expected to be helpful, given the high proportion of Spanish-speaking elderly residing in the United States and the fact that, at least in English speakers, this triad of categories has been shown to have high sensitivity and specificity in the detection of Alzheimer's disease (see Monsch et al., 1992).

METHODS

Research Participants

The normal control sample used in this study was drawn from a pool of 2332 community-dwelling individuals, ages 50 to 90 years, who presented for a free memory screening offered by the Wien Center for Alzheimer's Disease and Memory Disorders at either the Mount Sinai Medical Center in Miami Beach, Florida, or a satellite site in Aventura, Florida. Of these, 1455 (62%) were English speakers (ES) and 877 (38%) were Spanish speakers (SS). Participants learned about the screening program primarily through newspaper advertisement and were recruited between January, 1994 and March, 1999. Participants in the English-speaking group were required to speak English as their primary language and to have been born in the United States. Participants in the Spanish-speaking group were required to speak Spanish as their primary language and to have been born in a country where Spanish is the primary language. All persons included in the study were screened in their primary language.

The research participants in the present study were deemed cognitively normal after careful screening by a social worker or nurse who was fluent in the participant's primary language. The screen consisted of the Mini Mental Status Examination (MMSE; Folstein et al., 1975), the Hamilton Depression Rating Scale (HDRS; Hamilton, 1960), and questionnaires related to demographic information, medical and psychiatric history, and cognitive status. Consistent with recommendations by Kukull et al. (1994), all participants were required to have a total score of 27 or higher on the MMSE. In addition, participants were required to score a minimum cumulative score of 10 out of 12 points on four delayed-recall trials of the three words used on the MMSE. This cutoff has been shown to have over 96% sensitivity and over 90% specificity in differentiating between cases of dementia *versus* cases with no cognitive impairment (Loewenstein et al., 2000). Application of these psychometric criteria resulted in a total sample of 772 participants: 470 ES and 302 SS. Of these, separate scores for each individual category were available for 702 participants: 424 ES and 278 SS. Among these ES, 420 (99%) were classified by the examiner as White, 3 ($<$1%) as African American, and 1 ($<$1%) as Asian American. Among the SS, 269 (97%) were classified by the examiner as White, 2 ($<$1%) as Black, and 7 (3%) as "other." All individuals in the English-speaking group and most individuals in the Spanish-speaking group resided in the United States. A very small proportion of the SS resided in a Latin American country.

Procedure

Participants were asked to name as many different types of animals, vegetables, and fruits, in that order, as they could. Time was limited to a 60-s period for each category. The number of correct, nonrepeated responses for each individual category constituted the raw score for the specific category. The total category fluency score was calculated by adding the number of correct responses for the pooled categories.

Statistics

An analysis of variance (ANOVA) was conducted first to determine whether age and education had a significant main effect on total category fluency scores. Because there were only 14 SS at or above the age of 80, it was not possible to compare their performance, especially with regards to interaction terms, with the 105 ES who were in this age range. Thus, they were excluded from this analysis and the rest of the participants were grouped into three age ranges: 50 to 59, 60 to 69, and 70 to 79 years of age. Individuals with less than 8 years of education were also excluded from these analyses because of the disparity between the number of ES ($n = 1$) *versus* SS ($n = 29$) with less than 8 years of education. For the interested reader, means and standard deviations for *animals*, *vegetables*, *fruits*, and total fluency scores in SS with 7 or less years of education are provided: 15.6 $\pm$ 4.1, 11.9 $\pm$ 3.6, 11.2 $\pm$ 4.0, and 38.8 $\pm$ 10.1, respectively. For ES with 80 to 89 years of age, scores on these measures were 13.3 $\pm$ 3.7, 11.7 $\pm$ 3.3, 11.1 $\pm$ 3.4, and 36.0 $\pm$ 7.8, respectively. The rest of the sample was grouped into three educational levels: 8 to 12 years, 13 to 16 years, and 17 years or more. In total, scores on 553 subjects—316 ES and 237 SS—were included in subsequent univariate and multivariate analyses of variance or covariance.

To evaluate specific performance on each of the fluency measures (i.e., *animals*, *vegetables*, and *fruits*), a repeated measures multivariate analysis of covariance (MANCOVA) was conducted with age and education entered in as covariates. The three category fluency scores were treated as related dependent measures. Univariate main effects and interaction terms were only examined following significant multivariate F at $p < .05$. *Post-hoc* tests on group means or any interaction terms were conducted using the Tukey's b procedure.

A series of stepwise linear regression analyses were performed using all 702 participants to examine the relative and incremental effects of age, education, gender, lan-

Table 1. Demographic characteristics of the sample

Variable	English speakers ($n = 316$)	Spanish speakers ($n = 237$)
Age		
50–59	37	64
60–69	107	97
70–79	172	76
Mean age (± *SD*) (years)	69.1 (6.9)*	64.9 (7.7)
Education (years)		
8–12	112	105
13–16	154	94
17+	50	38
Mean education (± *SD*) (years)	14.4 (2.5)*	13.4 (3.2)
Gender (% female)	74.0%	69.2%
MMSE scores (*M* ± *SD*)	28.9 (1.0)	28.7 (1.0)
Total 4-Trial Recall (*M* ± *SD*)	11.2 (0.8)*	11.4 (0.7)
HDRS (*M* ± *SD*)	5.8 (4.2)*	7.9 (5.0)

Note. HDRS = Hamilton Depression Rating Scale.
$^*p < .001$.

guage, and scores on the HDRS on each category fluency measure.

RESULTS

Demographic information of the sample used in the analysis of variance and multivariate analysis of covariance is presented in Table 1. English speakers were older, had higher levels of educational attainment, and had lower scores on the HDRS than SS. Regarding the 237 participants in the Spanish-speaking group, 139 (58.6%) were born in Cuba and 98 (41.4%) were born in other Spanish-speaking countries. Of the latter, 31 participants were from Colombia (13.1%), 13 were from Peru (5.5%), 11 were from Puerto Rico (4.6%), 11 were from Argentina (4.6%), and 31 (13.1%) were from other Latin American countries. One participant was from Spain. Cubans did not differ from other SS in years of education, total MMSE scores, total recall scores, or scores on the HDRS, although they were older (*M* age = 66.8 ± 7.1) than the other SS (*M* age = 62.3 ± 7.9). Since Cubans did not differ from the other SS on total or separate verbal fluency scores, the groups were combined for subsequent analyses.

Tables 2 and 3 present the scores for ES and SS, respectively, for individual categories and for total scores for age, education, and gender. Tables 4 and 5 present the scores for individual categories and for total scores by gender and age and by gender and education for ES and SS, respectively.

An initial 3 × 3 × 2 × 2 (Education × Gender × Age × Language) ANOVA was conducted to examine the main ef-

Table 2. Fluency scores in English speakers

	Animals		Vegetables		Fruits		Total Fluency Score	
Variable	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Age (years)								
50–59 ($n = 37$)	18.4	4.9	16.0	4.1	16.0	4.1	50.4	10.6
60–69 ($n = 107$)	17.1	4.2	14.4	3.9	13.7	3.7	45.2	9.6
70–79 ($n = 172$)	15.2	4.3	13.6	3.5	12.5	3.1	41.3	8.4
Education (years)								
8–12 ($n = 112$)	15.0	4.3	14.2	3.8	13.0	3.1	42.2	8.7
13–16 ($n = 154$)	16.3	4.0	14.0	3.7	13.3	3.9	43.6	9.5
17+ ($n = 50$)	18.8	5.4	14.7	3.9	13.9	3.6	47.4	10.7
Gender								
Male ($n = 82$)	16.2	4.6	11.9	2.8	11.9	3.3	40.0	8.7
Female ($n = 234$)	16.3	4.5	15.0	3.7	13.8	3.6	45.0	9.5
Total *N*	16.2	4.5	14.2	3.8	13.3	3.6	43.7	9.6

Table 3. Fluency scores in Spanish speakers

	Animals		Vegetables		Fruits		Total Fluency Score	
Variable	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Age (years)								
50–59 ($n = 64$)	16.3	3.9	13.0	3.6	13.2	3.3	42.6	8.4
60–69 ($n = 97$)	17.2	5.3	13.1	4.0	13.4	3.4	43.6	10.0
70–79 ($n = 76$)	16.3	4.4	12.3	3.6	12.8	3.6	41.3	9.7
Education (years)								
8–12 ($n = 105$)	15.8	4.6	12.6	3.7	12.4	3.3	40.9	9.5
13–16 ($n = 94$)	17.1	4.1	13.1	3.9	13.8	3.5	44.0	9.6
17+ ($n = 38$)	17.7	5.7	12.8	3.4	13.5	3.2	44.0	8.8
Gender								
Male ($n = 73$)	16.6	5.7	11.3	3.8	12.2	3.5	40.1	10.7
Female ($n = 164$)	16.7	4.2	13.5	3.6	13.5	3.3	43.7	8.8
Total *N*	16.7	4.7	12.8	3.6	13.1	3.4	42.6	9.5

fects and interaction terms on total category fluency. There were no significant three-way interaction terms. The four-way interaction effect could not be reliably calculated because of a limited number of participants in various cells. There was a significant overall effect for the combined two-way interaction terms [$F(13,520) = 2.23$, $p < .009$], but the only single interaction term to approach significance was the Age $\times$ Gender effect [$F(2,520) = 2.95$, $p < .053$]. *Post-hoc* Tukey's *b* tests revealed that, in general, women demonstrated a greater decrease in verbal fluency scores than did males as a function of age. The only significant main effect was for gender [$F(1,520) = 2.17$, $p < .001$]. As a result, a series of ANOVAs were subsequently conducted that allowed a more in-depth analysis of main effects and interaction terms by allowing previously categorical independent variables (e.g., age and education) to take the form of interval level covariates.

Effects of Age and Education on Total Category Fluency Score

To determine the effects of age and education on the total category fluency score, a 3 $\times$ 3 ANOVA was conducted. There were significant main effects for age [$F(2,543) = 3.87$, $p < .03$], and education [$F(2,543) = 6.02$, $p < .01$]. No Age $\times$ Education interaction was found [$F(4,543) = 1.96$, $p > .10$]. Because Spanish-speaking participants were significantly younger and less educated than ES (see Table 1), both age and education were employed as covariates in subsequent analyses to examine possible language and gender effects.

Table 4. Fluency scores in English speakers by gender and age, and by gender and education

	Men: Age (years)			Women: Age (years)		
	50–59 ($n = 7$)	60–69 ($n = 30$)	70–79 ($n = 45$)	50–59 ($n = 30$)	60–69 ($n = 77$)	70–79 ($n = 127$)
Task	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
Animals	16.4 (3.3)	16.4 (4.9)	16.0 (4.7)	18.9 (5.1)	17.3 (3.9)	15.0 (4.2)
Fruits	12.3 (2.3)	11.7 (3.5)	11.9 (3.4)	16.9 (3.9)	14.4 (3.5)	12.7 (3.0)
Vegetables	11.7 (1.7)	11.8 (2.8)	12.0 (3.0)	17.0 (3.8)	15.4 (3.8)	14.2 (3.5)
Total Fluency	40.3 (4.5)	40.0 (9.7)	39.8 (8.6)	52.7 (10.2)	47.2 (8.8)	41.9 (8.3)
	Men: Education (years)			Women: Education (years)		
	8–12	13–16	17+	8–12	13–17	17+
Task	($n = 25$)	($n = 42$)	($n = 15$)	($n = 87$)	($n = 112$)	($n = 35$)
Animals	15.6 (4.4)	16.1 (4.4)	17.4 (5.8)	14.8 (4.3)	16.4 (3.9)	19.4 (5.2)
Fruits	11.9 (3.4)	11.7 (3.3)	12.3 (3.3)	13.3 (2.9)	13.9 (4.0)	14.6 (3.6)
Vegetables	12.2 (2.3)	11.7 (3.1)	12.0 (3.0)	14.8 (3.9)	14.8 (3.6)	15.9 (3.7)
Total Fluency	39.8 (8.3)	39.4 (8.8)	41.7 (9.3)	42.9 (8.8)	45.2 (9.3)	49.9 (10.4)

Table 5. Fluency scores in Spanish speakers by gender and age, and by gender and education

	Men: Age (years)			Women: Age (years)		
Task	50–59 (n = 15) *M* (*SD*)	60–69 (n = 32) *M* (*SD*)	70–79 (n = 26) *M* (*SD*)	50–59 (n = 49) *M* (*SD*)	60–69 (n = 65) *M* (*SD*)	70–79 (n = 50) *M* (*SD*)
Animals	15.5 (3.4)	18.0 (7.2)	15.4 (4.2)	16.6 (4.1)	16.7 (4.0)	16.7 (4.5)
Fruits	11.1 (3.0)	12.7 (3.9)	12.4 (3.2)	13.8 (3.1)	13.7 (3.2)	13.0 (3.8)
Vegetables	11.5 (3.4)	11.0 (3.7)	11.6 (4.2)	13.5 (3.6)	14.1 (3.7)	12.7 (3.3)
Total Fluency	38.3 (7.8)	41.7 (12.3)	39.3 (9.9)	43.9 (8.2)	44.6 (8.6)	42.4 (9.5)
	Men: Education (years)			Women: Education (years)		
Task	8–12 (n = 39) *M* (*SD*)	13–16 (n = 21) *M* (*SD*)	17+ (n = 13) *M* (*SD*)	8–12 (n = 66) *M* (*SD*)	13–16 (n = 73) *M* (*SD*)	17+ (n = 25) *M* (*SD*)
Animals	16.3 (5.4)	16.8 (5.2)	17.1 (7.7)	15.6 (4.2)	17.2 (3.8)	18.0 (4.5)
Fruits	12.2 (3.7)	12.8 (3.2)	11.5 (3.2)	12.6 (3.1)	14.1 (3.6)	14.5 (2.7)
Vegetables	11.8 (4.0)	10.6 (3.7)	10.9 (3.4)	13.1 (3.6)	13.8 (3.7)	13.8 (3.0)
Total Fluency	40.3 (10.9)	40.2 (10.6)	39.5 (10.9)	41.2 (8.6)	45.1 (9.1)	46.4 (6.5)

Effects of Gender and Language on Total Category Fluency Score

A 2 × 2 (Gender × Language) analysis of covariance (ANCOVA) was conducted with age and education entered in as covariates. As expected, the effect associated with the covariates was significant [$F(2,552) = 20.50$, $p < .001$]. After adjusting for the covariates on the model, there was a significant effect of gender [$F(1,552) = 23.66$, $p < .001$], but not language [$F(1,552) = 1.20$, $p > .27$]. Moreover, there was no Gender × Language interaction [$F(1,552) = 1.67$, $p > .19$]. In general, women generated more words (44.5 ± 9.2) than did men (40.0 ± 9.6).

Effects of Different Variables on Separate and Total Category Fluency Measures

To determine the effects of different variables on each of the different category fluency scores, a 2 × 2 (Gender × Language) repeated measures MANCOVA was conducted with the three different categories serving as the dependent measures and with age and education entered as covariates. Univariate analyses were only examined following a significant F at $p < .05$. The results indicated a significant effect of age on the number of exemplars generated for each of the three categories, with increasing age being associated with lower fluency scores in all categories. The covariate education was associated with the number of animals and fruits generated but not with the number of vegetables. Participants who were better educated generated more animals and fruits than their less educated counterparts. After adjusting for the aforementioned covariate effects, there was a statistically significant multivariate effect for language [Wilks $F(3,545) = 7.43$, $p < .001$], and gender [Wilks $F(3,545) = 23.81$, $p < .001$]. There was no statistically significant Gender × Language interaction [Wilks $F(3,545) = .87$, $p > .45$]. An examination of the specific univariate tests revealed a statistically significant effect of language for *vegetables* [$F(1,547) = 13.29$, $p < .001$] but not for *animals* or *fruits*. English-speaking participants generated more *vegetable* exemplars than Spanish-speaking participants (14.2 ± 3.8 *vs*. 12.8 ± 3.8). Univariate tests for gender revealed a gender effect for *vegetables* [$F(1,547) = 57.39$, $p < .001$] and *fruits* [$F(1,547) = 23.47$, $p < .001$], but not for *animals*. Women generated a greater number of *vegetable* exemplars than men (14.4 ± 3.7 *vs*. 11.6 ± 3.3) as well as more *fruit* exemplars (13.7 ± 3.5 *vs*. 12.0 ± 3.4).

Linear Regression Models

A series of stepwise linear regression analyses were conducted to more directly examine the relative contribution of age, education, gender, language, and depressive symptomatology on different category fluency measures. As compared to the analysis of variance model, regression analysis offers the advantage that participants with the lowest educational attainment (i.e., less than 8 years of education) or in the highest age range (80–89 years of age) do not have to be excluded because of prohibitively small cell sizes when conducting comparative analyses.

Table 6 presents the regression estimates, incremental increases in explained variance, and associated p values for each regression analysis. As shown in the table, there was a significant effect of age, education, and language on animal fluency, with younger, more educated, Spanish-speaking participants achieving higher scores. These variables explained 10.3% of the total variance in animal fluency scores. It should be noted that, although statistically significant, the effect of language accounted for less than 1% of the incremental increase in the explained variance associated with the model.

Table 6. Estimates of stepwise regression analysis for fluency measures

Task	Estimate *B*	*SE*	Beta	Cumulative *R*	Cumulative R^2	*p*
Animals						
Age	−0.115	.02	−.229	.262	6.9%	<.001
Education	0.269	.05	.189	.312	9.7%	<.001
Language	0.777	.37	.082	.321	10.3%	<.040
Vegetables						
Age	−0.102	.02	−.244	.319	10.2%	<.001
Gender	2.506	.29	.301	.361	13.1%	<.001
Language	−1.263	.29	−.162	.402	16.1%	<.001
HDRS	−0.105	.03	−.126	.420	17.6%	<.001
Fruits						
Age	−0.109	.01	−.273	.280	7.8%	<.001
Education	0.118	.04	.105	.336	11.3%	<.004
Gender	1.536	.28	.194	.355	12.6%	<.001
HDRS	−0.061	.03	−.077	.363	13.2%	<.040
Total Category Fluency						
Age	−0.322	.04	−.303	.306	9.4%	<.001
Education	0.399	.10	.133	.359	12.9%	<.001
Gender	4.181	.74	.198	.387	15.0%	<.001
HDRS	−0.226	.07	−.107	.401	16.1%	<.003

Note. HDRS = Hamilton Depression Rating Scale.

Gender and HDRS score failed to enter the model as predicting factors.

For vegetable fluency, age, gender, language, and HDRS scores were significant predictors of performance, explaining 17.6% of the total variance. Younger age, female gender, English language, and a lower HDRS score were associated with higher vegetable fluency scores. Educational attainment did not enter the model as a significant predictor variable.

Regarding fruit fluency, the variables age, education, gender, and HDRS score significantly predicted performance, explaining 13.2% of the total variance. Younger age, higher educational attainment, female gender, and a lower HDRS score were associated with a higher score on the fruit fluency task. Of note, although HDRS scores entered into the regression equation, this predictor accounted for less than 1% of the incremental increase in the explained variance in the model.

Age, education, gender, and HDRS scores were all related to total category fluency score, accounting for 16.1% of the variability. Younger age, higher educational attainment, female gender, and a lower HDRS score were associated with higher total category fluency scores. Again, although statistically significant, HDRS scores only accounted for an additional 1% of the explained variance in the model. Language did not enter the model as a predictor.

DISCUSSION

The present study provides normative data from a large cohort of cognitively normal elderly ES and SS for the set of category fluency measures most frequently used in neuropsychological studies in the United States. In addition to providing norms for total category fluency scores in ES and SS, the present normative set extends norms published to date by including wider age and education ranges and by comparing the relative importance of various participant characteristics on scores for each individual fluency measure. Normative data for the SS are expected to be especially helpful, given the sizable number of native Spanish-speaking elderly currently residing in the United States and the projected increase in their number during the next decades.

The results of the present study indicate that age is the strongest predictor of fluency scores for each individual measure (i.e., *animals*, *vegetables*, and *fruits*) and for the total number of exemplars retrieved from all three categories. In other words, increasing age was associated with lower fluency scores in all these measures. Education is also a strong predictor of performance, with increasing educational attainment being associated with higher total category fluency scores and with number of exemplars retrieved from the categories *animals* and *fruits*. Gender also predicts total category fluency scores as well as generation of exemplars belonging to the categories *vegetables* and *fruits*. In fact, gender was the best predictor of total fluency score for both ES and SS, with women outperforming men, after adjusting for age and education. Using the same categories as in the present study, Monsch et al. (1992) also reported higher total fluency scores in women compared to men in an English-speaking sample. On the other hand, Lucas et al. (1998) failed to find a gender effect, although they found that age and education correlated with total fluency scores in their English-speaking sample.

Language differentially predicted the scores on some fluency measures, with ES having a higher score on vegetable fluency and SS having slightly higher scores on animal fluency. As noted above, the HDRS scores entered the regression equation for total fluency scores and for scores on vegetable and fruit fluency, although its incremental explanatory power as a predictor was small (0.6–1.5%). The mechanism by which depressive symptomatology may differentially affect some category fluency measures but not others remains unclear. It has been suggested that the effect of depressed mood is more pronounced on more challenging cognitive tasks as compared to less challenging tasks (Caine, 1986). A number of studies evidence that regardless of primary language, more exemplars are generated for the category *animal* than for *fruit* or *vegetable*, suggesting that retrieval of exemplars from the former category is less difficult (see Ardila et al., 1994; Bayles et al., 1989; Bolla et al., 1998). Thus, the generation of vegetable and fruit exemplars most likely constitutes a more effortful task than generation of animal names, making the former fluency measures more susceptible to the effects of depressed mood.

In the present study, animal fluency was found to be associated with age and education, but not gender. Similar findings have been reported in both English- (Bolla et al., 1998, Crossley et al., 1997, Tombaugh et al., 1999) and Spanish-speaking samples (Ostrosky-Solis et al., 1999). Fluency for animals was also found to be affected by education and age in normal Hispanic, Chinese, and Vietnamese elderly who were tested in their native language (Kempler et al., 1998). The present study extends previous findings in ES and SS by showing that once the effects of age, education, and language have been taken into account, animal fluency is not affected by depressive symptomatology, as evidenced by scores on the HDRS.

Fluency score for the category fruit was associated with age, education, and gender, but not language. In other words, in both ES and SS, younger age, higher educational attainment, and female gender were associated with higher fruit fluency scores. Similar findings for fruit fluency in ES had been previously reported by Bolla et al. (1998), with the present study extending these findings to SS. In addition, the present study shows that depressive symptomatology may decrease scores, albeit minimally, on the fruit fluency measure.

The category in which language had a more pronounced effect was vegetables, with ES scoring higher than SS. In addition, gender and age predicted performance, with female participants outperforming male participants, and younger participants scoring higher than older subjects. Bolla et al. (1998) showed similar effects of gender and age in fluency for vegetables in ES. The present study extends their results by showing that depressive symptomatology may decrease performance on this task. More importantly, present results evidence that vegetable fluency scores are affected by language or one of its correlates. Although it is possible that inherent differences of the English *versus* the Spanish language are responsible for this difference, it is unlikely that such an effect would selectively affect some categories but not others. A more plausible explanation is that language is acting as a marker of other related factors, including sociocultural differences between ES and SS. Some of these differences may be related to nutritional preferences and to familiarity with a wider variety of vegetables in individuals from the United States as compared to Latin American countries. This may apply to elderly SS residing in the United States, as nutritional preferences and customs may be resilient to change, even after years of residence in a foreign country. Further support for a sociocultural explanation comes from the observation that, regardless of primary language, women outperformed men on vegetable and fruit fluency while evidencing similar scores on animal fluency. This may be related to the fact that, in the United States as well as in Latin American countries, women are more likely to be involved in food procurement and preparation than men.

A number of potential limitations can be identified in the present study. First, utilization of the same cut-off score of 27 on the MMSE for ES and SS may have created a selection bias potentially resulting in a relatively higher functioning Spanish-speaking than English-speaking group. In fact, Spanish-speaking participants in the present study scored slightly higher on the four-word delayed recall measure than their English-speaking counterparts (11.4 ± 0.7 *vs*. 11.2 ± 0.8). The extent to which different cut-off scores on the MMSE should be used to equate English- *versus* Spanish-speaking groups remains a controversial issue (see Bird et al., 1987; Escobar et al., 1986; Mungas et al., 1996).

The majority of SS participants in the present study resided in the United States and more than half of them were of Cuban origin. It is not known, however, whether the above mentioned findings generalize to SS who have different demographic characteristics or who reside in primarily Spanish-speaking countries. Moreover, information was not available on variables that may affect performance on neuropsychological tasks, including degree of acculturation and bilingualism, years of residence in the United States, years of school completed in the United States, and country of residence. In addition, it is not possible to ascertain the extent to which the present results generalize to younger cohorts, to verbal fluency paradigms that are based on other categories, and to languages other than English and Spanish. Finally, the present norms are based on the administration of the three fluency measures in the order of *animals–vegetables–fruits* as compared to the order *animals–fruits–vegetables* used in other studies. Whether the order of administration of these tasks affects performance remains an empirical question.

It was not possible to generate normative tables for Age $\times$ Education $\times$ Gender interactions for each of our language groups, as this would have resulted in extremely small cell sizes that would have precluded any interpretive significance. In clinical practice, the neuropsychologist could consult Table 6 to decide which factor should be given the most weight when determining which normative table to use for a given verbal fluency measure.

Although the use of all three categories is recommended, especially when assessing for the presence and type of dementia (see Monsch et al., 1992), animal fluency may be the measure of choice when only one category is to be sampled. More normative studies have been published on this category than on any other category fluency task, allowing the clinician to select an appropriate normative set based on relevant characteristics of the patient and on the procedure used in the administration and scoring of the task. In addition, although affected by age and education, animal fluency is not affected by gender or depressed mood, and is only minimally affected by the language of the examinee.

The present study suggests that different category fluency measures are not equivalent and that they are differentially affected by various variables related to the individual. Thus, in addition to providing information about total category fluency score, future normative studies should evaluate the impact of demographic variables on each category used and provide adjusted norms, if applicable. In addition, the present study provides evidence that some category fluency measures may be more culturally fair than others.

REFERENCES

Ardila, A. (1995). Directions of research in cross-cultural neuropsychology. *Journal of Clinical and Experimental Neuropsychology*, *17*, 143–150.

Ardila, A., Roselli, M., & Puente, A.E. (1994). *Neuropsychological evaluation of the Spanish speaker*. New York: Plenum Press.

Bayles, K.A., Salmon, D.P., Tomoeda, C.K., Jacobs, D., Caffrey, J.T., Kazniak, A.W., & Tröster, A.I. (1989). Semantic and letter category naming in Alzheimer's patients: A predictable difference. *Developmental Neuropsychology*, *5*, 335–347.

Binetti, G., Magni, E., Cappa, S.F., Padovani, A., Bianchetti, A., & Trabucchi, M. (1995). Semantic memory in Alzheimer's disease: An analysis of category fluency. *Journal of Clinical and Experimental Neuropsychology*, *17*, 82–89.

Bird, H.R., Canino, G., Stipec, M.R., & Shrout, P. (1987). Use of the Mini-Mental State Examination in a probability sample of a Hispanic population. *Journal of Nervous and Mental Diseases*, *175*, 731–737.

Bolla, K.I., Gray, S., Resnick, S.M., Galante, R., & Kawas, C. (1998). Category and letter fluency in highly educated older adults. *Clinical Neuropsychologist*, *12*, 330–338.

Caine, E.D. (1986). The neuropsychology of depression: The pseudodementia syndrome. In I. Grant & K.M. Adams (Eds.), *Neuropsychological assessment of neuropsychiatric disorders* (pp. 221–243). New York: Oxford University Press.

Carew, T.G., Lamar, M., Cloud, B.S., Grossman, M., & Libon, D.J. (1997). Impairment in category fluency in ischemic vascular dementia. *Neuropsychology*, *11*, 400–412.

Crossley, M., D'Arcy, C.D., & 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 of vascular type. *Journal of Clinical and Experimental Neuropsychology*, *19*, 52–62.

Day, J.C. (1996). *Population projection of the United States by age, sex, race, & Hispanic origin: 1995–2050* (U.S. Bureau of the Census, Current Population Reports, P25-1130). Washington, DC: U.S. Government Printing Office.

Echemendia, R.J., Harris, J.G., Congett, S.M., Diaz, M.L., & Puente, A.E. (1997). Neuropsychological training and practices with Hispanics: A national survey. *Clinical Neuropsychologist*, *11*, 229–243.

Escobar, J.I., Burnam, A., Karno, M., Forsythe, A., Landsverk, J., & Golding, J.M. (1986). Use of the Mini-Mental State Examination (MMSE) in a community population of mixed ethnicity: Cultural and linguistic artifacts. *Journal of Nervous and Mental Diseases*, *174*, 607–614.

Folstein, M.F., Folstein, S.E., & McHugh, P.R. (1975). Mini-Mental State: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*, 189–198.

Goodglass, H. & Kaplan, E. (1983). *The assessment of aphasia and related disorders* (2nd ed.). Philadelphia: Lea and Febiger.

Gurd, J.M. & Ward, C.D. (1989). Retrieval from semantic and letter-initial categories in patients with Parkinson's disease. *Neuropsychologia*, *27*, 743–746.

Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery, and Psychiatry*, *23*, 56–62.

Hodges, J.R., Salmon, D.P., & Butters, N. (1990). Differential impairment of semantic and episodic memory in Alzheimer's and Huntington's diseases: A controlled prospective study. *Journal of Neurology, Neurosurgery, and Psychiatry*, *53*, 1089–1095.

Huff, J.F., Corkin, S., & Growdon, J.H. (1986). Semantic impairment and anomia in Alzheimer's disease. *Brain and Language*, *28*, 235–249.

Jacobs, D.M., Sano, M., Albert, S., Schofield, P., Dooneief, G., & Stern, Y. (1997). Cross-cultural neuropsychological assessment: A comparison of randomly selected, demographically matched cohorts of English- and Spanish-Speaking Older Adults. *Journal of Clinical and Experimental Neuropsychology*, *19*, 331–339.

Johnson-Selfridge, M.T., Zalewski, C., & Aboudarham, J. (1998). The relationship between ethnicity and word fluency. *Archives of Clinical Neuropsychology*, *13*, 319–325.

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.

King, D.A. & Caine, E.D. (1990). Depression. In J. L. Cummings (Ed.), *Subcortical dementia* (pp. 218–230). New York: Oxford University Press.

Kukull, W A., Larson, E.B., Teri, L., Bowen, J., McCormick, W., & Pfanschmidt, M.L. (1994). The Mini-Mental State Examination score and the clinical diagnosis of dementia. *Journal of Clinical Epidemiology*, *47*, 1061–1067.

Loewenstein, D.A., Arguelles, T., Arguelles, S., & Linn-Fuentes, P. (1994). Potential cultural bias in the neuropsychological assessment of the older adult. *Journal of Clinical and Experimental Neuropsychology*, *16*, 623–629.

Loewenstein, D.A., Arguelles, T., Barker, W.W., & Duara, R.D. (1993). A comparative analysis of neuropsychological test performance of Spanish-speaking and English-speaking patients with Alzheimer's disease. *Journal of Gerontology: Psychological Sciences*, *48*, 142–149.

Loewenstein, D.A., Barker, W.W., Harwood, D.G., Luis, C., Acevedo, A., Rodriguez, I., & Duara, R. (2000). Utility of a modified Mini-Mental Examination with extended delayed recall in screening for mild cognitive impairment and dementia among community dwelling elders. *International Journal of Geriatric Psychiatry*, *15*, 434–440.

Lucas, J.A., Ivnik, R.J., Smith, G.E., Bohac, D.L., Tangalos, E.G., Graff-Radford, N.R., & Petersen, R.C. (1998). Mayo's older Americans normative studies: Category fluency norms. *Journal of Clinical and Experimental Neuropsychology*, *20*, 194–200.

Mahurin, R.K., Espino, D.V., & Holifield, E.B. (1992). Mental status testing in elderly Hispanic populations: Special concerns. *Psychopharmacological Bulletin*, *28*, 391–399.

Massman, P.J., Delis, D.C., Butters, N., Dupont, R.M., & Gillin, J.C. (1992). The subcortical dysfunction hypothesis of memory deficits in depression: Neuropsychological validation in a subgroup of patients. *Journal of Clinical and Experimental Neuropsychology*, *14*, 687–706.

Mattis, S. (1988). *Dementia Rating Scale: Professional manual*. Odessa, FL: Psychological Assessment Resources.

Mitrushina, M.N., Boone, K.B., & D'Elia, L.F. (1999). *Handbook of normative data for neuropsychological assessment*. New York: Oxford University Press.

Monsch, A.U., Bondi, M.W., Butters, N., Salmon, D.P., Katzman, R., & Thal, L.J. (1992). Comparisons of verbal fluency tasks in the detection of dementia of the Alzheimer type. *Archives of Neurology*, *49*, 1253–1258.

Monsch, A.U., Seifritz, E., Taylor, K.I., Ermini-Funfschilling D., Stahelin, H.B., & Spiegel, R. (1997). Category fluency is also predominantly affected in Swiss Alzheimer's disease patients. *Acta Neurologica Scandinavica*, *95*, 81–84.

Mungas, D., Marshall, S.C., Weldon, M., Haan, M., & Reed, B.R. (1996). Age and education correction of Mini-Mental State Examination for English- and Spanish-speaking elderly. *Neurology*, *46*, 700–706.

Ober, B.A., Dronkers, N.F., Koss, E., Delis, D.C., & Friedland, R.P. (1986). Retrieval from semantic memory in Alzheimer-type dementia. *Journal of Clinical and Experimental Neuropsychology*, *8*, 75–92.

Ostrosky-Solis, F., Ardila, A., & Roselli, M. (1999). NEUROPSI: A brief neuropsychological test battery in Spanish with norms by age and educational level. *Journal of the International Neuropsychological Society*, *5*, 413–433.

Pasquier, F., Lebert, F., Grymonprez, L., & Petit, H. (1995). Verbal fluency in dementia of frontal type and dementia of Alzheimer type. *Journal of Neurology, Neurosurgery, and Psychiatry*, *58*, 81–84.

Ponton, M.O., Satz, P., Herrera, L., Ortiz, F., Urrutia, C.P., Young, R., D'Elia, L.F., Furst, C.J., & Namerow, N. (1996). Normative data stratified by age and education for the Neuropsychological Screening Battery for Hispanics (NeSBHIS): Initial Report. *Journal of the International Neuropsychological Society*, *2*, 96–104.

Randolph, C., Braun, A.R., Goldberg, T.E., & Chase, T.N. (1993). Semantic fluency in Alzheimer's, Parkinson's, and Huntington's Disease: Dissociation of storage and retrieval failures. *Neuropsychology*, *7*, 82–88.

Reed, J. & Ramirez, R.R. (1998). *Current population reports: The Hispanic population in the United States: March 1997 (Update)* (U.S. Bureau of Census, No. P20-511). Washington, DC: U.S. Government Printing Office.

Rey, G.J. & Benton, A. (1991). *Examen de Afasia Multilingue* [Multilingual Aphasia Examination–Spanish]. Iowa City, IA: AJA Associates.

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

Stern, Y., Richards, M., Sano, M., & Mayeux, R. (1993). Comparison of cognitive changes in patients with Alzheimer's and Parkinson's disease. *Archives of Neurology*, *50*, 1040–1045.

Stricks, L., Pittman, J., Jacobs, D.M., Sano, M., & Stern, Y. (1998). Normative data for a brief neuropsychological battery administered to English- and Spanish-speaking community-dwelling elders. *Journal of the International Neuropsychological Society*, *4*, 311–318.

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. (1993). Differential effects of aging on two verbal fluency tasks. *Perceptual and Motor Skills*, *76*, 465–466.

Tröster, A.I., 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.

Unverzagt, F.W., Morgan, O.S., Thesiger, C.H., Eldemire, D.A., Luseko, J., Pokuri, S., Hui, S.L., Hall, K.S., & Hendrie, H.C. (1999). Clinical utility of CERAD neuropsychological battery in elderly Jamaicans. *Journal of the International Neuropsychological Society*, *5*, 255–259.

U.S. Bureau of the Census. (1997). *How we are changing: Demographic state of the nation: 1997* (Series P23-193). Washington, DC: U.S. Government Printing Office.

Welsh, K.A., Butters, N., Hughes, J.P., Mohs, R.C., & Heyman, A. (1992). Detection and staging of dementia in Alzheimer's disease (1992). *Archives of Neurology*, *49*, 448–452.