

---

# How does the homophone meaning generation test associate with the phonemic and semantic fluency tests? A quantitative and qualitative analysis

---

GITIT KAVÉ,<sup>1</sup> AYELET AVRAHAM,<sup>2</sup> DORIT KUKULANSKY-SEGAL,<sup>2,3</sup> AND ORLY HERZBERG<sup>2,3</sup>

<sup>1</sup>Department of Education and Psychology, The Open University, Ra'anana, Israel

<sup>2</sup>Department of Pediatric Rehabilitation, Safra Children's Hospital, Sheba Medical Center, Ramat Gan, Israel

<sup>3</sup>Department of Communication Disorders, Tel Aviv University, Tel Aviv, Israel

(RECEIVED July 24, 2006; FINAL REVISION November 19, 2006; ACCEPTED November 20, 2006)

## Abstract

The current study examines whether the Homophone Meaning Generation Test (HMGT; Warrington, 2000) is correlated with the phonemic and/or the semantic fluency tests and compares its association with the number of switches and clusters and the mean cluster size of these two tasks. One hundred healthy Hebrew speakers (18–35 years of age; mean = 24.9) generated meanings for 24 homophones and provided words beginning in three different letters, as well as words belonging to three semantic categories. Results show that the HMGT score correlated significantly and similarly with the total score of both the phonemic and the semantic fluency tests. There was a significant correlation between the HMGT and the number of phonemic switches and clusters, but not between the HMGT and the mean phonemic cluster size. The HMGT correlated with the number of semantic switches and clusters, as well as with the mean semantic cluster size. The findings suggest that the relationship between these tests is mediated by a shared executive component, attesting to the HMGT's utility in tapping into mechanisms of shifting and mental flexibility. (*JINS*, 2007, *13*, 424–432.)

**Keywords:** Verbal fluency, Word generation, Switching and clustering, Executive functions, Language testing, Hebrew

## INTRODUCTION

In an attempt to devise a new test that would help in diagnosing executive dysfunction, Warrington (2000) has suggested to look at verbal switching in a Homophone Meaning Generation Task (HMGT). On this task, participants are required to generate multiple meanings for each of a series of eight homophones: *tick*, *tip*, *slip*, *form*, *plain*, *bored*, *right*, and *sent* (Crawford & Warrington, 2002). It is assumed that performance on this task measures the ability to switch between alternative verbal concepts, and patients with anterior brain lesions have been found to be more impaired on the HMGT than patients with posterior lesions (Warrington, 2000).

According to Warrington (2000), unlike the more traditional Wisconsin Card Sorting Test, in which there are only three categories to switch between, the HMGT requires multiple switching. Normal English speakers provide one to six distinct meanings for each homophone and can make up to five switches per target (see Table 1 in Crawford & Warrington, 2002). While different authors disagree as to whether homophones have one or many phonological representations (Caramazza et al., 2001, 2004; Jescheniak et al., 2003; Jescheniak & Levelt, 1994; Miozzo et al., 2004), it is agreed that every homophone must have multiple conceptual representations. Thus, when participants are asked to provide as many different meanings as possible for an auditory target presented out of context, they are encouraged to switch from the most frequent meaning of that stimulus to other concepts associated with it. This process requires mental control and flexibility, considered to be executive functions.

---

Correspondence and reprint requests to: Gitit Kavé, Department of Education and Psychology, The Open University, The Dorothy de Rothschild Campus, 108 Ravutski Street, P.O. Box 808, Ra'anana 43104, Israel. Email: gkave@012.net.il

The HMGT has been devised to test executive functions through verbal fluency. However, being a new test, it is much less commonly used in neuropsychological and language assessment than the more familiar phonemic/letter or semantic/category fluency tests. On these tests, individuals are required to generate as many different words that begin with a certain letter or as many different category exemplars within a limited time, and they have been extensively studied in various languages (Benito-Cuadrado et al., 2002; Chan & Poon, 1999; Gladsjo et al., 1999; Kavé, 2005; Kempler et al., 1998) and in various neurological and healthy populations (Barr & Brandt, 1996; Bokas & Goldberg, 2003; Kozora & Cullum, 1995; Tombaugh et al., 1999). We believe that the utility and validity of the HMGT will be significantly improved once its association with phonemic and semantic fluency tests is more clearly understood. Hence, the main purpose of the current study is to examine the performance of healthy participants on the HMGT, the phonemic fluency test, and the semantic fluency test and to further elucidate the cognitive mechanisms underlying the HMGT.

Word retrieval on all verbal fluency tests depends on lexical knowledge as well as on effective search processes that require set shifting (Troyer et al., 1997). It has been suggested that semantic fluency may be more impaired in individuals with temporal brain damage, whereas phonemic fluency may be more impaired in individuals with frontal brain damage (Rosser & Hodges, 1994; Troyer et al., 1998). This suggestion reflects the assumption that retrieval of words by semantic categories requires greater reliance on lexical stores than retrieval of words by letters and that retrieval of words by letters requires greater reliance on executive functions than on lexical stores. However, this distinction is not without its problems.

Indeed, it has been repeatedly shown that individuals with temporal brain lesions are more impaired on the semantic fluency test than on the phonemic fluency test. A meta-analysis of studies of persons with Alzheimer's disease, who have severe lexical-conceptual disorders, demonstrated greater difficulties in semantic fluency relative to phonemic fluency (Henry et al., 2004). Focal temporal damage has also been associated with a lesser deficit on phonemic fluency and a larger deficit on semantic fluency (Henry & Crawford, 2004). Nevertheless, while persons with temporal lesions show a semantic fluency deficit, individuals with frontal lesions are often impaired on the semantic fluency test as well (Rogers et al., 2006).

It has also been shown that persons with frontal lobe lesions produce significantly fewer words on phonemic fluency tests than do healthy controls, and perform worse on that test than do persons with nonfrontal lesions (Alvarez & Emory, 2006). Thus, argue Alvarez and Emory (2006), an intact frontal cortex, especially on the left side, is required for successful performance on the phonemic fluency task. However, other brain areas, including subcortical circuitry, also subserve the phonemic fluency task, and phonemic fluency deficits are often found in individuals with

nonfrontal lesions (due, for example, to reduced articulation speed). Most importantly, persons with focal frontal lesions are reported to have a comparable impairment on both phonemic and semantic fluency tests (Henry & Crawford, 2004).

As the HMGT has been found to be more impaired in individuals with anterior brain damage than in individuals with posterior damage (Warrington, 2000), it is possible that it would be more highly correlated with the phonemic fluency task than with the semantic fluency task. Alternatively, since the HMGT requires handling of a large vocabulary, it could be more highly correlated with the semantic fluency test, which is known to be particularly affected by lexical damage. However, it is highly likely that there will be a strong association between the HMGT and both fluency tasks, since persons with frontal lesions have been found to be equally impaired on both fluency tests (Henry & Crawford, 2004), presumably because both tasks require reliance on intact executive functioning. An association between the HMGT and the semantic fluency test can result either from the involvement of lexical knowledge or from reliance on an executive component in both tests. Yet, an investigation of total fluency scores cannot easily differentiate among these alternatives.

To better understand the mechanisms involved in word retrieval on verbal fluency tests, some authors have used qualitative rather than quantitative methods to examine the cognitive strategies underlying these tasks (Koren et al., 2005; Kosmidis et al., 2004; Sauzéon et al., 2004; Troster et al., 1998; Troyer et al., 1997, 1998; Tucha et al., 2005; Woods et al., 2004). Instead of comparing the total number of words provided on the phonemic and semantic tasks, these authors have looked at two components termed *switching* and *clustering*. According to Troyer et al. (1998), when generating words on the phonemic and semantic fluency tasks, participants produce clusters of phonemically or semantically related words and, once a subcategory is exhausted, they switch to another subcategory. Thus, performance on these tasks relies on (1) an executive component (i.e., switching) responsible for strategic search, response initiation, monitoring, shifting, and flexibility; and (2) an associate component (i.e., clustering) that reflects the semantic organization of memory stores (Troyer et al., 1997, 1998; Troyer, 2000).

It is assumed that anterior brain regions play a more important role in switching than in clustering. Troyer et al. (1998) examined this hypothesis in persons with focal brain lesions, finding individuals with frontal lobe lesions to switch less frequently than healthy participants and to produce normal cluster size on both the phonemic and the semantic tasks. In contrast, individuals with temporal lobe lesions exhibited normal switching and clustering on the phonemic task, but were impaired in switching on the semantic task. Although persons with temporal lobe lesions showed no marked deficit in cluster size, those who had left temporal lesions produced smaller clusters than those who had right temporal lesions. This study suggested that phonemic clustering was

less dependent on the integrity of lexical stores than was semantic clustering and that the most discriminating index among the patient groups was the number of switches on the phonemic fluency task, which was impaired only in persons with frontal lesions.

Further evidence supporting the assumption that switching is an executive function, whereas clustering is more dependent on lexical abilities, especially within the semantic task, comes from research of various nonfocal neuropsychological disorders. For example, Woods et al. (2004) found that persons with dementia due to human immunodeficiency virus, which affects subcortical–frontal pathways, produce fewer switches than do healthy participants, but their clusters are of a similar size as those produced by the control group. The same was true also for adults with attention deficit hyperactivity disorder (Tucha et al., 2005) and for individuals with multiple sclerosis (Troster et al., 1998). On the other hand, individuals with Alzheimer's disease produced smaller clusters than normal (Epker et al., 1999; Troster et al., 1998). In addition, persons with schizophrenia, who suffer from disproportionate semantic fluency impairment relative to phonemic fluency (Bokat & Goldberg, 2003; Kremen et al., 2003), have been found to also have a disproportionate decrease in the number of clustered words (Bozikas et al., 2005).

As these studies show, the examination of switching and clustering has helped clarify the relative contribution of executive strategies and semantic stores to the performance on verbal fluency tests. It is thus suggested that the investigation of the association between the HMGT and the more familiar phonemic and semantic fluency tests should focus not only on the total number of words generated in each task, but also on the analysis of switching and clustering. Accordingly, the aims of the current study are twofold: (1) to examine whether performance on the HMGT is differentially or equally correlated with total performance on the phonemic and semantic fluency tests; and (2) to determine whether the HMGT is more highly correlated with the switching component than with the clustering component on either fluency test.

## METHODS

### Participants

The sample consisted of 100 volunteers, 54 of them women, 18–35 years of age (mean age = 24.9;  $SD = 5.1$ ). Their level of education ranged between 12 and 19 years, with a mean of 13.8 ( $SD = 1.7$ ). All participants were native Hebrew speakers, recruited through places of employment, university classes, and word of mouth. Persons with a known history of learning disorders, psychiatric disturbances, neurological disease, or head trauma were not included in the study. Participant recruitment was conducted in accordance with institutional research guidelines.

### Materials

#### Homophone Meaning Generation Test

Because the test was conducted in Hebrew, different target stimuli from the eight homophones used by Warrington (2000) had to be selected. To increase the number of possible switches, 24 homophones were chosen, each having at least 3 possible meanings (with a range of 3–10 possible meanings). Half of the targets were also homographs (i.e., all their meanings are spelled the same way). Each homograph was matched to a nonhomograph that had the same number of possible meanings (see Table 1).

#### Procedure

The two types of homophones were pseudorandomly mixed in one list, and the same list was administered to all participants. Each participant was tested individually, and all responses were written verbatim. There was no time limitation, and participants indicated to the examiner when they could think of no more meanings. The HMGT was administered following the two fluency tasks.

#### Scoring

Every distinct meaning was given one point according to a predetermined list of possible meanings, and the total test score consisted of the number of all distinct meanings generated for the 24 homophones. Repetitions (e.g., *et* 'pen': fountain pen, blue pen) and irrelevant meanings (*et*: 'at' in English) were excluded from the total score. Two independent raters first coded the responses generated by 30 participants. The correlations between their coding were  $r = .99$  for distinct meanings,  $r = .88$  for repeated meanings, and  $r = .93$  for irrelevant meanings. All correlations were significant at the .01 level. The rest of the sample was coded by either one of these two raters.

#### Phonemic and Semantic Fluency Tests

##### Procedure

Participants were asked to provide as many words as possible within 60 s on each of three letters (phonemic test) and three categories (semantic test). The phonemic fluency test was administered first and then the semantic fluency test and the order of letters, as well as the order of semantic categories, was constant across participants. Responses were written verbatim, with errors or repetitions subsequently excluded from the total score. When a questionable response was provided, clarifications were invited at the end of the 1-min interval.

*Phonemic fluency:* This was assessed by obtaining the number of words generated in 1 min for the letters *bet* (/b/), *gimel* (/g/), and *shin* (/š/). Instructions were as follows: "I want you to say as many Hebrew words as possible that

**Table 1.** Number of possible meanings per target on the Hebrew HMGT and mean number of distinct meanings generated per target across participants

	Homographs			Nonhomographs		
	Targets	Possible meanings	Mean no. of distinct meanings generated	Targets	Possible meanings	Mean no. of distinct meanings generated
1	keren	9	3.52 ( $\pm 0.98$ )	kara	10	3.61 ( $\pm 0.98$ )
2	parax	5	2.38 ( $\pm 0.72$ )	et	5	3.25 ( $\pm 0.76$ )
3	gamal	4	2.25 ( $\pm 0.67$ )	maxa	4	2.30 ( $\pm 0.73$ )
4	safa	3	2.71 ( $\pm 0.52$ )	atsar	3	1.92 ( $\pm 0.60$ )
5	ritsa	4	2.48 ( $\pm 0.63$ )	eres	5	2.70 ( $\pm 0.81$ )
6	shok	4	2.34 ( $\pm 0.59$ )	mana	5	2.91 ( $\pm 0.60$ )
7	tsipa	4	2.56 ( $\pm 0.54$ )	kina	4	2.82 ( $\pm 0.64$ )
8	rav	5	2.92 ( $\pm 0.39$ )	mila	5	2.80 ( $\pm 0.71$ )
9	ruax	10	2.56 ( $\pm 0.98$ )	kala	10	3.86 ( $\pm 1.07$ )
10	saraf	9	2.57 ( $\pm 0.77$ )	ala	9	2.25 ( $\pm 0.73$ )
11	tsir	8	2.70 ( $\pm 0.94$ )	amad	8	2.24 ( $\pm 0.61$ )
12	ma'ala	10	2.86 ( $\pm 0.90$ )	tava	8	2.96 ( $\pm 0.79$ )
Sum		75	3185		76	3362
Mean		6.25	2.65		6.33	2.80
SD		2.7	0.34		2.5	0.58
Range		3–10	2.25–3.52		3–10	1.92–3.86

Note. HMGT = Homophone Meaning Generation Test.

begin with a certain letter. You may say any word except for names of people and places, such as *Tomer* or *Tel Aviv*. Also, you should use different words rather than the same word with a different ending. For example, if you say *tapuz* ('orange'), don't also say *tapuzim* ('oranges'). If you say a verb, use the simplest form *halax* ('he went') and not *halaxti* ('I went') or *holex* ('he goes'). Please don't say words that are attached to other words, such as *mi-shamayim* ('from the sky') or *la-kise* ('to the chair')."

**Semantic fluency:** This test was assessed by obtaining the number of words generated in 1 min for each of the following three semantic categories: animals, fruits and vegetables, and vehicles. Fruits and vegetables were treated as one category to avoid the ambiguity between botanical definitions and common usage (as in 'avocado'). It was specified that for the category of vehicles only types of transportation should be provided while brand names were unacceptable.

### Scoring

When homophones were provided, the second mention was counted only if the participant pointed out the alternate meaning explicitly (i.e., *gamal* 'camel', 'repaid'). Words inflected in both masculine and feminine forms (e.g., *gever-gveret* 'mister-mistress'; *sus-susa* 'horse-mare') were counted as one, whereas an animal and its offspring were counted as separate words (e.g., *para* 'cow' and *egel* 'calf'). Synonyms were counted as two (*matos* and *aviron* 'airplane'). Names of subcategories on the semantic test (e.g.,

bird) were not given credit if specific items within that subcategory (e.g., dove, eagle) were also provided. Slang terms (e.g., *shluk* 'sip'), as well as foreign words (e.g., bandana, gangster), were generally acceptable.

**Clustering and switching:** In line with the guidelines of Troyer et al. (1997), repetitions and mistakes were included in the scoring of clustering and switching. An item that appeared in two clusters was coded in both. For example, the "cat" in "dog, cat, tiger, lion" was counted both as part of the cluster of pets and as part of the cluster of felines. In cases in which a small cluster was embedded within a larger cluster, only the larger cluster was counted. Thus, if farm birds were generated among other farm animals, only one cluster was counted (i.e., horse, cow, chicken, duck, turkey = one cluster). Semantic clusters generated within the phonemic task, as well as phonemic clusters generated within the semantic task, were not scored.

**Phonemic clustering:** A cluster was counted when two consecutive words shared the first consonant and vowel (*gezer-geshem*), shared the first and second consonant but differed in the vowel of the opening syllable (*gina-ganav*), rhymed (*shamayim-shinayim*), or included duplication (*barbur-bilbul*).

**Semantic clustering:** Where possible, subcategories were based on previous studies (Kosmidis et al., 2004; Troyer et al., 1997). Guidelines were formulated for consistency's sake, but flexibility was allowed for the coding of associated words that did not fall under the list of predefined

clusters. In the *Animal* category, clusters were coded according to habitat, zoological family, and family relation, which were further classified into relevant subcategories. In the *Fruit and Vegetable* category, clusters were coded according to either fruits or vegetables, with subcategories further defined by season, botanical family, manner of eating, and so on. In the *Vehicle* category, clusters were coded according to land, water, or air-borne means of transportations, with further classifications within land vehicles defined by common use.

Four variables were derived for each fluency test on the basis of the aforementioned criteria:

1. *Total fluency score*: All words excluding repetitions and errors, summed across the three letters for the phonemic task and across the three categories for the semantic task.

2. *Mean cluster size*: Following Troyer et al. (1997), the number of words in a cluster was counted from the second word. That is, a cluster of two words was coded as 1; a cluster of three words was coded as 2, and so forth. A mean of all clusters of two words or more was computed for every person for each letter or semantic category. These means were then averaged across the three letters to yield the mean phonemic cluster size of each participant, and across the three semantic categories to yield the mean semantic cluster size of each participant.

3. *Number of switches*: The number of switches between clusters of two words or more, between a cluster and a single word generated outside a cluster, and among those out-of-cluster single words (as in Troyer et al., 1997) was counted for every person for each letter and for each semantic category. Switches produced by each participant were summed across the three letters to yield the total phonemic number of switches score, and across the three semantic categories to yield the total semantic number of switches score.

4. *Number of clusters*: The number of clusters was counted separately, without single words, to examine participants' use of word association. As noted by Koren et al. (2005), the presence of single words may indicate that participants are in fact unable to use an associative strategy; thus, a measure that leaves out the single words is essential when focusing on the tendency to produce related words.

To calculate inter-rater reliability, responses generated by 30 participants on both fluency tasks were first coded for clustering and switching by two independent raters and the correlations between their scoring were derived. On the phonemic task, correlations between the two raters were  $r = .79$  for mean cluster size,  $r = .99$  for number of switches, and  $r = .93$  for number of clusters. On the semantic task, correlations between the two raters were  $r = .97$  for mean cluster size,  $r = .96$  for number of switches, and  $r = .96$  for number of clusters. All correlations were significant at the .01 level. The rest of the sample was coded by either one of these two raters.

## RESULTS

Table 1 presents the mean number of distinct meanings provided for each homophone on the HMGT. To arrive at this measure, repeated meanings and irrelevant meanings were first deleted from the data set. These responses accounted for 4.5% and 4.3% of all meanings, respectively, leaving 91.2% of the generated responses for further analyses. Distinct meanings were divided into three separate measures: total HMGT score, homographs, and nonhomographs. A paired-sample *t* test showed that participants provided significantly more meanings for the nonhomographs relative to the homographs:  $t(99) = 4.272, p < .05$  (see Tables 1 and 2). Table 2 presents means, standard deviations, and range of scores on the HMGT across stimuli, as well as on the phonemic and semantic fluency variables.

Because education level can reflect vocabulary size, its association with the HMGT was examined. As this association was found to be significant ( $r = .389, p < .05$ ), edu-

**Table 2.** Raw scores on the HMGT and on the phonemic and semantic fluency variables

Test	Variable	Mean	SD	Range
HMGT	Total distinct meanings	65.47	7.66	45–81
	Homographs: distinct meanings	31.85	4.19	21–41
	Nonhomographs: distinct meanings	33.62	4.51	22–45
	Total repeated & irrelevant meanings	6.34	5.70	0–36
Phonemic fluency	Total three letters	43.19	9.39	12–68
	Mean cluster size	1.45	0.32	1.0–3.2
	Sum of switches	27.65	6.50	6–42
	Sum of clusters	11.39	4.03	3–26
Semantic fluency	Total three letters	60.72	10.60	30–88
	Mean cluster size	1.75	0.31	1.2–2.8
	Sum of switches	31.69	5.86	14–48
	Sum of clusters	17.56	3.70	9–25

Note. HMGT = Homophone Meaning Generation Test.

cation was controlled for when analyzing the correlations between the HMGT and the fluency variables. Table 3 presents partial correlations between the HMGT and the fluency variables, net of education. There was a significant correlation between the total score on the HMGT and the total score on the phonemic fluency test ( $r = .433$ ), as well as between the total score on the HMGT and the total score on the semantic fluency test ( $r = .410$ ). The difference between these two correlation coefficients was not statistically significant:  $z = .274$ ,  $p > .05$ .

An examination of the correlations between the HMGT and the switching and clustering variables reveals a different picture for each fluency task. On the phonemic fluency task, there was no correlation between the mean cluster size and the HMGT, while the correlations between the number of switches and the HMGT ( $r = .434$ ), as well as between the number of clusters and the HMGT ( $r = .401$ ), were statistically significant. On the semantic fluency task, mean cluster size correlated significantly with the HMGT ( $r = .207$ ), and so did the number of switches ( $r = .297$ ) and the number of clusters ( $r = .288$ ). Although the former correlation was smaller than the latter two, the coefficients did not differ significantly (mean cluster size and HMGT *vs.* number of switches and HMGT:  $z = .605$ ,  $p > .05$ ; mean cluster size and HMGT *vs.* number of clusters and HMGT:  $z = .584$ ,  $p > .05$ ).

It is important to note that on both fluency tasks the correlations between the total number of switches or the total number of clusters and the number of words produced for the homographs were slightly higher than the equivalent correlations with the number of words produced for the nonhomographs (bottom two rows in Table 3). However, the differences between the relevant coefficients pertaining to single- and multiple-spelling homophones were too small to reach statistical significance.

## DISCUSSION

A correlation analysis revealed that the total score on the HMGT was significantly and equally associated with the total score on both the phonemic and the semantic fluency

tests in a group of healthy Hebrew speakers. The HMGT was originally constructed as a test of executive functions that requires directed search and flexibility, and persons with anterior brain damage performed more poorly on this test relative both to healthy controls and to persons with posterior brain damage (Warrington, 2000). It was thus plausible to hypothesize that the HMGT would correlate more highly with the phonemic fluency task, a task that requires more mental flexibility because it cannot rely on a search within existing conceptual categories. However, because the HMGT requires manipulation of various conceptual representations, it was also reasonable to expect that it would be highly correlated with the semantic fluency task, as this task is heavily affected by lexical knowledge (Rogers et al., 2006), much more so than the phonemic fluency task.

Alternatively, because it has been suggested that persons with focal frontal lesions are similarly impaired on both the phonemic and the semantic fluency tests (Henry & Crawford, 2004), comparable correlations between the HMGT and the two word fluency tests were likely to arise. This finding could indicate that both fluency tests involve a process that relies on an intact frontal lobe, and the association between all three tests could be a product of a shared executive component necessary for successful performance on these tasks. Based only on the total fluency scores, though, it is impossible to rule out the possibility that the HMGT is related to the phonemic fluency test through shared executive mechanisms, whereas its association with the semantic fluency test reflects a shared dependence on lexical-conceptual stores.

Results of the qualitative analysis of the verbal fluency measures speak directly to this issue. The present findings suggest that performance on the HMGT is more highly related to the number of switches or the number of clusters than to the mean cluster size across the two fluency tasks. In fact, the mean cluster size on the phonemic task was not correlated with the HMGT at all. The switching component is considered an executive function because it involves strategic search, shifting, and mental flexibility, whereas the clustering component is assumed to rely on semantic stores (Troyer et al., 1997, 1998; Troyer, 2000). Thus, it appears that the HMGT is associated with the phonemic task through

**Table 3.** Partial correlations between the HMGT and the fluency variables (controlling for education)

	Phonemic fluency				Semantic fluency			
	Total three letters	Mean cluster size	Sum of switches	Sum of clusters	Total three categories	Mean cluster size	Sum of switches	Sum of clusters
Total	.433**	-.091	.434**	.401**	.410**	.207*	.297**	.288**
Homographs	.454**	-.062	.445**	.412**	.423**	.236*	.301**	.281**
Nonhomographs	.309**	-.094	.318**	.293**	.298**	.131	.221*	.223*

Note. HMGT = Homophone Meaning Generation Test.

\*Significant at the .05 level.

\*\*Significant at the .01 level.

a shared executive component rather than through a measure of vocabulary. This was not the case on the semantic fluency task, in which the correlation between the HMGT and the mean cluster size was statistically significant, and although smaller than the correlations with the number of switches or the number of clusters, not significantly so. Importantly, however, the HMGT was associated not only with the clustering component on the semantic task, but also with the switching component.

Despite the fact that the correlation between the HMGT and the switching variable is likely to be the result of a shared executive component, some criticism of this variable has been raised in the literature. Specifically, Abwender et al. (2001) have pointed out that since switches in the Troyer et al. (1997) analysis include not only shifting from one cluster to another but also shifting among single words, switches do not reflect an executive process but simply a failure to cluster. However, in the current study, we looked also at the number of clusters and demonstrated essentially the same picture for both the number of switches and the number of clusters. The correlations between either the number of switches or the number of clusters and the HMGT further suggest that switching from one cluster to another requires mental flexibility and does not indicate a mere failure to cluster.

Why is the mean phonemic cluster size not correlated with the HMGT? Cluster size is assumed to reflect an associate component that relies on the semantic organization of memory stores (Troyer et al., 1997, 1998; Troyer, 2000). By definition, semantic clustering reflects the organization of knowledge into conceptual categories, whereas phonemic clustering depends on similarities of sound. Indeed, Hughes and Bryan (2002) found no association between phonemic cluster size and independent vocabulary scores, suggesting that the phonemic cluster size was not particularly affected by semantic funds. In addition, studies that attempted to predict the total word output on the phonemic task through fluency variables found that the number of switches was more important than the mean cluster size in this prediction (Kosmidis et al., 2004; Troyer et al., 1997).

Unlike the mean cluster size on the phonemic task, mean semantic cluster size was significantly correlated with the total HMGT score in the current study, even when controlling for education level, which was used as a proxy for vocabulary size. This finding is in line with previous studies that reported significant correlations between the total semantic fluency output and the mean semantic cluster size (Fossati et al., 2003; Kosmidis et al., 2004; Troyer et al., 1997). However, if mean cluster size is related to the HMGT because both measures assess lexical knowledge, why isn't the correlation between them higher? While members of a semantic cluster necessarily belong to a similar conceptual field, this is not the case with regard to homophone representations, especially in Hebrew. Note that many English homophones may have semantically related meanings that differ only in terms of their part of speech (a form = a shape, to form = to shape), whereas in Hebrew this is usu-

ally not the case (e.g., *kala* = bride, wove, hit, light in weight, etc.). Thus, performance on the HMGT cannot be attributed to spreading of activation through subcategories within the semantic lexicon.

Another finding of the current study is that participants generate more distinct meanings for nonhomographs than for homographs. This finding contrasts with Warrington's (2000) results, possibly reflecting the difference in total number of homophone targets used in both studies, or the different number of possible meanings of each target stimulus, which was larger in the current study. It is also possible that this discrepancy is a product of the differences between English and Hebrew, whether related to the number of semantic fields to which homophone meanings belong or to the differences between the two orthographic systems. When responses were examined by the nature of the homophone target (single vs. multiple spellings), it appeared that the homographs were more highly related to the switching component than were the nonhomographs, although the comparison of correlation coefficients did not reach statistical significance. This trend might represent the greater difficulty involved in switching flexibly among the various meanings of a homograph relative to a nonhomograph. It could be the case that multiple orthographic representations facilitate flexible search within the semantic system. Obviously, this observation deserves further investigation, in other languages as well as with clinical populations.

In conclusion, the present findings suggest that all three tests tap into mechanisms of shifting and mental flexibility within the mental lexicon, even if to a slightly different degree. The attested correlations between the HMGT and both fluency tests among healthy participants, its specific association with the switching component, and its sensitivity to frontal brain damage (Warrington, 2000) make the HMGT a good test for use in the evaluation of executive deficits. In addition, the fact that the HMGT has no time constraints allows for an examination of executive functioning that does not depend on speed of processing. This advantage could be especially useful in assessing disorders such as multiple sclerosis or depression, in which the need to tease apart these two cognitive factors is often a challenge (Henry & Beatty, 2006; Henry & Crawford, 2005). Performing the analyses conducted here on data from individuals with executive dysfunction is required to corroborate the current findings. Nonetheless, although the HMGT still has to be more thoroughly investigated, especially with patient populations, it holds much promise as a simple quantitative test of mental flexibility that could be easily incorporated into clinical neuropsychological practice.

## ACKNOWLEDGMENTS

The authors thank the participants of this study for volunteering their time. Many thanks are also extended to Merav Liran, Inbar Morag, Hadas Nagar-Turgeman, Shani Waidergorean, and Rivka Ziv for their help in data collection. There are no financial or other relationships that could be interpreted as a conflict of interests.

## REFERENCES

- Abwender, D.A., Swan, J.G., Bowerman, J.T., & Connolly, S.W. (2001). Qualitative analysis of verbal fluency output: Review and comparison of several scoring methods. *Assessment, 8*, 323–336.
- Alvarez, J.A. & Emory, E. (2006). Executive function and the frontal lobes: A meta-analytic review. *Neuropsychology Review, 16*, 17–42.
- Barr, A. & Brandt, J. (1996). Word-list generation deficits in dementia. *Journal of Clinical and Experimental Neuropsychology, 18*, 810–822.
- Benito-Cuadrado, M.M., Esteba-Castillo, S., Bohm, P., Cejudo-Bolivar, J., & Pena-Casanova, J. (2002). Semantic verbal fluency of animals: A normative and predictive study in a Spanish population. *Journal of Clinical and Experimental Neuropsychology, 24*, 1117–1122.
- Bokat, C.E. & Goldberg, T.E. (2003). Letter and category fluency in schizophrenic patients: A meta-analysis. *Schizophrenia Research, 64*, 73–78.
- Bozikas, V.P., Kosmidis, M.H., & Karavatos, A. (2005). Disproportionate impairment in semantic verbal fluency in schizophrenia: Differential deficit in clustering. *Schizophrenia Research, 74*, 51–59.
- Caramazza, A., Bi, Y., Costa, A., & Miozzo, M. (2004). What determines the speed of lexical access: Homophone or specific-word frequency? A reply to Jescheniak et al. (2003). *Journal of Experimental Psychology: Learning Memory and Cognition, 30*, 278–282.
- Caramazza, A., Costa, A., Miozzo, M., & Bi, Y. (2001). The specific-word frequency effect: Implications for the representation of homophones in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 1430–1450.
- Chan, A.S. & Poon, M.W. (1999). Performance of 7- to 95-year-old individuals in a Chinese version of the Category Fluency Test. *Journal of the International Neuropsychological Society, 5*, 525–533.
- Crawford, J.R. & Warrington, E.K. (2002). The Homophone Meaning Generation Test: Psychometric properties and a method for estimating premorbid performance. *Journal of the International Neuropsychological Society, 8*, 547–554.
- Epker, M.O., Lacritz, L.H., & Munro Cullum, C. (1999). Comparative analysis of qualitative verbal fluency performance in normal elderly and demented populations. *Journal of Clinical and Experimental Neuropsychology, 21*, 425–434.
- Fossati, P., Guillaume, le B., Ergis, A.M., & Allilaire, J.F. (2003). Qualitative analysis of verbal fluency in depression. *Psychiatry Research, 117*, 17–24.
- Gladsjo, J.A., Schuman, C.C., Evans, J.D., Peavy, G.M., Miller, S.W., & Heaton, R.K. (1999). Norms for letter and category fluency: Demographic corrections for age, education, and ethnicity. *Assessment, 6*, 147–178.
- Henry, J.D. & Beatty, W.W. (2006). Verbal fluency deficits in multiple sclerosis. *Neuropsychologia, 44*, 1166–1174.
- Henry, J.D. & Crawford, J.R. (2004). A meta-analytic review of verbal fluency performance following focal cortical lesions. *Neuropsychology, 18*, 284–295.
- Henry, J.D. & Crawford, J.R. (2005). A meta-analytic review of verbal fluency deficits in depression. *Journal of Clinical and Experimental Neuropsychology, 27*, 78–101.
- Henry, J.D., Crawford, J.R., & Phillips, L.H. (2004). Verbal fluency performance in dementia of the Alzheimer's type: A meta-analysis. *Neuropsychologia, 42*, 1212–1222.
- Hughes, D.L. & Bryan, J. (2002). Adult age differences in strategy use during verbal fluency performance. *Journal of Clinical and Experimental Neuropsychology, 24*, 642–654.
- Jescheniak, J.D. & Levelt, W.J. (1994). Word frequency effects in speech production: Retrieval of syntactic information and phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 824–843.
- Jescheniak, J.D., Meyer, A.S., & Levelt, W.J. (2003). Specific-word frequency is not all that counts in speech production: Comments on Caramazza, Costa, et al. (2001) and new experimental data. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 432–438.
- Kavé, G. (2005). Phonemic fluency, semantic fluency, and difference scores: Normative data for adult Hebrew speakers. *Journal of Clinical and Experimental Neuropsychology, 27*, 690–699.
- 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.
- Koren, R., Kofman, O., & Berger, A. (2005). Analysis of word clustering in verbal fluency of school-aged children. *Archives of Clinical Neuropsychology, 20*, 1087–1104.
- Kosmidis, M.H., Vlahou, C.H., Panagiotaki, P., & Kiosseoglou, G. (2004). The verbal fluency task in the Greek population: Normative data, and clustering and switching strategies. *Journal of the International Neuropsychological Society, 10*, 164–172.
- Kozora, E. & Cullum, C.M. (1995). Generative naming in normal aging: Total output and qualitative changes using phonemic and semantic constraints. *The Clinical Neuropsychologist, 9*, 313–320.
- Kremen, W.S., Seidman, L.J., Faraone, S.V., & Tsuang, M.T. (2003). Is there disproportionate impairment in semantic or phonemic fluency in schizophrenia? *Journal of the International Neuropsychological Society, 9*, 79–88.
- Miozzo, M., Jacobs, M.L., & Singer, N.J.W. (2004). The representation of homophones: Evidence from anomia. *Cognitive Neuropsychology, 21*, 840–866.
- Rogers, T.T., Ivanou, A., Patterson, K., & Hodges, J.R. (2006). Semantic memory in Alzheimer's disease and the frontotemporal dementias: A longitudinal study of 236 patients. *Neuropsychology, 20*, 319–335.
- 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.
- Sauzéon, H., Lestage, P., Raboutet, C., N'Kaoua, B., & Claverie, B. (2004). Verbal fluency output in children aged 7–16 as a function of the production criterion: Qualitative analysis of clustering, switching processes, and semantic network exploitation. *Brain and Language, 89*, 192–202.
- 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.
- Troster, 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.



- Troyer, A.K. (2000). Normative data for clustering and switching on verbal fluency tasks. *Journal of Clinical and Experimental Neuropsychology*, *22*, 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.P., & Stuss, D.T. (1998). Clustering and switching on verbal fluency: The effects of focal frontal- and temporal-lobe lesions. *Neuropsychologia*, *36*, 499–504.
- Tucha, O., Mecklinger, L., Laufkotter, R., Kaunzinger, I., Paul, G.M., Klein, H.E., & Lange, K.W. (2005). Clustering and switching on verbal and figural fluency functions in adults with attention deficit hyperactivity disorder. *Cognitive Neuropsychiatry*, *10*, 231–248.
- Warrington, E.K. (2000). Homophone meaning generation: A new test of verbal switching for the detection of frontal lobe dysfunction. *Journal of the International Neuropsychological Society*, *6*, 643–648.
- Woods, S.P., Conover, E., Rippeth, J.D., Carey, C.L., Gonzalez, R., Marcotte, T.D., Heaton, R.K., Grant, I., & HIV Neurobehavioral Research Center (HNRC) Group. (2004). Qualitative aspects of verbal fluency in HIV-associated dementia: A deficit in rule-guided lexical-semantic search processes? *Neuropsychologia*, *42*, 801–809.