Normative Data Collection and Comparison of Performance on the Poreh Naming Test to the Boston Naming Test

Orion R. Biesan
Cleveland State University

Follow this and additional works at: https://engagedscholarship.csuohio.edu/etdarchive

Recommended Citation
Biesan, Orion R., "Normative Data Collection and Comparison of Performance on the Poreh Naming Test to the Boston Naming Test" (2012). ETD Archive. 597.
https://engagedscholarship.csuohio.edu/etdarchive/597

This Thesis is brought to you for free and open access by EngagedScholarship@CSU. It has been accepted for inclusion in ETD Archive by an authorized administrator of EngagedScholarship@CSU. For more information, please contact library.es@csuohio.edu.
NORMATIVE DATA COLLECTION AND COMPARISON OF PERFORMANCE ON THE POREH NAMING TEST TO THE BOSTON NAMING TEST

ORION R. BIESAN

Bachelor of Science in Neuroscience/Psychology
Baldwin Wallace University
December, 2009

Submitted in partial fulfillment of the requirements for the degree MASTER OF ARTS IN PSYCHOLOGY at the CLEVELAND STATE UNIVERSITY August, 2012
This thesis has been approved

for the Department of Psychology

and the College of Graduate Studies by

Thesis Committee Chairperson, Dr. Amir Poreh, Ph.D.

Department & Date

Committee Member, Dr. Naohide Yamamoto, Ph.D.

Department & Date

Committee Member, Dr. Steve Slane, Ph.D.

Department & Date
ACKNOWLEDGEMENT

To my family, friends, and the professors who supported me and provided me with guidance through my journey for knowledge, thank you!
NORMATIVE DATA COLLECTION AND COMPARISON OF PERFORMANCE ON THE POREH NAMING TEST TO THE BOSTON NAMING TEST

ORION R. BIESAN

ABSTRACT

Although word-finding difficulty is commonly self-reported by older adults, there are no clinical instruments able to reliably distinguish normal age-related effects from pathology in word-finding impairments. The purpose of this study is two-fold: (1) design and evaluate the validity of the Poreh Naming Test, a novel electronic confrontation naming test used to evaluate naming difficulties in demented populations and (2) to investigate the effect of normal aging word-finding abilities on confrontation naming tests, using both accuracy and response latency as performance indices. A community sample was used with each participant over the age of 65 or younger participants reporting health problems shown to interfere with confrontation naming test performance also received the St. Louis University Mental Status Exam. The 57-item Poreh Naming Test used in this study was analyzed and refined to a 30-item test. Items were defined as easy, medium, or hard based on latency and proportion of the sample that correctly named the item. The Poreh Naming Test was found to be a valid measure of word-finding abilities
and was shown to better distinguish between mental status exam groups than the Boston Naming Test. However, the findings of this study do not support the hypotheses that normal aging has a negative impact on word-finding skills.

Cognitive status was the best predictor for accuracy and latency on the confrontation naming tasks and no effect of age was found on accuracy or latency in either confrontation naming test.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>viii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>ix</td>
</tr>
<tr>
<td>CHAPTER</td>
<td></td>
</tr>
<tr>
<td>I INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>II LITERATURE REVIEW</td>
<td>4</td>
</tr>
<tr>
<td>III METHOD</td>
<td>20</td>
</tr>
<tr>
<td>IV RESULTS</td>
<td>28</td>
</tr>
<tr>
<td>V DISCUSSION</td>
<td>47</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>56</td>
</tr>
<tr>
<td>APPENDICES</td>
<td></td>
</tr>
<tr>
<td>APPENDIX A: Original Poreh Naming Test Items</td>
<td>70</td>
</tr>
<tr>
<td>APPENDIX B: The Health and Communication</td>
<td>71</td>
</tr>
<tr>
<td>Screening Questionnaire</td>
<td></td>
</tr>
<tr>
<td>APPENDIX C: St. Louis University Mental</td>
<td>72</td>
</tr>
<tr>
<td>Status Examination</td>
<td></td>
</tr>
<tr>
<td>APPENDIX D: Scoring Rules for the Verbal</td>
<td>75</td>
</tr>
<tr>
<td>Fluency Tests</td>
<td></td>
</tr>
<tr>
<td>APPENDIX E: Demographic Survey</td>
<td>77</td>
</tr>
<tr>
<td>APPENDIX F: Second Phase of Test Construction – Item</td>
<td>78</td>
</tr>
<tr>
<td>Order for Poreh Naming Tests</td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX G: Procedure of Administration........................................80

APPENDIX H: Correlation’s Table.........................................................83

APPENDIX I: IRB Approval Form..........................................................85
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Demographic and background information for all subjects</td>
<td>29</td>
</tr>
<tr>
<td>II. St. Louis University Mental Status Examination Classifications and Demographics</td>
<td>35</td>
</tr>
<tr>
<td>III. Frequency of Health Problems in Sample</td>
<td>36</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

1. Mean latency times for Boston Naming Test items…………………39
2. Mean latency times for the Poreh Naming Test items………………..40
3. Proportion of the sample which correctly identified each
   Boston Naming Test item…………………………………………….41
4. Proportion of the sample which correctly identified each
   Poreh Naming Test item………………………………………………42
5. Latency times for St. Louis University Mental Status
   Examination derived groups for the Boston Naming Test…………43
6. Latency times for St. Louis University Mental Status
   Examination derived groups for the Poreh Naming Test……………44
7. Latency times for refined 30-item Poreh Naming Test………………45
8. Proportion of the sample which correctly identified each item
   on the refined 30-item Poreh Naming Test………………………….46
CHAPTER 1

INTRODUCTION

Although word-finding difficulty is commonly self-reported by older adults, there are no clinical instruments able to reliably distinguish normal age-related effects from pathology in word-finding impairment. Currently, the gold standard for evaluating word-finding abilities is the Boston Naming Test (also known as the BNT), which was developed 40 years ago and has not been improved upon since. Some of the criticisms of the current BNT include (a) some of the items are lower frequency words which results in a clear bias in favor of better educated people with a wider vocabulary (Hawkins & Bender, 2002), (b) the test is indirectly biased towards people of a higher socioeconomic status who can afford better education (Jefferson et al., 2007), (c) the pictures used in the BNT are simple black and white images, which can make recognition of the items more difficult due to ambiguity (Adlington, Laws, & Gale, 2009), and (d) latency is often not precisely measured or reported in the literature (Tsang & Lee, 2003). An examinee has 20 seconds to respond to the item, but this is not recorded and is done with a stopwatch. Time is not recorded because the examinee merely has to
finish in the time allowed, and there is no benefit to finishing an item quickly, nor is there any penalty for taking longer.

The Poreh Naming Test (referred to as the PNT) was created to address each of these concerns. The PNT is a computerized confrontation naming test where items are presented to participants on a computer screen. The benefits of a computerized version are primarily that administration of the test is more standardized between administrators, and that the software automatically measures latency, responses, and errors. In this study, 27 new items were added based on the methods of Miotto, Lucia, Camargo, and Scaff, (2010) and Tsang and Lee (2003). New items were included that had high word-frequency, perceived high familiarity, cultural relevance, and perceived to have high difficulty independent of word-frequency to make the PNT a test of actual naming abilities, as opposed to a test of vocabulary.

The goals of this study were two-fold: to evaluate the validity of the PNT and evaluate the effects of normal aging on word-finding abilities. Community populations were sampled and given a battery containing the BNT, the PNT, a semantic verbal fluency task, and phonemic verbal fluency tasks. Any participant over age 65 or younger participants reporting health problems known to interfere with word-finding abilities were also given a mental status exam to screen for mild neurocognitive impairment disorder-like performance or dementia-like performance. Participants were grouped based on mental status examination score
ranges to investigate the validity of the PNT and compare group differences on
the battery, particularly differences between BNT and PNT performances. In
addition, participants were grouped into three age groups: less than 65 years old,
between 65 and 80 years old, and greater than 80 years old. Performances on the
battery were compared to investigate the effects of normal aging.
CHAPTER II
LITERATURE REVIEW

Measuring Word-Finding Abilities. Confrontation naming tasks, particularly the Boston Naming Test (BNT), are the most common method to assess word-finding skills in clinical and normative populations (Nicholas, Barth, Obler, Au, & Albert, 1997; Schmitter-Edgecombe, Vesneski, & Jones, 2000; Strauss, Sherman, & Spreen, 2006). The BNT consists of presenting a person with a line-drawing of an object and asking the person to name each object presented. Originally designed for the detection of aphasia in clinical populations, the BNT is a well-researched confrontation naming test used to assess word-finding abilities in normal aging populations and clinical populations with neurodegenerative diseases (Goodglass, Kaplan, & Barresi, 2001; Thompson & Heaton, 1989).

The items on the BNT range from high-frequency, high familiarity objects (e.g. harmonica) to those that occur less frequently in word-frequency and are considered less familiar (e.g. trellis). If the subject does not know what the picture is (e.g. a harmonica), the test administrator can give them a stimulus cue.
(e.g. “it is a musical instrument you play by blowing”) with an additional 20 seconds following the cue (Kaplan, Goodglass, & Weintraub, 1983; Strauss et al., 2006). If the person is still unable to correctly name the object, a phonemic cue is provided (e.g. “it begins with H”) (Kaplan et al., 1983; Strauss et al., 2006). If the participant fails to provide the correct response in an additional 20 second period, the item is marked as incorrect and the test is continued until all items are completed or the participant fails to name six consecutive items (Strauss et al., 2006).

Demographic Factors Affecting Naming Ability. Many studies have been conducted to accumulate normative data for the BNT. In regards to accuracy on the BNT, the effects of demographic factors have been investigated in normative population sample performance. More specifically, the effects of gender, age, education, socioeconomic status, and race have been investigated in regards to their effect on BNT performance (Randolph, Lansing, Ivnick, Cullum, & Hermann, 1999).

The effect of gender on BNT performance remains equivocal. Several studies regarding the effect of gender on BNT performance have found that males outperform females (Jefferson et al., 2007; Randolph et al., 1999; Welch, Doneau, Johnson, & King, 1996) while other studies show no effect of gender (Barker-Collo, 2007; Fastenau, Denburg, & Mauer, 1998). While gender does not appear to yield consistent effect on total BNT performance, Randolph et al. (1999) and
Welch et al. (1996) found that gender differences were specific to groups of individual items. The number of items more frequently named correctly by males was approximately four times more items that of females with females with females only scoring higher than men on two items: asparagus and palette.

Welch et al. (1996) theorized that gender differences in BNT performance are attributable to the frequency at which certain items on the BNT are used in normal conversation by men compared to women due to occupational differences (e.g. compass, protractor, yoke). However, if gender effects were due to differences on these groups of items alone, it is unclear why several studies did not produce a gender effect utilizing the same stimuli. The overall findings and mixed results in the literature suggest that gender does not have a significant impact on BNT performance (Lezak, 2004; Zec Burkett, Markwell, & Larsen, 2007).

Education is a factor that has been consistently shown to influence BNT performance (Henderson, Frank, Pigatt, Abramson, & Houston, 1998; Kim, & Na, 1999; Steinberg, Bieliauskas, Smith, Langellotti, & Ivnik, 2005; Zec et al., 2007). Multiple regression analyses have found years of education to be the best predictor of BNT performance (Heaton, Avitable, Grant, & Matthews, 1999; Tombaugh & Hubley, 1997). Both Henderson et al. (1998) and Hawkins and Bender (2002) suggest that individuals with higher levels of education possess a wider vocabulary and thus have an unfair advantage leading to higher scores.
BNT performance was found to highly correlate with verbal intelligence, as measured by the Wechsler Adult Intelligence Scores-Revised (WAIS-R) indicating that the BNT may measure picture naming abilities and vocabulary rather than naming ability alone (Heaton et al., 1999; Steinberg et al., 2005; Thompson & Heaton, 1989).

Age appears to play a role on BNT performance with the literature showing that individuals over the age of 80 have the most difficulty in verbal naming abilities (Kent & Luszcz, 2002; MacKay, Connor, & Storandt, 2005). As the age of a population being studied increases, the score range on the BNT expands and standard deviations become larger (Nicholas, Brookshire, MacLennan, Schumacher, & Porrazzo, 1989; Zec et al., 2007). While there is some ambiguity as to the existence of natural age-related decline in naming ability, there is significantly more variance in performance on the BNT among older populations than younger populations which is suggestive of other mediating factors accounting for impaired performance on the BNT in older populations.

Research on the effect of race on BNT performance has yielded some evidence that minority populations perform poorer on confrontation naming tasks such as the BNT (Jefferson et al., 2007; Randolph et al., 1999; Saxton et al., 2000). Whitfield et al. (2000) found that multiple regression analysis of BNT scores between African Americans and European Americans produced different
prediction equations for BNT performance while failing to produce an alternative prediction equation on any other measure in their battery of tests in the study. Furthermore, populations of different ethnicities have been shown to make a significant amount of alternative naming errors on the BNT inadvertently impairing performance (e.g. mouth-organ for harmonica; walker for stilts) and accounting for a significant amount of incorrect answers on the BNT (Azrin et al., 1996; Calero, Arnedo, Navarro, Ruiz-Pedrosa, & Carnero, 2002; Cruice, Worrall, & Hickson, 2000). However, there is a significant interaction effect between race, socioeconomic status, and level of education leading to some debate in regards to the extent race may play a role in BNT performance (Henderson et al., 1998; Jefferson et al., 2007).

Several studies have addressed the effect of demographic variables on confrontation naming performance by creating different sets of norms based on a particular individual’s demographics and the normative performance of individuals with those similar demographics (Strauss et al., 2006). Others have created regression equations from normative data samples due to the need to control for non-pathological variable influences on neuropsychological test performance (Heaton et al., 1999). Regression-based corrections for demographics, however, are dependent on the representativeness of the population from which they were derived (Fastenau, 1998; Hawkins & Bender, 2002). The issue of demographic effects has not yet been resolved despite a great
deal of research investigating their effect on confrontation naming test performance.

*Additional Factors Affecting Naming Performance.* Word-frequency is the frequency at which words are assessed to occur in a language through speech and written materials. Word-frequency is a factor in word-finding abilities that positively and significantly correlates with an individual’s subjective familiarity of an item on picture naming tests (Brookshire & Nicholas, 1995; Snodgrass & Vanderwart, 1980). Naming accuracy on various picture naming tests, including the BNT, has been shown to positively correlate with word-frequency (Avila Lambdon Ralph, Parcet, Geffner, & Gonzalez-Darder, 2001). Items on the BNT with high word-frequency are named significantly more frequently than those with low word-frequency in clinical and normative populations (Brookshire & Nicholas, 1995; Hodgson & Ellis, 1998).

Name agreement is a necessity in the construction of confrontation naming tests. Items on the BNT where an object can only be given one correct name are said to have high name agreement whereas words with low name agreement possess multiple correct names (Hodgson & Ellis, 1998). In picture naming tests, words that have low name agreement take significantly longer to correctly name than words with high agreement, thus confounding the assessment of response latency (Budd, 2007; Hodgson & Ellis, 1998). A common issue is that a significant portion of errors on the BNT are the result of the use of common
alternative names for an object (e.g. mouthorgan for harmonica) (Azrin et al., 1996; Calero et al., 2002; Cruice et al., 2000). While revised versions of the BNT include acceptable alternatives for particular items, the revisions have not been shown to have a significant effect on reducing overall alternative naming errors since some of the alternative names used are incorrect and cannot be considered correct. For example, a rhinoceros is commonly misidentified as a hippopotamus (Budd, 2007).

Latency is a major factor involved in naming ability. Latency is defined as the time it takes an individual to correctly retrieve mental lexicons from semantic memory of the picture stimuli. Latency can be affected by numerous factors, including name-agreement, word-frequency, as well as age (Budd, 2007; Tsang & Lee, 2003). Most studies rely on accuracy alone, rather than accuracy and latency, despite latency times being shown to be slower in older populations (Goulet, Ska, & Kahn, 1994; Tsang & Lee, 2003). However, no normative data exist for latency and very little research has been conducted on latency as it is rarely reported in research utilizing the BNT. In studies where latency was not taken into account, an individual with a possible neurological deficit and one with no pathology could potentially both get credit for the same item although one theoretically could take as much time as needed if no restrictions on time were placed. Therefore, a key component of picture-naming is omitted which may distinguish normal aging processes from brain-damaged individuals (Tsang & Lee, 2003).
The simple black and white line drawings of particular objects in the BNT have been reported as visually complex and ambiguous, contributing to an increase in naming errors. The addition of color has been shown to reduce the number of errors not attributable to normal aging processes or cognitive impairment (Adlington et al., 2009; Laws & Hunter, 2006; Zannino et al., 2010). Therefore, the inclusion of color may aid in the semantic retrieval of the item and minimize irrelevant and distracting stimuli in items which may contribute to incorrect answers (Adlington et al., 2009). Using color in future versions of the BNT or novel confrontation naming tasks may aid in reducing errors that are not a sign of impaired word-finding abilities.

The variations in the literature in regards to adherence to the written administration and scoring instructions of the Boston Naming Test have also been the subject of debate. Lopez, Arias, Hunter, Charter, and Scott (2003) criticize the written administration and scoring instructions of the BNT as poorly written and allowing too wide a range of interpretations. Within the context of their study, it was found that different interpretations of scoring and administration can yield large differences in the total score of a participant. In addition, Ferman, Ivnik, and Lucas (1998) found large discrepancies in the scores of healthy older and demented populations when phonemic cues were counted as correct rather than failures. Budd (2007) discusses how variations in the type of BNT used (15, 30, 60, or 85 item test), administration, measurement of latency, and scoring makes it
difficult to compare studies in the literature since methodologies and test forms vary widely. Computerized confrontation naming tests may attenuate some of these problems by increasing the standardization of administration, scoring, and by including an automatic measure of latency rather having a researcher or clinician utilize an outside tool such as a stopwatch.

**Similar Tests and Shortened Versions.** Since the advent of the 60-item BNT, various researchers have formulated shortened and alternative tests to address issues regarding patients in clinical populations who have attention deficits by decreasing the amount of time necessary for testing and address issues regarding the test-retest effects (Kaplan et al., 1983; Lansing, Ivnik, Cullum, & Randolph, 1999; Roach, Schwartz, Martin, Grewal, & Brecher, 1996). The shortened tests, usually containing 15 or 30 items, have been shown to possess high-internal consistency and to significantly and positively correlate with the 60-item BNT in regards to differentiating between normal participants and those with anomia (Lansing et al., 1999). The scores from the short forms of the BNT significantly and positively correlate with 60-item BNT performance and were found to be equally reliable and valid in populations with aphasia or cognitive impairments (Del Toro et al., 2010). Demographic variables similarly affect shortened and original versions of the BNT (Lansing et al., 1999).

Other forms have been created to address some of the issues and criticisms of the BNT. The BNT-L is a 15-item short form of the BNT in which both latency
and accuracy scores are used to differentiate brain-damaged individuals from normal individuals and ascertain the effects of normal aging (Budd, 2007). Another version of the BNT, the incidental memory modification version of the BNT (memo-BNT), was created with the addition of free-recall, recognition of content, and recognition of temporal order to increase diagnostic accuracy within clinical populations (Karrasch et al., 2010). The memo-BNT was found to possess significantly more diagnostic sensitivity than the original BNT in differentiating between normal controls and patients with mild to moderate Alzheimer’s dementia compared to the original BNT and equivalent shortened forms (Karrasch et al., 2010).

The BNT has been translated into many different languages, including French, Korean, Danish, Portuguese, Finnish, Chinese, and Swedish. However, a majority of the BNT normative data regarding accuracy has been collected in North America (Kent & Luszcz, 2002). One problem with direct translation and use of the BNT is the effect of race and ethnicity on performance. To address this problem, alternative versions of the BNT were created with the intention of altering particular items of the BNT due to cultural differences (Kim & Na, 1999; Miotto, et al., 2010; Tsang & Lee, 2003). Barker-Collo (2001) found that participants from New Zealand who were matched to similar participants from North America did significantly worse than North Americans, suggesting cultural bias. While as many of the original BNT words are retained as possible, some
items are replaced with words that occur in higher frequency and are more culturally relevant to the population being studied (Miotto et al., 2010; Tsang & Lee, 2003). These alternative forms are created to reduce the effects of demographic variables, such as age and gender, and attenuate possible cultural biases of the BNT (Miotto et al., 2010; Tsang & Lee, 2003).

**Verbal Fluency.** Numerous studies compare both picture-naming tests, most commonly the BNT, with verbal fluency tests for the purpose of assessing word-finding skills in both clinical and normal populations (Thompson & Heaton, 1989). Verbal fluency tests, in addition to picture naming tests, are significantly useful in assessing word-finding deficits in clinical populations (Henry, Crawford, & Phillips, 2004). The purpose of verbal fluency tests is to measure the spontaneous production of words a participant can recall in a specific amount of time. There are two specific types of verbal fluency tests, phonetic and semantic (Strauss et al., 2006). Verbal fluency tests are conducted by prompting participants to name as many words as they can in one minute from a specific category. In phonemic fluency tests, participants are asked to name words that begin with a particular letter while semantic fluency tests have participants name words from a specific category (e.g. animals). Normative data exist for the total number of words generated in populations between the ages of 18 to 91 years old for both the phonetic and semantic versions of the test (Strauss et al., 2006).
There are three important components of verbal fluency test performance (Lezak, 2004; Strauss et al., 2006; Troyer, 2000; Troyer, Moscovitch, & Winocur, 1997; Troyer, Moscovitch, Winocur, Alexander, & Struss, 1998). The first performance index is the quantity of words produced in a 60 second time period (Lezak, 2004). The second performance index is clustering or the production of words into subgroups from memory and clustered together into semantic or phonemic subcategories (Troyer et al., 1997; Troyer et al., 1998). The output can be broken down into subcategories which are associatively related within semantic memory (Troyer et al., 1997; Troyer et al., 1998). Typically, an individual retrieves a cluster of items, or items semantically similar to each other, from semantic memory until a new cluster becomes available. Switching is the third performance index and is the process by which a participant utilizes the cognitive domain of executive functioning to switch to a new semantic cluster or return to a previously exhausted cluster. Once output for the subcategory is exhausted, the participant switches to a new subcategory (Troyer, 2000). Cognitive impairments would interfere with one’s ability to switch to a new cluster and cause the participant to produce larger cluster sizes (Troyer et al., 1997; Troyer et al., 1998). Normative data exist for verbal fluency performance in regards to clustering and switching for participants from 18 to 91 years old (Strauss et al., 2006; Troyer, 2000).
Verbal fluency tests provide several advantages to picture naming tests in detecting mild word-finding impairments (Dunn, Russell, & Drummond, 1989), distinguishing between different forms of dementia (Henry et al., 2004), and distinguishing between specific types of aphasia (e.g., fluent from non-fluent dysphasia) (Dunn et al., 1989). Significant correlations between verbal fluency skills and word-finding abilities in both healthy populations (Schmitter-Edgecombe et al., 2000) and clinical populations with dementia (Henry et al., 2004) are reported in the literature.

While verbal fluency tests are useful adjunct measures of word-finding skills, they lack the amount empirical support that exists for the BNT. However, both the BNT and verbal fluency tasks measure word-finding skills. Therefore, verbal fluency tasks are often used in conjunction with the BNT to assess aspects of word-retrieval in clinical and experimental populations.

*Mild Cognitive Impairment Disorder and Dementia.* In diagnosis of dementia, word-finding problems are commonly evaluated because it can be a symptom of all forms of dementia and is a clear sign of cognitive dysfunction (American Psychiatric Association, 2000). Dementia is a broad term used to describe a number of disorders characterized by the development of various cognitive deficits, including memory loss, as the result of various etiologies including general medical conditions (DSM-IV-TR). The research shows that individuals with dementia show significantly impaired performance on the BNT
Mild neurocognitive disorder (MNCD) is a disorder characterized by cognitive deficits beyond those attributable to the effects of normal aging (Petersen et al., 1999). MNCD is a condition that is clinically distinguishable from early Alzheimer’s dementia (Bennett, Schneider, Bienias, Evans, & Wilson, 2005). Approximately 7.5 to 15% of individuals with MNCD will develop dementia (De Jager, Hogervorst, Combrinck, & Budge, 2003). MNCD has been shown to impair performance on the BNT and verbal fluency tasks (De Jager et al., 2003).

St Louis University Mental Status Examination (SLUMS). The St. Louis University Mental Status Examination (SLUMS) is used to assess difficulty in orientation, memory, executive function, and attention (Tariq, Tumosa, Chibnall, Perry, & Morley, 2006). It is a 30 item test used for screening the elderly, or anyone suspected of having cognitive difficulty with more sensitivity than previous short form mental status exams such as the Mini-Mental Status Exam (MMSE) (Rosack, 2006). While the most common criticism of the SLUMS is that it is fairly new compared to other mental status exams such as MMSE (Farlow, Miller, & Pejovic, 2008; Rosack, 2006), the SLUMS is more sensitive than other shortened mental status exams in detecting mild neurocognitive impairment.
disorder (Osher, Wicklund, Rademaker, Johnson, & Weintraub, 2008; Rosack, 2006; Scazuferca, Almeida, Vallada, Tasse, & Menezes, 2009; Tariq et al., 2006).

Despite its popularity, the MMSE has been criticized and the SLUMS has become more prevalent in the literature (Zarit, Blazer, Orrell, & Woods, 2008). Some of the common criticisms of the MMSE include the variability in the detection of dementia, lack of accuracy in determining the severity of dementia in populations from lower socioeconomic statuses, populations with issues of illiteracy or limited literacy skills, education biases, and its inclusion of items that add noise rather than discrimination between normal and demented individuals (Ashford, Kolm, Colliver, Bekian, & Hsu, 1989; Brackhus, Laake, & Engedal, 1992; Tombaugh & McIntryye, 1992).

The SLUMS addresses some of the concerns associated with the MMSE by including separate scales for people who have less than a high school education and for those who have a high school education or greater. For the purposes of this study, it is imperative that the most sensitive tools be used to screen for mild cognitive impairments with a sample population in order to gather accurate data.

Hypotheses.

1. PNT accuracy scores and average latency would have significant positive correlations with performance on the BNT, BNT average latency, total output on the semantic verbal fluency, and total output on the phonemic verbal fluency establishing test validity and construct validity.
2. Total clusters and average cluster size, but not switching, in the semantic and phonemic fluency tasks would have significant positive correlations with performance on the BNT, PNT, and latency times.

3. The test-retest accuracy and latency on the PNT would have significant positive correlations with initial scores and latency establishing test-retest reliability.

4. Accuracy and latency times on the PNT would have a significant positive correlation with age and not significantly correlate to education.

5. Latency scores would be significantly different on both the BNT and PNT between older adults and younger adults. More specifically, the older adult group (>80 years of age or older) and the young-old group (66-80 years old) would have significantly higher average latency on the PNT and BNT compared to the youngest group (<65 years old).

6. Performance on the PNT, BNT, and total output on verbal fluency test would be significantly greater for those who score in the normal range than those who score in the MNCD range and DEM range on the SLUMS.

7. Age would have significant negative correlations with switching and total output and significant positive correlations with total clusters and average cluster size for the verbal fluency test.
CHAPTER III

METHOD

Participants. Forty-two participants received the PNT, a revised and improved visual naming test specifically designed to measure accuracy and latency (Appendix A). They also received the BNT, a phonemic Verbal Fluency task, and a semantic Verbal Fluency Task which is an item of the SLUMS. All participants were given a Health and Communication Screening Questionnaire with items adopted from Budd (2007) and Christensen, Multhaup, Nordstrom, and Voss (1991) to accurately define the population in regards to health factors that may affect naming ability (Appendix B). People over the age of 65 or individuals of a younger age who reported health problems that have been shown to affect BNT performance received the whole SLUMS (Appendix C) to make sure they are not suffering from dementia. A short demographic survey was also completed by each participant (Appendix E). See Appendix G for full procedure of administration. The scores of these forty-two participants were compared to establish validity. An eight person sub-sample of N was retested after a three month period to establish test-retest reliability.
The study was approved by the Cleveland State University Human Subjects Review Board. All participants were 18 years or older and informed consent was obtained from all research participants. Participants were recruited through the Parma Senior Center and Bay Village Senior Center between the months of February and May of 2012. Participants were tested individually in an office setting provided by the senior centers. Participants were not provided with reimbursement for participation and subject codes were assigned to all participants to keep all information anonymous outside of the informed consent. All participants were native English speakers. Exclusion criteria for the study included the following: being told by a physician that they had conditions or symptoms associated with cognitive decline, being a non-native English speaker, and/or previous head injury with loss of consciousness.

The sub-sample for retesting after a three month period (n=8) was selected primarily based on availability, with a smaller emphasis placed on finding people of differing age groups and education levels to retake the test.

Test Construction. The original PNT software and items were developed by Poreh (2009). Martincin (2010) administered the original PNT to a community sample and refined the original test into a 30-item test based on average latency times. In this study, 27 new items were added with the intention of adding items that were more difficult for participants to name in regards to latency and accuracy.
New items were selected using word-frequency. Word-frequencies were determined using the Corpus of Contemporary American English (Davies, 2010). Davies (2010) assessed word-frequency from both spoken (unscripted conversation radio shows and unscripted conversation from television shows) and non-spoken sources (books, short-stories, magazines, newspapers, and academic journals). Using the methods of Miotto et al. (2010) and Tsang and Lee (2003), the test was constructed including items of varying difficulty based on cultural relevance, word-frequency as assessed by the Corpus of Contemporary American English (Davies, 2010), high perceived difficulty, and high perceived familiarity. The aim was to include items that are well known by the general population, independent of age and years of education, and of varying difficulty.

Once data collection and analysis were completed, the test was refined from 57 items to 30 items. Items were discarded if determined to be ambiguous, have low-name agreement, or showed a clear bias for a particular gender, education level, or age group. Acceptable alternative names for items were evaluated based on the proportion of the sample which used the alternative names. Next, the latency and accuracy of each item on the PNT was examined. The items were then arranged according to difficulty into three categories (easy, moderate, and difficult). Difficulty was determined by accuracy and latency scores.

The Health and Communication Screening Questionnaire (HCSQ). The HCSQ (See appendix B) is a 26 item questionnaire, with select items adopted
from Budd (2007) and Christensen et al. (1991) which asks participants if they have any of the listed medical conditions. These medical conditions may confound their performance on the BNT, PNT, and verbal fluency tests.

_Poreh Naming Test (PNT)._ The PNT is a computer-based word-naming test administered on computer with a screen for the participant to view the stimuli, which was attached to a computer with the appropriate software for the administrator to oversee the test. The software and original items of the PNT were developed and tested by Poreh (2009). The participant viewed the stimuli on the electronic screen, while the administrators recorded his or her responses, latency, whether a phonemic or semantic error was made, and precisely what that error was. If a participant is having trouble with the stimulus item, a semantic cue was given (e.g. for broccoli “it is a vegetable”) and if they continue to have trouble, a phonemic cue was given (e.g. it begins with “br”) after a 20 second period.

Latency times were digitally recorded for each individual item but marked as incorrect for accuracy if the participant was unable to correctly name the item within 20 seconds of presenting the stimulus. Average latency was assessed using the methods of Tsang and Lee (2003) where average reaction time was quantified by summing all correctly named items by a participant, then subtracting items where cues were given or that were not correctly named, and finally dividing by the total number of items scored.
**Verbal Fluency Tasks.** The semantic and phonemic verbal fluency tasks were administered, scored, and analyzed by the “Quantified Process Scoring System” software (Poreh, 2009). For the semantic verbal fluency task, the subject was asked to name as many animals as possible in one minute. On the phonemic verbal fluency task, the subject was asked to name as many words as possible in one minute that began with a particular letter. All responses were recorded in order electronically. Four scores were obtained from each fluency task, including the total number of clusters, mean cluster size, the raw number of switches, and the total number of correct words generated (detailed rules for defining clusters, scoring cluster size, and switches for phonemic and animal are printed in Appendix D).

**Data Analysis.** Scores were analyzed using Pearson product-moment correlations comparing total score and average latency on the PNT compared to both BNT total score and average latency to establish test validity.

Pearson product-moment correlations were used to examine the relationships of the total output, the total number of clusters, average cluster size, and switches in semantic and phonemic verbal fluency to PNT and BNT accuracy and latency.

Test-retest reliability was assessed in a subset of participants (n=8) after a three month period using Pearson product-moment correlations and Spearman rank-order correlations comparing initial scores on the PNT, BNT, latency,
semantic verbal fluency performance, and phonemic verbal fluency performance to retested performance on the PNT, BNT, semantic verbal fluency performance, and phonemic verbal fluency performance.

Next, a stepwise multiple regression analysis was used to independently assess the relationship of accuracy and latency on the PNT to age, years of education, and SLUMS scores. Four independent analyses were conducted with PNT accuracy, PNT latency, BNT accuracy, and BNT latency being the dependent variables of each model, respectively. The independent variables for all models were age, education, SLUMS scores, and gender.

Post-hoc partial correlations were used to examine the role of age and education on verbal fluency performance controlling for the effect of SLUMS. Since SLUMS scores were found to be the most significant predictor of naming abilities in the previous analyses, it was necessary to evaluate the relationships between these variables after partialling out the effects of SLUMS scores.

A one-way analysis of covariance (ANCOVA) was used to assess whether there were differences in average latency times between age groups. The independent variable, age groups, involved three levels: the young group (<65 years old), the young-old group (65-80 years old), and the old group (>80 years old). The dependent variable was the average PNT latency time and the covariate was the SLUMS scores.
A Kruskal-Wallis one-way analysis of variance (ANOVA) was used to evaluate differences between SLUMS group on PNT and BNT performance. Follow-up Mann-Whitney U tests were conducted, controlling for Type I error across tests using the Bonferroni approach for an adjusted p value of 0.016. Pairwise differences were evaluated between the three SLUMS groups: normal, mild neurocognitive disorder (MNCD), and dementia (DEM).

A Kruskal-Wallis one-way analysis of variance (ANOVA) was used to evaluate differences between SLUMS group on semantic and phonemic verbal fluency performance. More specifically, differences in output, total clusters, average cluster size, and switching were evaluated. Follow-up Mann-Whitney U tests were conducted, controlling for Type I error across tests using the Bonferroni approach for an adjusted p value of 0.016, to evaluate pairwise differences among the three SLUMS groups: normal, mild neurocognitive disorder (MNCD), and dementia (DEM).

A spearman rank-order correlation was conducted to assess the relationship between item difficulty for both the BNT and PNT to SLUMS groups and gender to assess whether there were similar or dissimilar profiles of performance across items.

Finally, a Kruskal–Wallis one-way analysis of variance was used to determine if there were significant differences in performance on the PNT, BNT, and verbal fluency measures between health groups. Follow-up Mann-Whitney U
tests were conducted, controlling for Type I error across tests using the Bonferroni approach for an adjusted p value of 0.016, to evaluate pairwise differences among the three groups: no health problems, one to two health problems, and three or more health problems.
CHAPTER IV

RESULTS

For the community sample that scored normal on the SLUMS (n=20), the mean score on the PNT was 53.79 out of 57 items. The mean score on the BNT was 57.21 out of 60 items. For the sample which scored in the mild neurocognitive disorder range (MNCD) on the SLUMS (n=17), the mean PNT score was 49.94 out of 57 while the mean BNT score was 53.72 out of 60. For the dementia group (n=5), the mean PNT score was 37.4 out of 57 on the PNT and 37.6 out of 60 on the PNT. See Table I for demographic information and SLUMS performance of the total sample. Consistent with the first hypothesis, a Pearson product-moment correlation revealed a significant positive correlations between BNT accuracy and PNT accuracy (r=0.922, p<0.001) and BNT latency and PNT accuracy scores (r=0.767, p<0.001). Accuracy on the BNT significantly and negatively correlated with BNT average latency time (r= -0.888, p<0.001). Similarly, PNT accuracy significantly and negatively correlated with PNT average latency time (r= -0.916, p<0.001). Figures 1 and 2 show mean latency times for the total sample on individual BNT and PNT items, respectively.
Figures 3 and 4 show the proportion of the total sample which correctly name individual items on the BNT and PNT, respectively. Overall, these results establish the test validity of the PNT. See Appendix H for a list of all correlations and significance values.

Table I: Demographic and background information for all subjects

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>N (females/males)</td>
<td>42 (33/9)</td>
</tr>
<tr>
<td>Age (M ± SD), years</td>
<td>72.83 ± 11.64</td>
</tr>
<tr>
<td>Education (M ± SD), years</td>
<td>13.98 ± 3.10</td>
</tr>
<tr>
<td>SLUMS Score (M ± SD)</td>
<td>24.69 ± 5.31</td>
</tr>
</tbody>
</table>

As predicted, a Pearson product-moment correlation revealed that PNT accuracy significantly positively correlated with semantic fluency total output (r=0.603, p<0.001) and phonemic verbal fluency total output (r=0.498, p=0.001). Post-hoc analysis revealed that BNT accuracy also positively correlated significantly with semantic fluency total output (r=0.542, p<0.001) and phonemic verbal fluency total output (r=0.495, p=0.001). These results are consistent with the first hypothesis that the PNT is a valid measure of word-finding abilities. However, latency scores did not yield any unique or stronger relationships between variables. Only the directions of relationships changed.

Consistent with the second hypothesis, a Pearson product-moment correlation revealed that PNT accuracy scores positively and significantly correlated with total semantic clusters (r=0.378, p=0.014) and total phonemic clusters (r=0.421, p=0.006). Inconsistent with the prediction, PNT accuracy
scores were significantly negatively correlated with average semantic cluster size ($r = -0.379, p=0.013$), failed to correlate with average phonemic cluster size, and significantly positively correlated with both semantic fluency switching ($0.419$, $p=0.006$) and phonemic switches ($r=0.437, p=0.004$). Latency scores did not yield any unique or stronger relationships between variables. Only the directions of relationships changed. As predicted in the second hypothesis, BNT accuracy scores had significant positive correlations with total phonemic clusters ($r=0.411, p=0.007$). Inconsistent with the prediction, BNT accuracy scores were significantly negatively correlated with average semantic cluster size ($r= -0.494, p=0.001$), failed to correlate significantly to total semantic clusters or average phonemic cluster size, and significantly positively correlated with both semantic switches ($r=0.465, p=0.002$), and phonemic switches ($r=0.490, p=0.001$). Latency scores did not yield any unique or stronger relationships between variables. Only the directions of relationships changed.

Test-retest reliability was analyzed using Pearson product-moment correlations and Spearman rank-order correlations for the PNT, BNT, and verbal fluency tasks. Data was collected from a small subset of participants ($n=8$) who were retested after a three month period. Because Spearman and Pearson analyses yielded similar results, only the Pearson product-moment correlation results are presented. As predicted in the third hypothesis, there were significant positive correlations with initial and retest PNT accuracy scores ($r=0.990, p<0.001$) and
retest PNT scores and retest BNT accuracy scores \((r=0.985, p<0.001)\). In addition, significant and positive correlations were revealed with initial and retest BNT accuracy scores \((r=0.989, p<0.001)\), semantic fluency total output \((r=0.751, p<0.032)\), and phonemic fluency total output \((r=0.893, p=0.003)\). In regards to clustering and switching, there was a significant positive correlation between total semantic clusters \((r=0.827, p=0.011)\), total phonemic clusters \((r=0.773, p=0.024)\), average semantic cluster size \((0.734, p=0.038)\), and phonemic switches \((r=0.868, p=0.005)\). No significant correlations were found for semantic switches or average phonemic cluster sizes.

Stepwise regressions were conducted to evaluate the influence of age, education, gender, and SLUMS score on PNT and BNT total score and average latency time. Analysis revealed that SLUMS score was the most significant predictor of PNT total score \((R^2=0.765, F(1,40)=130.339, p=<0.001)\), PNT average latency \((R^2=0.675, F(1,40)=83.086, p<0.001)\), BNT performance \((R^2=0.710, F(1,40)=97.908, \ p<0.001)\), and BNT average latency \((R^2=0.577, F(1,40)=54.629, p<0.001)\). Consistent with the fourth hypothesis, gender, and education failed to significantly predict PNT total score, PNT average latency, BNT total score, or BNT average latency. However, the findings that age did not significantly predict PNT performance was not consistent with the fourth hypothesis.
A one-way analysis of covariance (ANCOVA) was conducted with the independent variable, age groups, possessing three levels: the young group (<65 years old), the young-old group (65-80 years old), and the old group (>80 years old), the dependent variable of average PNT latency time, and the covariate of SLUMS scores. Inconsistent with the fifth hypothesis, the ANCOVA was not significant for age-groups. However, there was significant main effect of SLUMS scores \((F(2,42) =20.14, \ p <.005)\). The strength of the relationship between SLUMS score and the dependent variable was strong, as assessed by \(\eta^2\), with the SLUMS scores accounting for 35.9 percent of the variance.

Post-hoc partial correlations were conducted to evaluate the relationship between both age and education to semantic and phonemic verbal fluency performance partialling out the effects of SLUMS scores. The partial correlations revealed a significant negative correlation between age and phonemic fluency total output \((r= -0.371, \ p=0.017)\). This analysis also yielded non-significant but suggestive negative correlations for semantic fluency total output \((r= -0.304, \ p=0.054)\) and phonemic switches \((r= -0.278, \ p=0.078)\) suggesting a trend that may not have been significant due to the small sample size used in this study. No significant correlations were found for education in partial correlations controlling for the effect of SLUMS scores.

A Kruskal–Wallis one-way analysis of variance tests revealed significant differences between SLUMS groups for PNT total scores \(\chi^2(2)=19.289,\)
p<0.001), BNT total scores ($\chi^2(2)=15.784$, p<0.001), PNT average latency ($\chi^2(2)=17.124$, p<0.001), BNT total scores ($\chi^2(2)=15.784$, p<0.001), BNT average latency ($\chi^2(2)=15.527$, p<0.001). Figures 5 and 6 show mean latency times for the normal, MNCD, and DEM groups on each of the BNT and PNT items, respectively.

Follow-up tests were conducted to evaluate pairwise differences among the three groups controlling for Type I error across tests by using the Bonferroni approach for an adjusted $p$ value of 0.016. As predicted in the sixth hypothesis, pairwise comparisons revealed significant difference between the DEM group and the Normal group in the following: PNT accuracy scores ($z=-3.413$, $p=0.001$), PNT average latency ($z=-3.397$, $p=0.001$), BNT accuracy scores ($z=-3.351$, $p=0.001$), and BNT average latency ($z=-3.262$, $p=0.001$). As predicted, significant differences between the DEM group and the MNCD group were as follows: PNT accuracy scores ($z=-3.260$, $p=0.001$), PNT average latency ($z=-3.135$, $p=0.002$), BNT accuracy scores ($z=-2.949$, $p=0.003$), and BNT average latency ($z=-2.468$, $p=0.014$). Differences in performance between the MNCD group and normal group were as follows: PNT total ($z=-2.894$, $p=0.004$), PNT average latency ($z=-2.469$, $p=0.014$), and BNT average latency ($z=-2.621$, $p=0.009$). Inconsistent with the prediction, BNT accuracy scores ($z=-2.305$, $p=0.021$) failed to reach significance at the adjusted 0.017 level. While PNT scores were significantly higher for the normal group compared to all other
SLUMS groups as predicted, BNT accuracy scores were not significantly different between the normal and MNCD group.

A Kruskal-Wallis one way analysis of variance revealed significant differences between SLUMS groups in the following components of the verbal fluency tasks: semantic fluency total output ($\chi^2(2)=12.629, p=0.002$), phonemic total output ($\chi^2(2)=6.656, p<0.036$), and phonemic clusters ($\chi^2(2)=9.090, p=0.011$). Semantic fluency switches ($\chi^2(2)=5.742, p=0.052$) and phonemic switches ($\chi^2(2)=5.459, p=0.065$) showed a trend of possible differences between groups but failed to reach significance. There were no significant differences in total semantic clusters, average semantic cluster size or average phonemic cluster size. See Table II for demographics of the SLUMS groups.

Follow-up tests were conducted to evaluate pairwise differences, among the three groups, controlling for Type I error across tests by using the Bonferroni approach for an adjusted p value of 0.017. The results of these tests indicated a significant difference between the DEM group and the Normal group in the following: semantic total output ($z=-3.341, p=0.001$), phonemic total output ($z=-2.484, p=0.013$), and phonemic clusters ($z=-2.384, p=0.017$). Significant differences between the DEM group and the MNCD group were as follows: semantic fluency total output ($z=-3.276, p=0.001$) and phonemic clusters ($z=-2.956, p=0.003$). Phonemic total output ($z=-2.274, p=0.023$) reached significance at the 0.05 level but failed to reach significance at the adjusted 0.017 level. No
differences were found between the normal and the MNCD group. These findings are not consistent with the prediction that the normal group would produce significantly more output on verbal fluency tasks than both the MNCD and DEM groups.

Table II: SLUMS Classifications and Demographics

<table>
<thead>
<tr>
<th></th>
<th>Males</th>
<th>Females</th>
<th>Age (M ± SD), years</th>
<th>Education (M ± SD), years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>5</td>
<td>15</td>
<td>67.32 (+12.06)</td>
<td>15.11 (+2.47)</td>
</tr>
<tr>
<td>MNCD</td>
<td>3</td>
<td>14</td>
<td>75.50 (+9.46)</td>
<td>13.11 (+2.68)</td>
</tr>
<tr>
<td>DEM</td>
<td>1</td>
<td>4</td>
<td>84.20 (+3.83)</td>
<td>12.80 (+5.40)</td>
</tr>
</tbody>
</table>

The proportion of correct responses was calculated for each of the 60 items for the BNT and 57 items on the PNT for each SLUMS group (e.g. if 65% of the MNCD participants correctly named “dart,” that item would have a score of 0.65 for the MNCD group). These numbers, which represent the item difficulty for each item for a particular group, were found to be highly correlated among groups utilizing a Spearman rank-order correlation. For the BNT, item difficulty was significantly correlated between the following groups: Normal and MNCD: (r=0.671, p<0.001); Normal and DEM: (r=0.641, p<0.001); MNCD and DEM: (r=0.733, p<0.001). For the PNT, item difficulty was also significantly correlated between the SLUMS groups: Normal and MNCD: (r=0.785, p<0.001); Normal and DEM: (r=0.697, p<0.001); MNCD and DEM: (r=0.73, p<0.001). Item difficulty was also significantly correlated between males and females for the PNT (r=0.776, p<0.001) and BNT (r=0.659, p<0.001). This indicates that the
profile of performance across items on both tests was highly similar across the SLUMS groups and between males and females.

A Kruskal–Wallis one-way analysis of variance was used to determine if there were significant differences in performance on the PNT, BNT, and verbal fluency measures between health groups. See Table III for the frequency at which health problems were reported in the sample and group sizes. No significant differences were found between groups for PNT latency or total items correct, BNT latency or total items correct, semantic and phonemic fluency total output, total clusters, average cluster size, or switches.

Table III: Frequency of Health Problems in Sample

<table>
<thead>
<tr>
<th>Health Problems</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Health Problems</td>
<td>27</td>
</tr>
<tr>
<td>One to Two Health Problems</td>
<td>9</td>
</tr>
<tr>
<td>Three or more Health Problems</td>
<td>6</td>
</tr>
</tbody>
</table>

Following data analysis, the refinement of the PNT was conducted using two steps. The first step was evaluating each item on the PNT in regards to the accuracy at which the items were correctly named. The only item found to be too difficult and possibly measuring vocabulary rather than naming abilities was “catapult” which was correctly named by only 33% of the total sample. For the second step, each item on the PNT was examined in relation to the response time and difficulty for the clinical population. This allowed for the original 57 item versions of the PNT to be refined into a 30 item tests. The first ten items on the test are considered “easy”, the middle ten items “medium”, and the final ten items
“hard”, so there would be a progression in difficulty for each task, as displayed in figures 7 and 8. Difficulty was used to establish which items were in the easy, medium, or hard groups, and was based on reaction time, the accuracy at which items were correctly named by the total sample, and the accuracy at which items were correctly named by those in the normal SLUMS group. Figure 7 shows mean reaction times of the total sample on individual items on the refined 30-item PNT. Figure 8 shows the proportion of the total sample that correctly named individual items on the refined 30-item PNT. For the easy group, items with a faster reaction time (M=3.22) and correctly identified by a high proportion of the total sample (M=92.84%) were used. For the medium group, items with moderate reaction times (M=5.415), that were correctly identified by a high proportion of the normal group (M=92.08%), and that were found to be more difficult to name by the overall sample (M=79.73%) were used. For the hard group, items with the slowest reaction times (M=6.595), that were correctly identified by a high proportion of the normal group (M=88.38%), and that were found to be more difficult to name by the overall sample (M=73.30%) were used. Of the 30 items retained, 4 were found to have acceptable alternatives and were retained. Each the alternative names were utilized by at least 20% of the population and believed to be influenced by demographic factors although future research would be required to validate demographic differences. See Appendix F for the new order of items,
proportion of the sample which correctly named the item, item latency, and acceptable alternatives.

Upon completing the refinement of the PNT, ex post facto analyses using Pearson product-moment correlations to evaluate the relationship between the refined 30-item PNT accuracy scores to the 57-item PNT accuracy scores, BNT accuracy scores, the SLUMS, verbal fluency tasks, and demographic variables. The refined 30-item PNT was found to have positive significant correlations with the full 57-item PNT (r=0.984, p<0.001), the BNT (r=0.917, p<0.001), and the SLUMS (r=0.853, p<0.001). In regards to semantic fluency, the 30-item PNT had significant positive correlations with total output (r=0.575, p<0.001), semantic clusters (r=0.502, p=0.001), and semantic switches (r=0.391, p=0.01). In regards to phonemic fluency, the 30-item PNT had significant positive correlations with phonemic total output (r=0.502, p=0.001), phonemic clusters (r=0.402, p=0.008), and phonemic switches (r=0.461, p=0.002) while again failing to correlate to average phonemic cluster size. Partial correlations controlling for the effect of SLUMS scores revealed no significant correlations with age or education.
Figure 1: Mean latency times for Boston Naming Test items.
Figure 2: Mean Latency times for Poroh Naming Test items.
Figure 3: Proportion of the sample which correctly identified each Boston Naming Test item.
Figure 4: Proportion of the sample which correctly identified each Poreh Naming Test item.
Figure 5: Latency times for St. Louis University Mental Status Examination derived groups for the Boston Naming Test.
Figure 6: Latency times for St. Louis University Mental Status Examination derived groups for the Poreh Naming Test
Figure 7: Latency Times for refined 30-item Poreh Naming Test
Figure 8: Proportion of the sample which correctly identified each item on the refined 30-item Poreh Naming Test.
CHAPTER V
DISCUSSION

In the community sample, the hypothesis that performance on the BNT, PNT, and verbal fluency total output would all correlate significantly was supported. When SLUMS group performances were compared, the PNT demonstrated better sensitivity in distinguishing between the normal and the MNCD groups. PNT, but not BNT accuracy scores and performance on verbal fluency tasks, were significantly different between the normal and MNCD groups with the normal group outperforming the MNCD group. This suggests that the PNT possesses a clear advantage over the BNT in detecting word-finding deficits that are able to differentiate between the effects of normal aging and cognitive impairments present in individuals with MNCD. When evaluating the validity of the PNT in a larger community sample, an additional focus of the research should be to replicate these findings. Future research should investigate the diagnostic sensitivity of the PNT in differentiating between normal controls and patients with MNCD.
In regards to the hypotheses of this experiment evaluating verbal fluency performance to confrontation naming test performance and evaluating group differences in performance, clustering and switching proved to be problematic to analyze using the method of scoring known as the Troyer scoring system (Troyer et al., 1997, 1998). A number of researchers have criticized and reviewed various aspects of the Troyer scoring system despite evidence that clustering and switching is impaired in patients with temporal lobe or frontal lobe lesions, respectively (Troyer et al., 1998). Abwender, Swan, Bowerman, and Connolly (2001) suggests that while positive correlations exist between clustering and total items produced, there is not enough evidence to postulate that clustering is a strategic process that leads to more output. In fact, it is suggested that more output could lead to what appears to be meaningful clusters occurring through chance. However, the results of this study show that there is a significant positive relationship between total semantic and phonemic clusters and PNT performance. In addition, the test-retest reliability was established for semantic output, phonemic output, total clusters, and average semantic cluster size, but not for average phonemic cluster size.

Alternative, but less widely used, methods of calculating clusters allow for a researcher to combine clusters phonemically related within the semantic fluency task (buck and duck), the generation of exemplars starting with subsequent letters of the alphabet, and the combination of semantically related items within the
phonemic fluency task (shirt, socks, sweater) (Abwender et al., 2001). Future studies should utilize these alternative clustering methods in order to better understand the relationship between clustering and confrontation naming performance. However, the findings of this study show that in contrast with the predicted outcome, average semantic cluster size was significantly negatively correlated with confrontation naming performance and average phonemic cluster size did not significantly correlate with confrontation naming test performance.

In regards to switching, Abwender et al. (2001) suggests distinguishing between what they referred to as “real switches” (switches between multi-word clusters) from “hard switches” (switches between isolated words). Only “real switches” would underlie the strategic executive processes while “hard switches” would be a measure of general processing speed. The evidence from this study and from the literature suggests that automatic software scoring clusters and switches may need to be modified in order to better represent the possible cognitive strategies used by participants and differentiate between “hard switches” and “real switches” to better understand the qualitative strategies utilized by participants in the process of verbal fluency tasks. The findings of this study show that switching significantly correlated to confrontation naming test performance. In addition, the test-retest reliability of phonemic switching, but not semantic switching, was established.

While further testing is necessary to evaluate the effects of demographic
variable influence on PNT performance in larger normative samples, the initial results of this study are consistent with the hypothesis education and gender would not significantly predict PNT scores. The finding that SLUMS scores were the most significant predictor of PNT performance further establishes that the PNT is a valid measure of word-finding abilities.

However, the hypothesis that differences between age-groups would be detected using latency on the PNT was not confirmed. Budd (2007) found results similar to this study using a 15-item version of the BNT and assessing age-related differences in latency in 1235 participants. It is also possible that age-group differences, similar to those found by Tsang and Lee (2003), could not be detected due to the small sample size of adults who scored in the normal range on the SLUMS. While this study found no significant difference between age groups in latency, future research with access to larger sample sizes is required particularly since latency on confrontation naming tasks has been scarcely investigated or reported in the literature.

The hypothesis that the young group (<65 years old) would significantly outperform the young-old group (65-80 years old) and the old group (>80 years old) in confrontation naming performance was not confirmed. One possibility is that age-related naming decline was not detected because this study utilized a cross-sectional design rather than a longitudinal design. Kent and Luszcz (2002) found that after following a large sample of community-dwelling participants that
naming declines were most detectable when participants reached the age of 80 to 84 years old over a six year period. Another possibility is that the sample size of adults who scored in the normal range on the SLUMS was too small to yield a significant effect of age. The final possibility is that age-related differences were not detected due to the use of the SLUMS, an instrument found to be highly sensitive to mild cognitive impairments and lacking in educational biases found in other mental status exams (For a listing of current measures found to be sensitive to mild cognitive impairment, see Ashford, 2008). Studies which utilize measures not sensitive to mild cognitive impairment or with educational biases, such as the MMSE, may incorrectly classify participants and may lead to poor performance on tasks such as the BNT being identified as an effect of age rather than mental status.

In regards to evaluating SLUMS group differences in verbal fluency performance, the normal group showed no significant differences in total output when compared to the MNCD group. However, both the normal group and the MNCD group produced significantly more output on the verbal fluency tasks compared to the DEM group. In regards to the dissociable variables of clustering and switching, those in the normal and MNCD group produced significantly more total phonemic clusters than the DEM groups. The findings of this study in verbal fluency performance were expected given that those with cognitive decline have a harder time retrieving information from semantic memory, show impaired
performance on tasks measuring executive functioning such as the Wisconsin Card sorting task (Lezak, 2004), and show biological abnormalities within prefrontal cortex and the temporal lobes which are involved in executive functioning and semantic memory, respectively (Henry et al., 2004; Tsang & Lee, 2003).

The hypothesis of the effect on age and verbal fluency performance was only partially substantiated. Older participants produced significantly less output on the phonemic fluency task. While older participants showed a trend of producing less output on the semantic verbal fluency task and switching less on the phonemic verbal fluency task, both findings failed to reach statistical significance. The findings of this study suggest that elderly participants tended to perform worse on phonemic fluency tasks than their younger peers in all facets of the tasks. The evidence supports the conclusion that semantic fluency is less difficult than phonemic fluency for healthy elderly adults (Lezak, 2004).

No differences between health groups were found in this study in naming performance. However, the presence of depression and severe anxiety was not assessed which have been shown to cause cognitive impairments similar to those seen in patients with early dementia (Castaneda, Tuulio-Henriksson, Marttunen, Suvisaari, & Lonngvist, 2008; Lezak, 2004, Wright & Persad, 2007). Overall, cognitive impairments associated with depression can be difficult to distinguish from cognitive impairments associated with dementia when only utilizing mental
status exams (Wright & Persad, 2007). Future studies looking into age-related
differences in confrontation naming and verbal fluency tasks should include a
short questionnaire, such as the Beck Depression Inventory and Beck Anxiety
Inventory, which would serve as convenient yet accurate measures of individual
differences that may affect performance on tasks assessing word-finding skills.

_Ceiling Effect._ As expected, responses of the community sample that
scored near or in the normal range on the mental status examine evidenced a
strong ceiling effect on the BNT and PNT. Psychometrically, the PNT and BNT
are negatively skewed and display extreme kurtosis making it difficult to detect
differences at the average to higher average levels of performance. Even the
oldest group, there were participants able to score nearly perfectly on both
measures. This was expected, as the tasks are intended to be a measure of naming
and not of vocabulary.

_Sample Limitations._ Data collection proved to be far more difficult that
originally imagined. Those that refused to take part in the study were asked in a
conversational manner what held them back from participating. Out of the 80
people that refused to engage in the research, 11.25% (n=9) reported that they
were already involved in another study at this time aimed at helping them avoid
memory loss. Concurrent research was being conducted at the senior center by
another major university on life-style changes that may have had a positive effect
on memory. The exclusion criteria of head injury was reported by 2.5% (n=2) of
the potential sample. Finally, 21.25% (n=17) of the population reported that English was not their primary or native language and were unable to participate as this was listed as an exclusion criteria. The remaining 65% (n=52) of possible participants did not provide a reason for declining participation. It could be inferred that some of these individuals may have had similar reasons for not participating thus limiting my sample.

Another major limitation of this study was the low number of male participants (n=9) and the low number of participants with less than a high school education (n=6). Therefore, larger studies are required to validate whether there are potential biases of items on the PNT. In addition, future research should conduct preliminary testing to obtain an estimated Wechsler Verbal Intelligence Quotient to determine if some of the biases of the BNT are present in the PNT. However, the purposes of this study were to gather preliminary data on the validity and reliability of the PNT. In this sample, there were no significant effects of gender or education on PNT performance.

Refinement of Measure. Despite having its limitations, these data have allowed us to refine our 58 item tests into more concise 30 item tests. These data also allowed us to eliminate less desirable images on the PNT. At this stage, it is necessary to continue investigation and repeat the data collection process. Within the sample, there were several items found to discriminate between those who scored within the normal group and those in either the MNCD or DEM group. In
addition, some items on the PNT were found to have acceptable alternative names hypothesized to be influenced by demographic factors although future research would be required to validate demographic differences. The items kept in the 30-item PNT and a list of acceptable alternative names for particular PNT items are found in Appendix F.

For the second round of data collection using the refined 30-item PNT, it would be wise to utilize a community sample with more of a focus on obtaining elderly participants. Next, a much larger community sample must be obtained with a larger number of individuals from each SLUMS derived group. In addition, it would be ideal to have a larger number of participants with more diverse education levels and a more equal sample size of males to better evaluate the effects of demographic variables on PNT performance. Once this sample is collected, all the same analyses that have already been done should be done once more within sample and then again between samples. If the findings of this study are replicated, the diagnostic sensitivity of the PNT in differentiating between normal participants and those with MNCD should be investigated.
REFERENCES


doi:10.1177/107319110100800308

doi:10.1080/13803390802488103


doi:10.1093/geronj/44.5.P139

doi:10.1006/brln.2001.2472


NY: Oxford University Press.


*Psychological Reports*, 92, 468-472.


doi:10.1080/01688639708403773.


Whitfield, K.E., Fillenbaum, G.G., Pieper, C., Albert, M.S., Berkman, L. F.,


## Appendix A- Original Items on the Poreh Naming Test

<table>
<thead>
<tr>
<th>1. Umbrella</th>
<th>23. Microscope</th>
<th>45. Easel</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. Swing</td>
<td>24. Hourglass</td>
<td>46. Spur</td>
</tr>
<tr>
<td>5. Moustache</td>
<td>27. Gorilla</td>
<td>49. Saddle</td>
</tr>
<tr>
<td>8. Accordion</td>
<td>30. Rhino</td>
<td>52. Pacifier</td>
</tr>
<tr>
<td>11. Anchor</td>
<td>33. Slingshot</td>
<td>55. Grenade</td>
</tr>
<tr>
<td>12. Trumpet</td>
<td>34. Whisk</td>
<td>56. Spatula</td>
</tr>
<tr>
<td>13. Cactus</td>
<td>35. Trowel</td>
<td>57. Scarecrow</td>
</tr>
<tr>
<td>14. Helicopter</td>
<td>36. Thimble</td>
<td></td>
</tr>
<tr>
<td>15. Tweezers</td>
<td>37. Antelope</td>
<td></td>
</tr>
<tr>
<td>16. Windmill</td>
<td>38. Tambourine</td>
<td></td>
</tr>
<tr>
<td>18. Iron</td>
<td>40. Gavel</td>
<td></td>
</tr>
<tr>
<td>20. Taj Mahal</td>
<td>42. Corkscrew</td>
<td></td>
</tr>
<tr>
<td>21. Grasshopper</td>
<td>43. Catapult</td>
<td></td>
</tr>
<tr>
<td>22 Crab</td>
<td>44. Dragonfly</td>
<td></td>
</tr>
</tbody>
</table>
Appendix B- The Health and Communication Screening Questionnaire

Please check the box next to any items which pertain to you:

1. Stroke or transient ischemic attack
2. History of seizures
3. Parkinson’s disease
4. Multiple sclerosis
5. Cerebral palsy
6. Huntington’s disease
7. Encephalitis
8. Meningitis
9. Brain surgery
10. Surgery to clear arteries to the brain
11. Diabetes that requires insulin to control
12. Hypertension that is not well controlled
13. Cancer other than skin cancer diagnosed within the past 3 years
14. Shortness of breath while sitting
15. Use of home oxygen
16. Heart attack with changes in memory, ability to talk, or solve problems lasting at least 24 hours afterward.
17. Kidney dialysis
18. Liver disease
19. Unconsciousness for more than one hour other than during surgery
20. Overnight hospitalization because of a head injury
21. Illness causing a permanent decrease in memory or other mental functions
22. Trouble with vision that prevents reading ordinary print even with glasses on
23. Difficulty understanding conversations because of hearing even if wearing a hearing aid
24. Inability to write own name
25. A diagnosed learning disability
26. English is not the native and primary language
Appendix C: Saint Louis University Mental Status Exam

Appendix C

SLUMS EXAMINATION

e-mail: aging@Slums.com Saint Louis Univ. Mental Status Exam

__/1 1. What day of the week is it?  ____/1 3. What State are we in
____/1 2. What is the year?

4.; Please remember these five objects. I will ask you what they are later:
Apple Pen Tie House Car

5. You have $100 and you go to the store and buy a dozen apples for 3 dollar and a tricycle for $20
__/1 How much did you spend? _____/2 How much do you have left?

6. Please name as many animals as you can in one minute:

_____________  ______________  ______________  ______________  ______________
_____________  ______________  ______________  ______________  ______________
_____________  ______________  ______________  ______________  ______________

0 0-4 animals  1 5-9 animals  2 10-14 animals  3 15+ animals

7. What are the five objects I asked you to remember? (1 point for each)
___Apple ___Pen ___Tie ___House ___Car Total score = _____

8. I am going to say a series of numbers and I would like you to say them to me backward
___0 ___1 ___1
87 649 8537 Total score = _____
9. On this page is a clock face. Please put in the hour markers and the time at ten minutes past eleven o'clock
   ___2   Hours marked okay   ___2   Time correct

10. Place an X on the triangle   ___/1      Which Figure is the largest   ___/1

11. I am going to tell you a story. Please listen carefully because afterwards, I'm going to ask you some questions about it

   Jill was a very successful stockbroker. She made a lot of money on the stock market. She then met Jack, a devastatingly handsome man. She married him and had three children. They lived in Chicago. She then stopped work and stayed at home to bring up her children. When they were teenagers, she went back to work. She and Jack lived happily ever after.

   ___2   What is the name of the woman?   ___2   What work did she do?
   ___2   When did she go back to work?   ___2   What State did she live in?

Total score       ______

<table>
<thead>
<tr>
<th>High School Education</th>
<th>Less than High School Education</th>
</tr>
</thead>
<tbody>
<tr>
<td>27-30</td>
<td>Normal</td>
</tr>
<tr>
<td>21-26</td>
<td>MNCD*</td>
</tr>
<tr>
<td>01-20</td>
<td>Dementia</td>
</tr>
<tr>
<td></td>
<td>25-30</td>
</tr>
<tr>
<td></td>
<td>20-24</td>
</tr>
<tr>
<td></td>
<td>01-19</td>
</tr>
</tbody>
</table>

* Mild Neurcoognitive Dementia
Appendix D: Scoring Rules for the Verbal Fluency Tests

Troyer et al. (1997) Method of Verbal Fluency Scoring

Scoring Rules for Clustering and Switching

*Total number of correct words generated.* This performance index was calculated using the sum of all words produced and subtracting both errors and repetitions.

*Mean cluster size.* Cluster size was calculated starting with the second word in a generated cluster. A single word was given a cluster size of 0, two words had a cluster size of 1, three words had a cluster size of 2, and so forth. Errors and repetitions were included. The mean cluster size was computed across the three phonemic trials and across the one semantic trial.

*Number of switches.* This was calculated as the total number of transitions between clusters, including single words, for the three phonemic trials combined and for the one semantic trial. Errors and repetitions were included.

**Phonemic fluency**
Clusters were calculated from consecutively generated words which shared any of the following characteristics:

  *First letters:* words beginning with same first two letters, such as “arm” and “art”
  *Rhymes:* words that rhyme with each other, such as “sand” and “stand”
  *First and last sounds:* words that differ only by a single vowel sound, regardless of the spelling, such as “sat,” “seat,” “soot,” “sight,” and “sought”
  *Homonyms:* words with two or more different spellings, such as “some” and “sum,” as indicated by the subject

**Semantic fluency**
Clusters on the semantic fluency consisted of consecutively generated words belonging to the same subcategories, as specified by Troyer et al. (1997). The listings of examples below for subcategories are not exhaustive.

**Animals**

  *African animals:* aardvark, antelope, buffalo, camel, chameleon, cheetah, chimpanzee, cobra, eland, elephant, gazelle, giraffe, gnu, gorilla, hippopotamus, hyena, impala, jackal, lemur, leopard, lion, manatee, mongoose, monkey, ostrich, panther, rhinoceros, tiger, wildebeest, warthog, zebra

  *Australian animals:* emu, kangaroo, kiwi, opossum, platypus, Tasmanian devil, wallaby, wombat
Arctic/Far North animals: auk, caribou, musk ox, penguin, polar bear, reindeer, seal

Farm animals: chicken, cow, donkey, ferret, goat, horse, mule, pig, sheep, turkey

North America animals: badger, bear, beaver, bobcat, caribou, chipmunk, cougar, deer, elk, fox, moose, mountain lion, puma, rabbit, raccoon, skunk, squirrel, wolf

Water animals: alligator, auk, beaver, crocodile, dolphin, fish, frog, lobster, manatee, muskrat, newt, octopus, otter, oyster, penguin, platypus, salamander, sea lion, seal, shark, toad, turtle, whale

Beasts of burden: camel, donkey, horse, llama, ox

Animals used for their fur: beaver, chinchilla, fox, mink, rabbit

Pets: budgie, canary, cat, dog, gerbil, golden retriever, guinea pig, hamster, parrot, rabbit

Birds: budgie, condor, eagle, finch, kiwi, macaw, parrot, parakeet, pelican, penguin, robin, toucan, woodpecker

Bovine: bison, buffalo, cow, musk ox, yak

Canine: coyote, dog, fox, hyena, jackal, wolf

Deers: antelope, caribou, eland, elk, gazelle, gnu, impala, moose, reindeer, wildebeest

Feline: bobcat, cat, cheetah, cougar, jaguar, leopard, lion, lynx, mountain lion, ocelot, panther, puma, tiger

General Scoring Rules

In cases where two categories overlapped, with some items belonging to both categories, some items belonging exclusively to the first category, and some items belonging exclusively to the second category, the overlapping items were assigned to both categories (e.g. for “dog, cat, tiger, lion,” the first two items were scored as pets, and the last three items were scored as feline. “Cat” was included in both the pet category and the feline category).

In cases where smaller clusters were embedded within a larger cluster, or two categories overlapped, but all items could correctly be assigned to a single category, only the larger, common category was used. For example, for “sly, slit, slim, slam” all begin with “sl,” but an additional cluster was not scored for the last two words which differ only by a vowel sound.
Appendix E: Demographic Survey

1. What is your gender? Please circle one. Male  Female

2. What is your age?  

3. The number of years of education you possess (How many years you formally attended school?)  

4. Please circle your ethnicity. If “other” is selected, please provide how you identify your ethnicity.

Caucasian  African-American  Native American

Hispanic  Asian  Indian  Middle Eastern

Other:  

77
Appendix F: Second Phase of Test Construction – Item Order for Poreh Naming Test

* = 20% or greater difference in accuracy between normal and MNCD groups
Bolded and underlined text = Alternative names provided below graph

<table>
<thead>
<tr>
<th>Item</th>
<th>Accuracy within Normal Group</th>
<th>Accuracy for Total Sample</th>
<th>Mean Latency Time</th>
<th>Difficulty</th>
<th>Original Item Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anchor</td>
<td>100</td>
<td>95.2</td>
<td>2.21</td>
<td>Easy</td>
<td>11</td>
</tr>
<tr>
<td>Helicopter</td>
<td>100</td>
<td>95.2</td>
<td>2.15</td>
<td>Easy</td>
<td>56</td>
</tr>
<tr>
<td>Swing</td>
<td>100</td>
<td>95.2</td>
<td>2.81</td>
<td>Easy</td>
<td>25</td>
</tr>
<tr>
<td>Broccoli</td>
<td>94.7</td>
<td>90.5</td>
<td>3.11</td>
<td>Easy</td>
<td>19</td>
</tr>
<tr>
<td>Slingshot</td>
<td>94.7</td>
<td>95.2</td>
<td>3.1</td>
<td>Easy</td>
<td>46</td>
</tr>
<tr>
<td>Tweezers</td>
<td>100</td>
<td>97.6</td>
<td>2.6</td>
<td>Easy</td>
<td>24</td>
</tr>
<tr>
<td>Globe</td>
<td>100</td>
<td>92.9</td>
<td>2.96</td>
<td>Easy</td>
<td>29</td>
</tr>
<tr>
<td>Microscope</td>
<td>94.7</td>
<td>95.2</td>
<td>4.54</td>
<td>Easy</td>
<td>38</td>
</tr>
<tr>
<td>Grasshopper</td>
<td>94.7</td>
<td>90.5</td>
<td>5.05</td>
<td>Easy</td>
<td>53</td>
</tr>
<tr>
<td>Windmill</td>
<td>94.7</td>
<td>80.9</td>
<td>3.67</td>
<td>Easy</td>
<td>51</td>
</tr>
<tr>
<td>Trumpet</td>
<td>89.5</td>
<td>80.9</td>
<td>4.03</td>
<td>Medium</td>
<td>1</td>
</tr>
<tr>
<td><strong>Podium</strong></td>
<td>100</td>
<td>90.4</td>
<td>4.48</td>
<td>Medium</td>
<td>58</td>
</tr>
<tr>
<td>Corkscrew</td>
<td>94.7</td>
<td>88.1</td>
<td>4.25</td>
<td>Medium</td>
<td>2</td>
</tr>
<tr>
<td>Spur*</td>
<td>94.7</td>
<td>78.6</td>
<td>5.5</td>
<td>Medium</td>
<td>18</td>
</tr>
<tr>
<td>Easel</td>
<td>89.5</td>
<td>69</td>
<td>6.43</td>
<td>Medium</td>
<td>39</td>
</tr>
<tr>
<td>Dragonfly</td>
<td>89.4</td>
<td>66.6</td>
<td>7.29</td>
<td>Medium</td>
<td>50</td>
</tr>
<tr>
<td>Yoyo</td>
<td>94.7</td>
<td>85.7</td>
<td>3.77</td>
<td>Medium</td>
<td>4</td>
</tr>
<tr>
<td>Whisk*</td>
<td>89.4</td>
<td>83.3</td>
<td>5.45</td>
<td>Medium</td>
<td>6</td>
</tr>
<tr>
<td>Tambourine</td>
<td>84.2</td>
<td>71.4</td>
<td>6.47</td>
<td>Medium</td>
<td>14</td>
</tr>
<tr>
<td>Avocado</td>
<td>94.7</td>
<td>83.3</td>
<td>6.48</td>
<td>Medium</td>
<td>48</td>
</tr>
<tr>
<td><strong>Antelope</strong></td>
<td>[<strong>84.2</strong>]</td>
<td>[69]</td>
<td>[6.25]</td>
<td>[Hard]</td>
<td>[43]</td>
</tr>
<tr>
<td><strong>Trowel</strong></td>
<td>[89.4]</td>
<td>[83.33]</td>
<td>[5.41]</td>
<td>[Hard]</td>
<td>[15]</td>
</tr>
<tr>
<td>Palette*</td>
<td>89.4</td>
<td>66.6</td>
<td>5.47</td>
<td>Hard</td>
<td>45</td>
</tr>
<tr>
<td>Hourglass*</td>
<td>84.2</td>
<td>69</td>
<td>6.25</td>
<td>Hard</td>
<td>43</td>
</tr>
<tr>
<td>Taj Mahal</td>
<td>89.4</td>
<td>76.2</td>
<td>6.09</td>
<td>Hard</td>
<td>47</td>
</tr>
<tr>
<td>Totem Pole*</td>
<td>100</td>
<td>83.3</td>
<td>4.52</td>
<td>Hard</td>
<td>7</td>
</tr>
<tr>
<td><strong>Snorkel</strong>*</td>
<td>94.7</td>
<td>76.1</td>
<td>7.74</td>
<td>Hard</td>
<td>41</td>
</tr>
<tr>
<td>Gavel*</td>
<td>89.4</td>
<td>71.4</td>
<td>7.5</td>
<td>Hard</td>
<td>3</td>
</tr>
<tr>
<td>Artichoke*</td>
<td>84.2</td>
<td>66.6</td>
<td>8.67</td>
<td>Hard</td>
<td>52</td>
</tr>
<tr>
<td>Bayonet*</td>
<td>68.4</td>
<td>54.8</td>
<td>8.99</td>
<td>Hard</td>
<td>34</td>
</tr>
</tbody>
</table>
Acceptable Alternative Names for Items

Item 12: Podium- pulpit, lecturn

Item 21: Antelope- gazelle, eeland, impala

Item 22: Trowel- spade

Item 27: Snorkel- scuba mask
Appendix G: Procedure of Administration

1. Greet participant. Give them my name and tell them that I am a graduate student at Cleveland State University who is looking for people to participate in a simple task for psychological assessment. Inform them that testing will take roughly twenty minutes and all responses will be completely confidential and they may discontinue testing at any time during the process. Present informed consent form, with extra copy given to participant for them to keep. Ask the participant if they are over the age of 18, to complete the demographic survey, and to complete the Health and Communication Questionnaire. Ask if they have any questions at this time.

2. Ask participant if English is his or her native language. If the participant reports a health problem or is 65 years of age or older, give the SLUMS. As a part of the battery or as part of the SLUMS, administer the semantic fluency task. The following instructions will be given “I want you to name as many animals as you can, as quickly as possible.” Begin the stop watch when the participant is ready and record the words as they go. Do not count the animal if it is mentioned twice.

3. Perform phonemic verbal fluency test – Say the following instructions “Now I will say a letter of the alphabet. Then I want you to give me as many words that begin with that letter as quickly as you can. For instance,
if I say “B”, you might give me “bad”, “battle”, “bed”, I do not want you to use words that are proper names such as “Boston”, “Bob”, or “Brazil” or numbers. Also, do not use the same word again with a different ending such as “eat” and “eating. You will begin when I say the letter”.

Participants will be asked if they understand the instructions. Tell the participant that you will keep track of the time and the words, and to try their best to continue if they feel they are stuck. The letter will then be provided. Begin stop watch when participant is ready and record the words as they go. Do not count the words if it is mentioned twice.

4. Upon completion, administer the Boston Naming Test and Poreh Naming Test. Alter order of test presentation with each new participant.

   a. Boston Naming Test – administer the 30 item short form using the odd numbered items. Instruct the participant that you are going to show them some picture, and to please tell you what each is. Adhere to the published directions of the BNT, including the 20 second time limit, and using the semantic and phonemic clues as needed. Participants are allowed to give multiple responses to an item within the time limit. If a participant gave multiple responses, they were asked to identify what their final response was to an item. Discontinuation rules were not adhered to and all items of the test were given to all participants.
b. Poreh Naming Test – Enter in personal data for the participant in the first screen. Next, the computer program will give the participant directions stating “I am going to show you some pictures. Please look at each carefully and tell me what it is.” As the participant answers, click on the button stating if they are correct, or if they are stuck click on the “semantic” or “phonemic” buttons, depending on which type of error the participant is making. If they are not sure what the item is, administer the semantic clue. If they seem to know what the item is but cannot come up with the name (e.g. if they state “oh it’s that musical instrument…” for “accordion”), administer the phonemic clue. If you are unable to differentiate immediately, administer the semantic clue. If they are still unsure, note that in the text box and move on the phonemic clue and repeat. Participants are allowed to give multiple responses to an item within the time limit. If a participant gave multiple responses, they were asked to identify what their final response was to an item.

5. Thank the participant for his or her time. Ask if they have any final questions.
Appendix H: Tables of All Correlation Values

(*) = Correlation is significant to the .05 level (2-tailed)  
(**) = Correlation is significant to the .01 level (2-tailed)

<table>
<thead>
<tr>
<th>Correlations With Total Sample</th>
<th>r=</th>
<th>p=</th>
</tr>
</thead>
<tbody>
<tr>
<td>57-item PNT and 30-item PNT</td>
<td>0.984(**)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>57-item PNT Accuracy and Latency</td>
<td>-0.922(**)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>30-item PNT Accuracy and Latency</td>
<td>-0.941(**)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>57-item PNT and BNT</td>
<td>0.922(**)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>30-item PNT and BNT</td>
<td>0.917(**)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>57-item PNT and SLUMS</td>
<td>0.875(**)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>30-item PNT and SLUMS</td>
<td>0.853(**)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>BNT &amp; SLUMS</td>
<td>0.843(**)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>57-item PNT and Semantic Total Output</td>
<td>0.603(**)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>30-item PNT and Semantic Total Output</td>
<td>0.575(**)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>BNT and Semantic Total Output</td>
<td>0.542(**)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>57-item PNT and Total Semantic Clusters</td>
<td>0.378(*)</td>
<td>0.014</td>
</tr>
<tr>
<td>30-item PNT and Total Semantic Clusters</td>
<td>0.376(*)</td>
<td>0.014</td>
</tr>
<tr>
<td>BNT and Total Semantic Clusters</td>
<td>0.237</td>
<td>0.13</td>
</tr>
<tr>
<td>57-item PNT and Semantic Average Cluster Size</td>
<td>-0.379(*)</td>
<td>0.013</td>
</tr>
<tr>
<td>30-item PNT and Semantic Average Cluster Size</td>
<td>-0.353(*)</td>
<td>0.024</td>
</tr>
<tr>
<td>BNT and Semantic Average Cluster Size</td>
<td>-0.494(**)</td>
<td>0.001</td>
</tr>
<tr>
<td>57-item PNT and Semantic Switching</td>
<td>0.419(*)</td>
<td>0.006</td>
</tr>
<tr>
<td>30-item PNT and Semantic Switching</td>
<td>0.391(*)</td>
<td>0.01</td>
</tr>
<tr>
<td>BNT and Semantic Switching</td>
<td>0.465(**)</td>
<td>0.002</td>
</tr>
<tr>
<td>57-item PNT and Phonemic Total Output</td>
<td>0.498(**)</td>
<td>0.001</td>
</tr>
<tr>
<td>30-item PNT and Phonemic Total Output</td>
<td>0.502(**)</td>
<td>0.001</td>
</tr>
<tr>
<td>BNT and Phonemic Total Output</td>
<td>0.495(**)</td>
<td>0.001</td>
</tr>
<tr>
<td>57-item PNT and Total Phonemic Clusters</td>
<td>0.421(**)</td>
<td>0.006</td>
</tr>
<tr>
<td>30-item PNT and Total Phonemic Clusters</td>
<td>0.402(**)</td>
<td>0.008</td>
</tr>
<tr>
<td>BNT and Total Phonemic Clusters</td>
<td>0.411(**)</td>
<td>0.007</td>
</tr>
<tr>
<td>57-item PNT and Average Phonemic Cluster Size</td>
<td>0.073</td>
<td>0.646</td>
</tr>
<tr>
<td>30-item PNT and Average Phonemic Cluster Size</td>
<td>-0.01</td>
<td>0.952</td>
</tr>
<tr>
<td>BNT and Average Phonemic Cluster Size</td>
<td>0.053</td>
<td>0.738</td>
</tr>
<tr>
<td>57-item PNT and Phonemic Switches</td>
<td>0.437(**)</td>
<td>0.004</td>
</tr>
<tr>
<td>30-item PNT and Phonemic Switches</td>
<td>0.461(**)</td>
<td>0.002</td>
</tr>
<tr>
<td>BNT and Phonemic Switches</td>
<td>0.490(**)</td>
<td>0.001</td>
</tr>
</tbody>
</table>
### Test-Retest Correlations

<table>
<thead>
<tr>
<th>Test-Retest Correlations</th>
<th>r</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>57-item PNT: Time 1 &amp; 2</td>
<td>0.990(**)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>BNT: Time 1 &amp; 2</td>
<td>0.989(**)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>57-item PNT retest and BNT retest</td>
<td>0.985(**)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Semantic total output: Time 1 &amp; 2</td>
<td>0.751(*)</td>
<td>0.032</td>
</tr>
<tr>
<td>Phonemic total output: Time 1 &amp; 2</td>
<td>0.893(**)</td>
<td>0.003</td>
</tr>
<tr>
<td>Semantic total clusters: Time 1 &amp; 2</td>
<td>0.827(*)</td>
<td>0.011</td>
</tr>
<tr>
<td>Phonemic total clusters: Time 1 &amp; 2</td>
<td>0.773(*)</td>
<td>0.024</td>
</tr>
<tr>
<td>Semantic average cluster size: Time 1 &amp; 2</td>
<td>0.734(*)</td>
<td>0.038</td>
</tr>
<tr>
<td>Phonemic average cluster size: Time 1 &amp; 2</td>
<td>0.44</td>
<td>0.275</td>
</tr>
<tr>
<td>Semantic switches: Time 1 &amp; 2</td>
<td>0.434</td>
<td>0.283</td>
</tr>
<tr>
<td>Phonemic switches: Time 1 &amp; 2</td>
<td>0.868(**)</td>
<td>0.005</td>
</tr>
</tbody>
</table>

### Partial Correlation Controlling for SLUMS

<table>
<thead>
<tr>
<th>Partial Correlation Controlling for SLUMS</th>
<th>r</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>57-item PNT and Age</td>
<td>0.011</td>
<td>0.948</td>
</tr>
<tr>
<td>30-item PNT and Age</td>
<td>0.029</td>
<td>0.856</td>
</tr>
<tr>
<td>57-item PNT and Education</td>
<td>0.098</td>
<td>0.542</td>
</tr>
<tr>
<td>30-item PNT and Education</td>
<td>0.196</td>
<td>0.218</td>
</tr>
<tr>
<td>57-item PNT Latency and Age</td>
<td>0.212</td>
<td>0.184</td>
</tr>
<tr>
<td>30-item PNT Latency and Age</td>
<td>0.175</td>
<td>0.274</td>
</tr>
<tr>
<td>Age and Phonemic Total Output</td>
<td>-0.371(*)</td>
<td>0.017</td>
</tr>
<tr>
<td>Age and Semantic Total Output</td>
<td>-0.304</td>
<td>0.054</td>
</tr>
<tr>
<td>Age and Phonemic Switches</td>
<td>-0.278</td>
<td>0.078</td>
</tr>
</tbody>
</table>
Appendix I: IRB Approval Form

Memorandum

To: Amir Poreh
   Psychology

From: Barbara Bryant
   IRB Recording Secretary

Date: January 3, 2012
Re: Results of IRB Review of your project number: #29407-POH-HS
   Co-Investigator: Oron Bissan
   Entitled: Normative data collection and comparison of performance on the Poreh Naming Test to the Boston Naming Test

The IRB has reviewed and approved your application for the above named project, under the category noted below. Approval for use of human subjects in this research is for one year from today. If your study extends beyond this approval period, you must contact this office to initiate an annual review of this research.

By accepting this decision, you agree to notify the IRB of: (1) any additions to or changes in procedures for your study that modify the subjects' risk in any way; and (2) any events that affect the safety or well-being of subjects. Notify the IRB of any revisions to the protocol, including the addition of researchers, prior to implementation.

Thank you for your efforts to maintain compliance with the federal regulations for the protection of human subjects.

Approval Category: Exempt Status: Project is exempt from further review under CFR 46.101 b(2)

Date: January 3, 2012

cc: Project file