

Verbal Naming Test for Use with Older Adults: Development and Initial Validation

Brian P. Yochim,^{1,2,3} Sherry A. Beaudreau,^{1,2,4} J. Kaci Fairchild,^{1,2} Maya V. Yutsis,^{1,2} Neda Raymond,¹ Leah Friedman,^{1,2}
AND Jerome Yesavage^{1,2}

¹VA Palo Alto Health Care System, Palo Alto, California

²Stanford University School of Medicine, Stanford, California

³National Jewish Health, Denver, Colorado

⁴University of Queensland, Australia

(RECEIVED September 12, 2014; FINAL REVISION February 2, 2015; ACCEPTED February 2, 2015; FIRST PUBLISHED ONLINE March 24, 2015)

Abstract

Naming or word-finding tasks are a mainstay of the typical neuropsychological evaluation, particularly with older adults. However, many older adults have significant visual impairment and there are currently no such word-finding tasks developed for use with older visually impaired populations. This study presents a verbal, non-visual measure of word-finding for use in the evaluation of older adults with possible dysnomia. Stimuli were chosen based on their frequency of usage in everyday spoken language. A 60-item scale was created and given to 131 older Veterans. Rasch analyses were conducted and differential item functioning assessed to eliminate poorly-performing items. The final 55-item scale had a coefficient alpha of 0.84 and correlated with the Neuropsychological Assessment Battery Naming test, $r = 0.84$, $p < .01$, Delis-Kaplan Executive Function System (D-KEFS) Category Fluency, $r = 0.45$, $p < .01$, and the D-KEFS Letter Fluency, $r = 0.40$, $p < .01$. ROC analyses found the measure to have sensitivity of 79% and specificity of 85% for detecting dysnomia. Patients with dysnomia performed worse on the measure than patients with intact word-finding, $t(84) = 8.2$, $p < .001$. Patients with no cognitive impairment performed significantly better than patients with mild cognitive impairment, who performed significantly better than patients with dementia. This new measure shows promise in the neuropsychological evaluation of word-finding ability in older adults with or without visual impairment. Future directions include the development of a shorter version and the generation of additional normative data. (*JINS*, 2015, 21, 239–248)

Keywords: Alzheimer disease, Anomia, Cognition disorders, Dementia, Language, Neuropsychology, ROC curve, Sensitivity and specificity

A VERBAL NAMING TEST FOR USE WITH OLDER ADULTS

Development and Initial Validation

Neuropsychologists who work with older adults typically assess word-finding, or naming, when evaluating for the possible presence of Alzheimer's disease (AD) or primary progressive aphasia, or when evaluating patients after a history of stroke or other lesions in the language-dominant hemisphere. The prevalence of dementia in adults age 71 and older has been estimated to be 14%, with rates increasing to 24% for individuals age 80–89 and 37% for those individuals age 90 and

older (Plassman et al., 2007). Both the National Institute on Aging (NIA)/Alzheimer's Association diagnostic guidelines (McKhann et al., 2011) and Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5; American Psychiatric Association, 2013) include impaired word-finding or naming in their diagnostic guidelines for dementia or neurocognitive disorder. Other types of dementia (e.g., primary progressive aphasia) are also characterized by a profound deficit in word-finding (Gorno-Tempini et al., 2011; Lezak, Howieson, Bigler, & Tranel, 2012). Damage to the left hemisphere, through strokes or other forms of acquired brain injury and temporal lobe epilepsy can lead to aphasia. Nearly all patients with aphasia have a word-finding deficit (Beeson & Rapcsak, 2006; Blumenfeld, 2010; Goodglass & Wingfield, 1997; Laine & Martin, 2006) that can be long-lasting even as aphasia symptoms improve (Goodglass & Wingfield, 1997). Evaluation of word-finding is thus paramount in clinical

Correspondence and reprint requests to: Brian Yochim, National Jewish Health, 1400 Jackson Street, B134, Denver, CO 80206. E-mail: yochimb@njhealth.org

neuropsychological practice for accurate diagnosis of neurocognitive disorders and proper evaluation of cognitive deficits following damage to the language-dominant hemisphere.

Neuropsychologists assess for word-finding difficulty with tests such as the naming subtest of the Neuropsychological Assessment Battery (NAB; Stern & White, 2003; White & Stern, 2003), the Boston Naming Test (BNT; Kaplan, Goodglass, & Weintraub, 1978, 2001), and the naming subtest of the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS; Randolph, 1998), in which an examiner presents photographs (NAB) or line drawings (BNT, RBANS) to patients (e.g., of a lion, a compass, etc.) and asks the patient to name them. Older patients often have visual impairment, however, which can interfere with the validity of these measures. Vision loss is among the top three causes of physical disability in older adults (Lee & Coleman, 2004), and the prevalence rate for vision loss ranges from 10% for adults aged 55–64 to over 20% after age 75 (Ryskulova et al., 2008). In addition, over 700,000 adults are completely blind (Ryskulova et al., 2008). Thus, a key cognitive ability of word-finding may not be adequately assessed for over 20% of older patients. The prevalence rate of blindness and low vision will increase in the coming years due to the growing population of older adults (Eye Diseases Prevalence Research Group, 2004). Additional cortical problems such as visual field cuts can pose challenges to valid test administration. Niemeier (2010) noted how visual impairment in the context of neuropsychological testing can lead to misdiagnosis, and in such situations neuropsychologists should use tests that involve alternative sensory modalities (e.g., audition) whenever possible. Thus, it would be advantageous to have a word-finding test that can be used with patients with visual impairment.

A non-visual naming test would also be an effective way to assess word-finding in patients who have intact vision. Evidence suggests that a non-visual, auditory/verbal measure of word-finding may detect mild cognitive impairment and/or dementia more effectively than picture-naming tests (Hodges and Patterson, 1995, Miller, Finney, Meador, & Loring, 2010). This may be due to aspects of the functional neuroanatomy of word-finding; word-finding and semantic processing is a complex ability often mediated by several neuroanatomical correlates including lateral temporal, temporoparietal, and anterior temporal regions (Chiang et al., 2014; Grossman et al., 2004). The anterior temporal lobes may serve as storehouses of general semantic knowledge (Brier, Maguire, Tillman, Hart, & Kraut, 2008), or knowledge of more specific concepts such as social conceptual processing (Simmons & Martin, 2009). Although some studies have linked picture-naming performance to anterior temporal lobe volume (Balthazar et al., 2010), other studies (Hamberger, Goodman, Perrine, & Tamny, 2001; Malow et al., 1996) suggest that picture naming is more closely associated with temporo-parietal brain regions, while anterior temporal regions may be more related to word-finding in response to verbal definitional prompts. As such, visual naming tests may not adequately assess aspects of word-finding in patients with

primarily anterior temporal dysfunction (e.g., subtypes of frontotemporal lobar degeneration, temporal lobe epilepsy, and focal acquired brain injuries).

In addition, verbal word-finding tasks may be a more ecologically-valid way to assess this ability. The word-finding problems associated with AD tend to occur in everyday conversational speech (Nebes, 1989). Indeed, dysnomia has been defined in the context of speech (Goodglass & Wingfield, 1997). As a result, the diagnostic criteria for AD focus on difficulty while speaking (McKhann et al., 2011), as opposed to difficulty identifying pictured objects. Existing word-finding measures, however, do not assess word-finding during actual speech but in the context of looking at a picture and naming it, potentially limiting their ecological validity. Furthermore, self-reported difficulty finding words while speaking has also correlated with word-finding in response to a verbal definition but not with naming in response to a picture (Hamberger, Seidel, McKhann, & Goodman, 2010). Picture-naming tests such as the BNT may tap into “inability to recognize... common objects” or “object agnosia” as described in the McKhann et al. (2011) criteria as examples of impaired visuospatial abilities (p. 265), which is distinct from word-finding difficulty. Therefore, a verbally-based, non-visual word-finding test may enable clinicians to more accurately detect difficulty in finding words while speaking.

Hamberger and Seidel (2003) created an auditory naming test in their studies on temporal lobe epilepsy. Their measure was developed in an epilepsy clinic and used with a sample of adults with temporal lobe epilepsy. Their normative data are available for 100 healthy adults with a mean age of 34.3 ($SD = 11.0$) and for 56 additional patients of similar age with epilepsy. The Hamberger and Seidel measure represents a valuable addition to the neuropsychological tool kit with a promising methodology. However, the normative sample does not include individuals older than 64, rendering it difficult for clinicians to determine whether an older patient's performance is normal or impaired. Also, some items may be objectionable to older patients (e.g., “what an old man uses to walk with”) or have become outdated (e.g., “the white stuff used to write on a blackboard”). In addition, the stimuli on this measure consist of words commonly used in everyday spoken language, with frequencies of usage significantly higher than the average frequency of usage of words on the BNT or NAB Naming subtest (Yochim, Rashid, Raymond, & Beaudreau, 2013). Other researchers have suggested that future confrontation naming tests should use predominantly low-frequency stimuli (Randolph, Lansing, Ivnik, Cullum, & Hermann, 1999) to render the scale more sensitive to emerging difficulty retrieving words quickly (Goodglass, Kaplan, & Barresi, 2001; Goodglass & Wingfield, 1997) and to word-finding errors (Kirshner, Webb, & Kelly, 1984; Skelton-Robinson & Jones, 1984). The aim of this study was to develop a similar non-visual word-finding measure that improved upon the contribution of Hamberger and Seidel (2003) by developing it in an older adult population and by using items of low frequency to increase its sensitivity.

Development of our word-finding measure, the Verbal Naming Test (VNT), followed several critical guidelines to improve upon other related measures. First, a unique aspect of our non-visual word-finding measure was the use of current word frequency ratings of *spoken* language in its development (Brysbaert & New, 2009), rather than relying upon frequency ratings for *written* language based on 1961 American literature (Francis & Kučera, 1982; Kučera & Francis, 1967; for more details about difference in these two word frequency systems see Yochim et al., 2013). Second, unlike most other existing naming tests (i.e., with the exception of the Action Naming Test; Obler & Albert, 1979), this one includes verbs as stimuli to be named in addition to nouns. Although naming of verbs has not been shown to change with age as much as the naming of nouns (Nicholas, Barth, Obler, Au, & Albert, 1997), aphasia has been shown to affect verb-naming and noun-naming differently (Kohn, Lorch, & Pearson, 1989; Miceli, Silveri, Nocentini, & Caramazza, 1988). Noun and verb naming also differ with regard to more temporal lobe activity for naming nouns and more frontal cortex involvement for naming verbs (Damasio & Damasio, 1992; Damasio & Tranel, 1993). The inclusion of 10 verbs on the proposed measure will allow some exploration of neuroanatomical differences in verb and noun naming, and the implications for detecting and tracking progression of AD.

Neuropsychological measures are available for assessing cognitive domains other than word-finding (e.g., attention, memory) without the need for intact vision. The absence of a non-visual measure of word-finding is, therefore, unique in the field of clinical neuropsychology. Such a measure would enable the evaluation of word-finding ability in patients with visual impairment, as well as over the telephone for patients who are unable to travel to a clinic, or in the emerging field of tele-neuropsychology (Cullum, Weiner, Gehrman, & Hynan, 2006; Grosch, Gottlieb, & Cullum, 2011). The aim of this project was to fill this void by creating a verbal (non-visual) measure of word-finding and to report on its preliminary psychometric properties.

METHODS

Participants

This study was conducted under approval from the Institutional Review Board in compliance with the Helsinki Declaration.

The sample included 92 patients in a neuropsychological assessment clinic, for whom diagnoses and other neuropsychological test data were available. Another 39 participants took the measure as part of an unrelated research study without receiving diagnoses or taking other neuropsychological measures used in this study, and these 39 participants were included in the sample used for Rasch analyses only. Self-reported ethnic background was 74.8% European American, 12.2% African American, 6.1% Hispanic American, 5.3% Asian American, and 1.6% Other. There were 123 males (94%) and 8 females (6%). Mean age, education, and test performance are provided in Table 1.

Measures

Patients in the neuropsychology clinic completed the following measures, and raw scores from these were used in correlation analyses:

NAB Naming Test

The Neuropsychological Assessment Battery (NAB) contains two forms of a measure of word-finding, described above, in which patients are shown a photograph and asked to name the pictured object. Form 1 was used in this study. The reliability of scores from this measure has been demonstrated in a nationally representative sample closely matched to U.S. Census data, and the validity of scores from this measure has been demonstrated in healthy participants of all ages as well as in a sample of patients with aphasia (mean age of 59 years, 93% male, and 82% European American; White & Stern, 2003). The validity of scores from the measure was also demonstrated in a sample of healthy adults with a mean age of 75.4 (Yochim, Kane, & Mueller, 2009). Patients who scored more than one SD below the published normative sample on this measure were defined as having dysnomia for analyses described below.

California Verbal Learning Test-Second Edition

The California Verbal Learning Test, second edition (CVLT-II; Delis, Kramer, Kaplan, & Ober, 2000) is a commonly-used measure of verbal memory (Rabin, Barr, & Burton, 2005) with extensive normative data. The test involves having an examinee learn a list of 16 words over five trials, and then recalling

Table 1. Participant characteristics ($N = 131$)

Measure	Mean (<i>SD</i>)	Range
Age	74.7 (11.5)	21–98
Education	14.5 (3.0)	7–21
Verbal Naming Test: Entire sample	47.1 (8.5)	0–55
Verbal Naming Test: Clinic sample without impairment ($n = 23$)	51.6 (3.6)	42–55
Verbal Naming Test: Clinic sample with MCI ($n = 30$)	48.1 (4.9)	34–54
Verbal Naming Test: Clinic sample with dementia ($n = 36$)	43.7 (5.6)	34–53

Note. Three outliers were removed from the clinic sample after Rasch analyses were completed with the entire sample.

the words 20 min later. Scores from this measure have high internal consistency based on split-half reliability, $r = .94$, in a national sample of 1,087 adults age 16–89 closely matched to U.S. Census data, and scores have demonstrated validity as a measure of verbal memory in a sample of 62 healthy adults with a mean age of 36.8 years, and in a variety of studies cited in the manual (Delis et al., 2000). The measure produces numerous scores, two of which were used in this study: the total words learned in Trials 1–5, and the number of words recalled after a long delay (Long Delay Free Recall).

D-KEFS Verbal Fluency

This subtest, used as a measure of language output and of executive functioning, is part of the Delis-Kaplan Executive Function System (D-KEFS; Delis, Kaplan, & Kramer, 2001). Participants are asked to verbally generate in 1 min as many words as possible that start with the same letter or belong to a certain semantic category. This task is completed three times with three different letters (F, A, and S) and two times with different categories (e.g., animals, boys' names). Internal consistency reliability for scores on this measure has been shown to range from .85 to .87 for letter fluency and .64 to .76 for category fluency among adults age 60–89 (Delis et al., 2001). Test-retest reliability has been shown to be .88 for letter fluency scores and .82 for category fluency among healthy adults age 50–89 (Delis et al., 2001).

D-KEFS Trail Making Test

The D-KEFS Trail Making Test involves a series of 5 conditions: visual scanning, number sequencing, letter sequencing, number-letter switching, and motor speed. This study incorporated the fourth condition, Number-Letter Switching, as a measure of complex attention and executive functioning. On the fourth condition, Number-Letter Switching, the participant draws a line connecting dots, alternating between dots containing numbers and dots containing letters. Scores from this measure have sufficient test-retest reliability in a sample of 36 healthy adults age 50–89 (Delis et al., 2001) and have been used extensively as a measure of executive functioning in studies on patients with frontotemporal dementia (e.g., with a mean age of 59.6, Huey et al., 2009) and patients with frontal lesions with a mean age of 65.5 (Yochim, Baldo, Nelson, & Delis, 2007). For this study, the total completion time was used as a measure of set-switching, with lower scores indicating better performance.

Judgment of Line Orientation

The Judgment of Line Orientation test (Benton, Sivan, Hamsher, Varney, & Spreen, 1994; Benton, Varney, & Hamsher, 1978) is a measure of visuospatial functioning. Participants are shown two lines and asked to pick which two lines, from a stimulus array below, match the spatial orientation of the initial two lines. Scores from the measure had high test-retest reliability ($r = 0.90$) in a sample of 37 patients (Benton et al., 1994), strong convergent validity

with other visuospatial measures among patients with cerebrovascular lesions with a mean age of 66 (Trahan, 1998), and a strong relationship with right parietal functioning among patients with focal brain damage and a mean age of 51 (Tranel, Vianna, Manzel, Damasio, & Grabowski, 2009).

Selection of Items for the Verbal Naming Test (VNT)

We enlisted Brysbaert and New's (2009) listing of words at various frequencies of usage to develop the Verbal Naming Test. A listing of words was generated with frequencies less than 6/1,000,000, which characterizes items 30–60 of the BNT (Yochim et al., 2013). This listing of words was then carefully scanned for possible items, in consultation with the authors on this study and other neuropsychologists. Special care was taken in the choice of stimuli to ensure the following: (1) Items should be spoken rarely enough to be difficult to generate, increasing the measure's sensitivity, while at the same time (2) being relatively free of relationships with education. (3) Items should be as culture-free and non-objectionable as possible (e.g., a noose is not included as a stimulus). (4) Items should be able to be defined with as short a definition as possible (e.g., "a baby cow"), to ensure comprehension difficulties do not interfere substantially. Pilot testing has shown that even patients with comprehension impairment generate phonemic paraphasic approximations of the target item, indicating they comprehended the stimulus prompt. (5) Items should have a definition prompt that should only elicit one word (i.e., avoiding stimulus words that have synonyms) to increase the standardization of administration and scoring. For example, the item "what you push a baby in to go for a walk" was eliminated because it unexpectedly generated several different responses (stroller, perambulator, carriage, buggy). (6) Items should not be limited to concrete nouns, but also include abstract nouns (e.g., "decade") and verbs. (7) Items should be avoided if the clearest definition included the target word, or a variation of the target word, itself (e.g., "the thing you use to clean your teeth" for toothbrush). Using these criteria, 60 items were generated, ranging in frequency from 0.55 to 5.29 per million, with two items at the beginning of the test with very high frequency (7.49 and 10.51), to ensure patients experience some success at the beginning of the task. Items were chosen such that there are approximately equivalent numbers of items at five frequency levels of 0.0–0.9, 1.0–1.9, 2.0–2.9, 3.0–3.9, and 4.0–5.3. Sixty items were selected to be similar to the widely used Boston Naming Test. In addition, 60 was thought to be an adequate number of items because elimination of problematic items would still leave enough items for the measure to be reliable and valid. Items were ordered in decreasing frequency, such that the measure becomes more difficult with each item. There are no discontinuous rules for administration at this time.

On the test, after a definitional prompt is given, patients have 10 s to independently generate the word, or an approximation

of the word such that the examiner can identify it as the target item. This ensures that regional variations in pronunciation of the word are not scored as incorrect. After 10 s, a phonemic cue is provided (e.g., “it starts with ‘um-’”). If the patient cannot generate the word within 10 s, the next item is administered. Clinicians should differentiate between articulation errors and phonemic or semantic paraphasic errors. The total score on the measure is the number of items correctly answered without phonemic cues. A delay of 10 s has been used on other established naming tests (i.e., the NAB Naming test, Stern & White, 2003). We also selected 10 s over longer delays provided by other naming tasks (e.g., 20 s on the BNT) because difficulty generating names within only 2 s has been shown to be a sensitive indicator of neurological disorders such as AD and temporal lobe epilepsy (Hamberger & Seidel, 2003; Miller et al., 2010) and healthy controls and patients are able to generate names within 7–8 s on naming tasks (Bell, Seidenberg, Hermann, & Douville, 2003). Moreover, the diagnostic guidelines for dementia due to AD include “hesitations” in finding words as a symptom of AD, and taking longer than 10 s to find a word would indicate the presence of such hesitations. A 10-s time limit also prevents unnecessarily tiring or discouraging of patients.

Statistical Analyses

Rasch analyses (Bond & Fox, 2007) were completed with the Winsteps software program. Items that showed evidence of bias were sought through Rasch analyses of differential item functioning between European American and non-European American participants. We dichotomized ethnicity in this way because there were not enough African American, Asian American, and Hispanic American participants to perform Rasch analyses specific to each ethnic group. Only eight females participated in this study, thus formal evaluation of bias related to sex was not possible. After these items were removed, other items showing statistically significant poor infit and outfit characteristics in Rasch analyses (i.e., less than 0.5 or greater than 1.5) were removed.

After Rasch analyses were completed, only the clinic patients were used for the following analyses. Correlations with demographic variables and with measures considered highly related (NAB Naming), semi-related (CVLT-II and D-KEFS Verbal Fluency), and unrelated (Judgment of Line Orientation and D-KEFS Trail Making) were calculated. *T* tests were calculated to determine whether patients with dysnomia performed significantly worse on the measure than patients without dysnomia. (Patients can have dysnomia but not dementia, and patients can have dementia without dysnomia. In our sample, 61% of the patients with dysnomia received diagnoses of dementia, 50% of the patients with dementia had dysnomia, and 18% of the sample had both dementia and dysnomia.) Logistic regression was conducted to determine the measure’s effectiveness in detecting dysnomia. ROC analyses were conducted to arrive at an optimal cut score to use for diagnosing patients as with or without dysnomia. Sensitivity, specificity, positive predictive

value, and negative predictive value were calculated. *T* tests were also calculated to determine if patients without cognitive diagnoses performed better than those with MCI, and if patients with MCI performed better than patients with dementia.

RESULTS

Two items that prompted multiple answers because of multiple synonyms were removed. In Rasch analyses, point-measure correlations ranged from 0.35 to 0.59. Only one item showed significant differential item functioning (Dif contrast = 1.71, $p < .05$) between European Americans and non-European Americans, and it was removed. Two other items showed poor infit (Mean square for both items = 0.33) and outfit (Mean square for both items = 0.01) that was statistically significant ($zstd = -2.0$) and they were removed, arriving at a scale with 55 items (see Appendix). Males (mean = 48.2; $SD = 5.4$) obtained similar scores as the eight females (mean = 48.1; $SD = 6.5$).

All subsequent analyses included only the clinic patients. Participants whose total scores on the naming task were more than two *SDs* below the mean ($n = 3$) were removed for all the following analyses, leading to a total of 89 participants for the following analyses. In the clinic sample, 23 patients were cognitively normal, 30 received diagnoses of MCI, and 36 received diagnoses of dementia due to various causes (11 due to AD, 8 due to AD and vascular disease, 5 due to vascular disease, 3 due to Parkinson’s disease, 4 due to multiple causes, and 5 unspecified).

The 55-item scale had a Cronbach’s alpha of 0.84. Correlations with other measures used in neuropsychological evaluations are presented in Table 2. As can be seen, the measure correlated highly with the NAB Naming test, measures of verbal fluency, and the CVLT-II.

Patients with dysnomia ($n = 28$; 32% of the clinic sample), defined by NAB Naming performance more than one *SD* below the mean, performed significantly worse on the Verbal Naming Test (mean = 41.9; $SD = 5.3$) than did those

Table 2. Correlations between Verbal Naming Test and other variables ($N = 89$)

Variable	Correlation
Age	-.08
Years of education	.31**
NAB Naming (form 1)	.84**
D-KEFS Category Fluency	.45**
D-KEFS Letter Fluency	.40**
CVLT-II Trials 1–5	.38**
CVLT-II Long Delay Free Recall	.30**
Judgment of Line Orientation	.23*
D-KEFS Trail Making, Number-Letter Switching	-.23*

* $p < .05$.

** $p < .01$.

without dysnomia (mean = 50.0; $SD = 3.6$); $t(84) = 8.2$; $p < .001$. Logistic regression, using VNT scores as a predictor of dysnomia, found a C index of 0.89. The VNT was a significant predictor of dysnomia, $Wald(1) = 22.8$, $p < .001$, with an overall correct classification rate of 83.7%. An ROC analysis was also conducted to determine how well the Verbal Naming Test detected dysnomia. This analysis found the area under the curve (AUC) to be 0.89. With a cut score of 47.5, sensitivity was 86%, specificity was 74%, positive predictive value was 62%, and negative predictive value was 92%. A cut score of 46.5 resulted in sensitivity of 79%, specificity of 85%, positive predictive value of 71%, and negative predictive value of 89%. We suggest the use of 46.5 as a cut score for diagnosing word-finding impairment, because it results in the least spread between sensitivity and specificity and positive and negative predictive value. This value is also approximately 1 SD below the mean VNT score of clinic patients with normal word-finding.

Patients with no cognitive impairment ($n = 23$) performed significantly better (mean = 51.6; $SD = 3.6$) than patients diagnosed with MCI ($n = 30$, mean = 48.1; $SD = 4.9$), $t(51) = 2.9$; $p < .01$. Patients diagnosed with MCI performed significantly better than patients diagnosed with dementia ($n = 36$; mean = 43.7; $SD = 5.6$); $t(64) = 3.4$; $p < .01$.

DISCUSSION

This study presents psychometrics for the newly developed Verbal Naming Test that improves upon existing word-finding tests in the following ways: (1) stimuli were chosen based on how frequently words are used in everyday spoken language; (2) stimuli were chosen only if rarely used in everyday conversation, increasing the sensitivity of the measure to mild word-finding deficits; (3) the measure can be used in patients with visual impairment; and (4) the measure can be administered in tele-neuropsychological evaluations. Tele-neuropsychology is a growing area in clinical neuropsychology (Cullum et al., 2006; Grosch et al., 2011), and this measure can be used in such evaluations, without concerns over how best to show visual stimuli to patients over the teleconference medium.

Preliminary data suggest that this measure has strong psychometric characteristics, with high internal consistency and validity. Convergent validity was demonstrated by its large correlation with another word-finding measure and medium-large correlations with related measures of verbal fluency and verbal memory, while discriminant validity was confirmed by weaker correlations with less-related constructs (i.e., measures of visuoception and set-switching). The measure also showed strong sensitivity, specificity, positive predictive value, and negative predictive value in detecting dysnomia with a cut score of 46.5 of 55 (i.e., with scores of 46 or below indicating impairment). There were significant differences in scores between patients with no impairment and with MCI, and between patients with MCI and with dementia, indicating the measure can be useful in detecting these conditions in older adults.

To our knowledge, this is the first word-finding test to incorporate frequency of usage in everyday spoken language in the choice of stimuli. This methodology should make this measure more externally valid, and the choice of rarely used words should increase its sensitivity to early word-finding deficits. As patients develop word-finding difficulty, these problems first occur when thinking of rarely used words, as opposed to commonly-used words (Kirshner et al., 1984; Skelton-Robinson & Jones, 1984).

Efforts were made to create a measure that is free from associations with education. Unfortunately, with the removal of three outliers from the data, the correlation with education became significant and medium in size (0.31), indicating that education accounts for 9.6% of the variance in performance. In contrast, the NAB Naming Test did not correlate significantly ($r = .19$; $p = .09$) with education in this sample, although the NAB Naming Test has correlated ($r = 0.32$; $p < .05$) with education in a non-clinical community sample (Yochim et al., 2009). The BNT has consistently been found to correlate with education (for review, see Strauss, Sherman, & Spreen, 2006), accounting for 10–13% of the variance in BNT performance. It is difficult, if not impossible, to create a word-finding test that is completely independent of education, as patients can only have familiarity with the stimuli through some degree of education. Future investigations with this measure will determine if specific items are highly impacted by education, and these items can be removed to arrive at a shorter form. Future normative data can also be stratified by education.

Another limitation is that, while the participants were from fairly diverse ethnic backgrounds, the study included only eight women (five of whom were clinic patients), because it was conducted in a Veterans' health care system. In this sample, males and females on average obtained similar Verbal Naming Test scores, although prior work (Randolph et al., 1999; Welch, Doineau, Johnson, & King, 1996; Zec, Burkett, Markwell, & Larsen, 2007) has found males and females to perform differently on the BNT. Future research with larger numbers of men and women varying in age, education, and ethnicity will determine if there is differential item functioning based on these demographic variables and the measure's generalizability to patients with diverse ethnic, educational, and socioeconomic backgrounds. The measure would also benefit from the development of versions in other languages. Like other naming measures, it should only be used cautiously, if at all, for patients who speak English as a second language. Clinicians should use their best judgment in determining whether a patient has been fluent enough in English, for a long enough period of time, for this test to be a valid measure of a patient's word-finding ability. Research on the use of this measure with younger populations would also help to ensure it works well with younger adults, although most common causes of impaired word-finding (e.g., Alzheimer's disease, strokes in the language-dominant hemisphere) occur primarily with older adults. Furthermore, the development of an alternate form in future studies will increase the usefulness of the measure for tracking change in word-finding ability over time while limiting practice effects.

Like all neuropsychological measures, sensory (e.g., hearing) impairment must be accounted for in the interpretation of performance. The use of phonemic cues assists clinicians in this regard. If a patient correctly produces a target word after a phonemic cue, it can be assumed that the patient correctly heard the item prompt. Pilot testing has shown that even patients with comprehension impairment generate phonemic paraphasic approximations of the target item, or will often say “I know what it is but I can’t think of it”, indicating they comprehended the stimulus prompt. If pure word deafness or auditory agnosia are suspected, it is recommended that clinicians supplement this measure with a picture-naming test. Fortunately, both of these conditions are known to be rare (Bauer, 2012), and likely to be detected and observed across measures administered as part of a larger battery (e.g., verbal memory testing). If a clinician observed impaired performance on the VNT in the context of completely preserved spontaneous speech, then syndromes such as these two would be in the diagnostic differential.

Rasch analyses were conducted to evaluate infit and outfit of items, and to assess for the presence of differential item functioning between European American and non-European American participants. Only two items showed significantly poor infit and outfit, and one item showed differential item functioning by ethnicity. While our study benefitted from 26% of our sample identifying as ethnic minorities, future research should continue to assess whether particular items have different difficulty levels for various ethnic groups, particularly for the rapidly growing Hispanic American population.

This study presents a new measure to add to the toolkit for clinical neuropsychologists, especially those who perform evaluations for older adults. Because approximately 20% of older adults have some degree of visual impairment (Ryskulova et al., 2008), neuropsychologists need assessment tools that do not require strong visual perception. This measure may also prove to be useful for assessing word-finding in all patients, regardless of their visual acuity. Other cognitive domains (e.g., memory) can be assessed in patients with visual impairment, and this measure will enable the assessment of another critical cognitive domain (i.e., word-finding) in such patients. Other advantages include the lack of a need to use a stimulus booklet to administer the measure, and the measure’s brief administration time (mean of 9.4 min, $SD = 2.7$ in our sample), which is shorter than the typical BNT administration time (10–20 min; Strauss et al., 2006). The measure will also facilitate the increased use and development of tele-neuropsychological evaluations, which will help increase the provision of services to patients in rural settings who may not have access to neuropsychology clinics.

ACKNOWLEDGMENTS

There are no conflicts of interest to report. This study was financially supported by the Sierra Pacific Mental Illness Research, Education, and Clinical Center (MIRECC) at the VA Palo Alto Health Care System (B.P.Y.); an Alzheimer’s Association New Investigator

Research Grant (S.B., NIRG-09-133592); and the Department of Defense Telemedicine and Advanced Technology Research Center (TATRC, J.K.F., W81XWH-12-1-0584).

REFERENCES

- American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: American Psychiatric Association.
- Balthazar, M.L.F., Yasuda, C.L., Pereira, F.R.S., Bergo, F.P.G., Cendes, F., & Damasceno, B.P. (2010). Coordinated and circumlocutory semantic naming errors are related to anterolateral temporal lobes in mild AD, amnesic mild cognitive impairment, and normal aging. *Journal of the International Neuropsychological Society*, *16*, 1099–1107. doi:10.1017/S1355617710000998
- Bauer, R.M. (2012). Agnosia. In K.M. Heilman & E. Valenstein (Eds.), *Clinical neuropsychology* (5th ed., pp. 238–295). New York: Oxford University Press.
- Beeson, P.M., & Rapcsak, S.Z. (2006). The aphasia. In P.J. Snyder, P.D. Nussbaum & D.L. Robins (Eds.), *Clinical neuropsychology: A pocket handbook for assessment* (2nd ed., pp. 436–459). Washington, DC: American Psychological Association.
- Bell, B.D., Seidenberg, M., Hermann, B.P., & Douville, K. (2003). Visual and auditory naming in patients with left or bilateral temporal lobe epilepsy. *Epilepsy Research*, *55*, 29–37. doi:10.1016/S0920-1211(03)00110-4
- Benton, A.L., Sivan, A.B., de Hamsher, K.S., Varney, N.R., & Spreen, O. (1994). *Contributions to neuropsychological assessment* (2nd ed.). Orlando, FL: Psychological Assessment Resources, Inc.
- Benton, A.L., Varney, N.R., & Hamsher, K. (1978). Visuospatial judgment: A clinical test. *Archives of Neurology*, *35*, 364–367.
- Blumenfeld, H. (2010). *Neuroanatomy through clinical cases* (2nd ed.). Sunderland, MA: Sinauer Associates.
- Bond, T.G., & Fox, C.M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences* (2nd ed.). New York: Routledge.
- Brier, M.R., Maguire, M.J., Tillman, G.D., Hart, J., & Kraut, M.A. (2008). Event-related potentials in semantic memory retrieval. *Journal of the International Neuropsychological Society*, *14*, 815–822. doi:10.1017/S135561770808096X
- Brysaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*, 977–990.
- Chiang, H.-S., Mudar, R.A., Spence, J.S., Pudhiyidath, A., Eroh, J., DeLaRosa, B., ... Hart, J. (2014). Age-related changes in feature-based object memory retrieval as measured by event-related potentials. *Biological Psychology*, *100*, 106–114. http://dx.doi.org/10.1016/j.biopsycho.2014.05.010
- Cullum, C.M., Weiner, M.F., Gehrman, H.R., & Hynan, L.S. (2006). Feasibility of telecognitive assessment in dementia. *Assessment*, *13*(4), 385–390. doi:10.1177/1073191106289065
- Damasio, A.R., & Damasio, H. (1992). Brain and language. *Scientific American*, *267*, 88–95.
- Damasio, A.R., & Tranel, D. (1993). Nouns and verbs are retrieved with differently distributed neural systems. *Proceedings of the National Academy of Sciences, of the United States of America*, *90*, 4957–4960.

- Delis, D., Kaplan, E., & Kramer, J. (2001). *Delis-Kaplan Executive Function System (D-KEFS)*. San Antonio, TX: The Psychological Corporation.
- Delis, D., Kramer, J., Kaplan, E., & Ober, B. (2000). *California Verbal Learning Test – 2nd edition: Manual*. San Antonio, TX: The Psychological Corporation.
- Eye Diseases Prevalence Research Group. (2004). Causes and prevalence of visual impairment among adults in the United States. *Archives of Ophthalmology*, *122*, 477–485.
- Francis, W., & Kučera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar*. Boston: Houghton Mifflin.
- Goodglass, H., Kaplan, E., & Barresi, B. (2001). *The assessment of aphasia and related disorders* (3rd ed.). Austin, TX: Pro-Ed.
- Goodglass, H., & Wingfield, A. (Eds.). (1997). *Anomia: Neuroanatomical and cognitive correlates*. San Diego, CA: Academic Press.
- Gorno-Tempini, M.L., Hillis, A.E., Weintraub, S., Kertesz, A., Mendez, M., Cappa, S.F., ... Grossman, M. (2011). Classification of primary progressive aphasia and its variants. *Neurology*, *76*, 1006–1014.
- Grosch, M.C., Gottlieb, M.C., & Cullum, C.M. (2011). Initial practice recommendations for teleneuropsychology. *The Clinical Neuropsychologist*, *25*(7), 1119–1133. doi:10.1080/13854046.2011.609840
- Grossman, M., McMillan, C., Moore, P., Ding, L., Glosser, G., Work, M., & Gee, J. (2004). What's in a name: Voxel-based morphometric analyses of MRI and naming difficulty in Alzheimer's disease, frontotemporal dementia and corticobasal degeneration. *Brain*, *127*, 628–649. doi:10.1093/brain/awh075
- Hamberger, M.J., Goodman, R.R., Perrine, K., & Tamny, T. (2001). Anatomic dissociation of auditory and visual naming in the lateral temporal cortex. *Neurology*, *56*, 56–61.
- Hamberger, M.J., & Seidel, W.T. (2003). Auditory and visual naming tests: Normative and patient data for accuracy, response time, and tip-of-the-tongue. *Journal of the International Neuropsychological Society*, *9*, 479–489.
- Hamberger, M.J., Seidel, W.T., McKhann, G.M., & Goodman, R.R. (2010). Hippocampal removal affects visual but not auditory naming. *Neurology*, *74*, 1488–1493. doi:10.1212/WNL.0b013e3181dd40f0
- Hodges, J.R., & Patterson, K. (1995). Is semantic memory consistently impaired early in the course of Alzheimer's disease? Neuroanatomical and diagnostic implications. *Neuropsychologia*, *33*(4), 441–459.
- Huey, E.D., Goveia, E.N., Paviol, S., Pardini, M., Krueger, F., Zamboni, G., ... Grafman, J. (2009). Executive dysfunction in frontotemporal dementia and corticobasal syndrome. *Neurology*, *72*(5), 453–459.
- Kaplan, E.F., Goodglass, H., & Weintraub, S. (1978). *The Boston Naming Test: Experimental edition*. Boston: Kaplan & Goodglass.
- Kaplan, E.F., Goodglass, H., & Weintraub, S. (2001). *The Boston Naming Test* (2nd ed.). Philadelphia: Lippincott Williams & Wilkins.
- Kirshner, H.S., Webb, W.G., & Kelly, M.P. (1984). The naming disorder of dementia. *Neuropsychologia*, *22*(1), 23–30.
- Kohn, S.E., Lorch, M.P., & Pearson, D.M. (1989). Verb finding in aphasia. *Cortex*, *25*, 57–69.
- Kučera, H., & Francis, W. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
- Laine, M., & Martin, N. (2006). *Anomia: Theoretical and clinical aspects*. New York, NY: Psychology Press.
- Lee, A.G., & Coleman, A.L. (2004). Research agenda-setting program for geriatric ophthalmology. *Journal of the American Geriatrics Society*, *52*, 453–458.
- Lezak, M.D., Howieson, D.B., Bigler, E.D., & Tranel, D. (2012). *Neuropsychological assessment* (5th ed.). New York, NY: Oxford University Press.
- Malow, B.A., Blaxton, T.A., Sato, S., Bookheimer, S.Y., Kufta, C. V., Figlozzi, C.M., & Theodore, W.H. (1996). Cortical stimulation elicits regional distinctions in auditory and visual naming. *Epilepsia*, *37*(3), 245–252.
- McKhann, G.M., Knopman, D.S., Chertkow, H., Hyman, B.T., Jack, C.R., Kawas, C.H., ... Phelps, C.H. (2011). The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging – Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia*, *7*, 263–269. doi:10.1016/j.jalz.2011.03.005
- Miceli, G., Silveri, M.C., Nocentini, U., & Caramazza, A. (1988). Patterns of dissociation in comprehension and production of nouns and verbs. *Aphasiology*, *2*, 351–358.
- Miller, K.M., Finney, G.R., Meador, K.J., & Loring, D.W. (2010). Auditory responsive naming versus visual confrontation naming in dementia. *The Clinical Neuropsychologist*, *24*, 103–118. doi:10.1080/13854040903045074
- Nebes, R.D. (1989). Semantic memory in Alzheimer's disease. *Psychological Bulletin*, *106*(3), 377–394. doi:10.1037/0033-2909.106.3.377
- Nicholas, M., Barth, C., Obler, L.K., Au, R., & Albert, M.L. (1997). Naming in normal aging and dementia of the Alzheimer's type. In H. Goodglass & A. Wingfield (Eds.), *Anomia: Neuroanatomical and cognitive deficits* (pp. 5–27). San Diego, CA: Academic Press.
- Niemeier, J.P. (2010). Neuropsychological assessment for visually impaired persons with traumatic brain injury. *Neurorehabilitation*, *27*, 275–283. doi:10.3233/NRE-2010-0609
- Obler, L.K., & Albert, M.L. (1979). Action naming test. *Unpublished experimental edition*.
- Plassman, B.L., Langa, K.M., Fisher, G.G., Heeringa, S.G., Weir, D.R., Ofstedal, M.B., ... Wallace, R.B. (2007). Prevalence of dementia in the United States: The Aging, Demographics, and Memory Study. *Neuroepidemiology*, *29*, 125–132. doi:10.1159/000109998
- Rabin, L.A., Barr, W.B., & Burton, L.A. (2005). Assessment practices of clinical neuropsychologists in the United States and Canada: A survey of INS, NAN, and APA Division 40 members. *Archives of Clinical Neuropsychology*, *20*, 33–65.
- Randolph, C. (1998). *Repeatable battery for the assessment of neuropsychological status*. San Antonio, TX: The Psychological Corporation.
- Randolph, C., Lansing, A.E., Ivnik, R.J., Cullum, C.M., & Hermann, B.P. (1999). Determinants of confrontation naming performance. *Archives of Clinical Neuropsychology*, *14*(6), 489–496.
- Ryskulova, A., Turczyn, T., Makuc, D., Cotch, M.F., Klein, R.J., & Janiszewski, R. (2008). Self-reported age-related eye diseases and visual impairment in the United States: Results of the 2002 National Health Interview Survey. *American Journal of Public Health*, *98*, 454–461.
- Simmons, W.K., & Martin, A. (2009). The anterior temporal lobes and the functional architecture of semantic memory. *Journal of the International Neuropsychological Society*, *15*, 645–649. doi:10.1017/S1355617709990348
- Skelton-Robinson, M., & Jones, S. (1984). Nominal dysphasia and the severity of senile dementia. *British Journal of Psychiatry*, *145*, 168–171.

- Stern, R.A., & White, T. (2003). *Neuropsychological Assessment Battery: Administration, Scoring, and Interpretation Manual*. Lutz, FL: Psychological Assessment Resources.
- Strauss, E., Sherman, E.M.S., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary* (3rd ed.). New York: Oxford University Press.
- Trahan, D.E. (1998). Judgment of line orientation in patients with unilateral cerebrovascular lesions. *Assessment, 5*(3), 227–235.
- Tranel, D., Vianna, E., Manzel, K., Damasio, H., & Grabowski, T. (2009). Neuroanatomical correlates of the Benton facial recognition test and judgment of line orientation test. *Journal of Clinical and Experimental Neuropsychology, 31*(2), 219–233. doi:10.1080/13803390802317542
- Welch, L.W., Doineau, D., Johnson, S., & King, D. (1996). Educational and gender normative data for the Boston Naming Test in a group of older adults. *Brain and Language, 53*, 260–266.
- White, T., & Stern, R.A. (2003). *Neuropsychological assessment battery: Psychometric and technical manual*. Lutz, FL: Psychological Assessment Resources, Inc.
- Yochim, B., Baldo, J., Nelson, A., & Delis, D.C. (2007). D-KEFS Trail Making Test performance in patients with lateral prefrontal cortex lesions. *Journal of the International Neuropsychological Society, 13*(04), 704–709.
- Yochim, B.P., Kane, K.D., & Mueller, A.E. (2009). Naming test of the Neuropsychological Assessment Battery: Convergent and discriminant validity. *Archives of Clinical Neuropsychology, 24*, 575–583. doi:10.1093/arclin/acp053
- Yochim, B.P., Rashid, K., Raymond, N., & Beaudreau, S.A. (2013). How frequently are words used on naming tests used in spoken conversation? *The Clinical Neuropsychologist, 27*(6), 973–987. doi:10.1080/13854046.2013.797501
- Zec, R.F., Burkett, N.R., Markwell, S.J., & Larsen, D.L. (2007). A cross-sectional study of the effects of age, education, and gender on the Boston Naming Test. *The Clinical Neuropsychologist, 21*, 587–616. doi:10.1080/13854040701220028

APPENDIX: VERBAL NAMING TEST

VERBAL NAMING TEST

1ST EDITION (VNT-1)

Name: _____

Date: _____

SAY: “Now we are going to do something different. I’m going to describe an object or a verb and I want you to tell me the name of what I am describing.”

After each prompt, allow the examinee **10 s** to respond. If an incorrect response is given, say “**No, it’s something else**” and allow the examinee the remainder of the initial 10 s to respond. If no correct response is provided during the initial 10 s, provide the phonemic cue, saying “**It starts with the sound...** (underlined part of word)”. If after **10 s** from the phonemic cue they have not provided the correct word, proceed to the next item.

Say "WHAT IS THE NAME OF..." (can stop saying this after participant gets the nature of the task.)	✓ if correct 10 sec	PHONEMIC CUE	✓ if correct 10 sec
1 The part of your shirt that goes around your neck		<u>C</u> ollar	
2 The thing you hold over your head when it rains		<u>U</u> mbrella	
3 The country where the Great Pyramids are		<u>E</u> gypt	
4 The animal in the desert with a hump on its back		<u>C</u> amel	
5 What you do when you put your nose up to a flower		<u>S</u> nell	
6 A patch of facial hair men grow between their nose and mouth		<u>M</u> ustache	
7 What a ship does if it can no longer float		<u>S</u> ink	
8 A structure you drive over to cross a river		<u>B</u> ridge	
9 A period of ten years		<u>D</u> ecade	
10 A small amount of money left for the waiter at a restaurant		<u>T</u> ip	
11 What you use to sweep the floor		<u>B</u> room	
12 A baby cat		<u>K</u> itten	
13 The item of clothing to wrap around your neck in the winter		<u>S</u> carf	
14 A piece of land surrounded by water		<u>I</u> sland	
15 What you do with a razor		<u>S</u> have	
16 A large animal in Africa with a trunk		<u>E</u> lephant	
17 What you use to chop wood		<u>A</u> x	
18 The thing you put on a dog to go for a walk		<u>L</u> eash	
19 What you do to water to make it hot and steaming		<u>B</u> oil	
20 What you do with your money with charities or the church		<u>D</u> onate	
21 What ice does when it gets hot		<u>M</u> elt	
22 What you wipe your mouth with when eating		<u>N</u> apkin	
23 What you use to measure how many inches something is		<u>R</u> uler	
24 What you put your head on to sleep at night		<u>P</u> illow	
25 A long, severe snowstorm		<u>B</u> lizzard	
26 The part of your shirt that covers your arms		<u>S</u> leeves	
27 The tool used to collect leaves on the ground		<u>R</u> ake	
28 A pool of water on the ground		<u>P</u> uddle	
29 The kind of mountain that explodes with lava		<u>V</u> olcano	
30 The animal in Australia that hops around and has a pouch		<u>K</u> angaroo	
31 The African animal that's like a horse and has black and white stripes		<u>Z</u> ebra	
32 The person who works at a drugstore to fill prescriptions		<u>P</u> harmacist	
33 A device that measures the temperature		<u>T</u> hermometer	
34 A collection of thousands of stars		<u>G</u> alaxy	
35 A device used to help you add and subtract numbers		<u>C</u> alculator	
36 A moving set of stairs		<u>E</u> scalator	
37 The thing on a car that connects two wheels		<u>A</u> xle	
38 What a fish uses to breathe		<u>G</u> ills	
39 What you do when you close one eye but keep the other one open		<u>W</u> ink	
40 What someone sings into to make their voice louder		<u>M</u> icrophone	
41 What you do to a pencil or knife when it becomes dull		<u>S</u> harpen	
42 A place people go to gamble money		<u>C</u> asino	
43 A small hill made of sand		<u>D</u> une	
44 When you take a deep breath when you're sleepy or bored		<u>Y</u> awn	
45 What a horse does when it runs really fast		<u>G</u> allop	
46 A toy that has a string and floats in the air when it is windy		<u>K</u> ite	
47 A baby cow		<u>C</u> alf	
48 An animal in Africa with a very long neck		<u>G</u> iraffe	
49 What you wear while cooking that prevents food from getting on your clothes		<u>A</u> pron	
50 A book that is made up of different maps		<u>A</u> tlas	
51 A desert plant that has spikes.		<u>C</u> actus	
52 The poison a snake uses to kill its prey.		<u>V</u> enom	
53 The document you receive when you graduate from high school.		<u>D</u> iploma	
54 A kitchen appliance that cleans plates and glasses.		<u>D</u> ishwasher	
55 The river in Egypt that is the longest in the world		<u>N</u> ile	
TOTAL CORRECT WITHOUT CUE		TOTAL CORRECT WITH CUE	