國立臺灣大學理學院心理學研究所

碩士論文

Graduate Institute of Psychology

College of Science

National Taiwan University

Master Thesis

臺灣健康成人之視覺命名測驗及物體命名測驗常模研究

A Normative Study on the Visual Naming Test and the Object Naming Test in Healthy Taiwanese Adults

吳瑋庭

Wei-Ting Wu

指導教授：花茂棽 博士

Advisor: Mau-Sun Hua, Ph. D.

中華民國 105 年 6 月

June, 2016

doi:10.6342/NTU201600919

摘要

背景：命名障礙常見於中風、癲癇及神經退化性疾患患者。本研究之命名測驗組包含視覺命名測驗及物體命名測驗，是一套用來評估口語表達功能的神經心理測驗工具，亦作為觀察患者語言功能有無變化之工具；且同時具有測量時間短、指導語簡單以及包含多種類型與難度題項之優點，然而目前台灣缺乏此測驗組之常模資料。**目的：**本研究主要目的為建立台灣健康成年人命名測驗組之常模，主要目標為：一、探討人口學變項對命名測驗表現的影響；二、檢驗命名測驗組的心理計量特性；三、本常模於台灣成年人之適用性。**方法：**本研究共收集台灣 322 位 16 至 90 歲、教育程度 0 至 18 年之台灣成年人於命名測驗組之常模資料，同時分層隨機取 30 位受試者進行魏氏成人智力量表第三版之詞彙及類同分測驗等資料以檢驗其效度。此外，比較臨床病人與健康成年人在本測驗之表現以檢驗效度。部分樣本同時收集再測資料以進行信度檢驗。**結果：**教育程度為主要影響測驗表現的人口學變項，年齡及性別則無顯著影響。此外，研究顯示本測驗組具有良好的建構效度及中等程度的再測信度。而其常模資料具有良好代表性、新進性與適切性。**結論：**本研究檢驗人口學變項對命名測驗組表現之影響，且提供人口學變項效果校正後之常模資料及切截點。儘管命名測驗組之再測信度不如預期，結果仍顯示其具有良好的效度。未來應重複驗證臺灣樣本於命名測驗組之再測信度，並建構質性計分系統，以利於臨床上測量口語表達能力之變化、偵測失語症症狀之出現及評估神經認知復健之成效。

關鍵字：視覺命名測驗、物體命名測驗、神經心理測驗、口語表達測驗、常模研究、人口學變項、信效度、校正分數

doi:10.6342/NTU201600919

A Normative Study on the Visual Naming Test and the Object Naming Test in Healthy Taiwanese Adults

Wei-Ting Wu

Abstract

Background: Naming impairment is mostly found in patients with stroke, epilepsy and neurodegenerative diseases. The Battery of Naming Tests, which includes the Visual Naming Test of the Multilingual Aphasia Examination and the Object Naming Test of the Neurosensory Center Comprehensive Examination for Aphasia, is a good task for measuring both oral expression performance and core linguistic functions. In addition, this battery has several advantages, such as a short administration time, simple instructions, and items of different categories and levels. However, there is currently a lack of normative data for the Battery of Naming Tests in the Taiwanese population. **Objective**: The present study aimed to obtain normative data for the Battery of Naming Tests. The specific objectives were as follows: (1) to assess the influence of demographic variables on test performance, (2) to establish psychometric properties, and (3) to determine norm appropriacy. **Methods**: Participants (N = 322) were recruited through stratified sampling by current age (ranging from 16 to 90 years old), education (ranging from 0 to 18 years), and area of residence. Thirty participants also completed the Vocabulary and Similarities subtests of the WAIS-III for

doi:10.6342/NTU201600919

investigation of the validity, and a clinical patient group was also recruited for verification of the validity. Test-retest reliability was derived from a subgroup of thirty participants. **Results**: Education significantly influenced performance of naming tests. The results also showed that the Battery of Naming Tests had sound construct validity and moderate test-retest reliability. The normative data showed good representativeness, recency, and relevance. **Conclusion**: The present study determined the influence of demographic effects on performance on the Battery of Naming Tests, established appropriate normative data, and developed a reference table for percentile ranks and a cut-off point. Despite the moderate test-retest reliability, this study verified the adequate validity of the Battery of Naming Tests. Further investigations on the test-retest reliability and the qualitative scoring system appear necessary in order to detect changes in oral expression and aphasic symptoms, and monitor the outcome of neurocognitive rehabilitation

Keywords: the Visual Naming Test, the Object Naming Test, neuropsychological assessment, oral expression tests, normative data, demographic variables, reliability, validity, adjusted score

doi:10.6342/NTU201600919

Table of Contents

doi:10.6342/NTU201600919

Introduction

From infancy, we learn to say the names of people and objects in order to convey our needs. As we grow older, the ability to express oneself verbally is also crucial for performance in the academic, occupational, and social domains (Hamberger, 2015). Oral expression belongs to the four basic types of language performance, along with aural comprehension, reading and writing (Benton, 1968). From a greeting to a delicate speech, oral expression is as ubiquitous and sophisticated as other aspects of language.

Among the word-level expression, naming, can be a sampling of one's verbal expression performance. Naming is thought to contain all the stages of word production from conceptualization and selection of a lexical item to phonological encoding and preparation for articulation (Levelt, Roelofs, & Meyer, 1999). As a form of verbal output, naming requires retrieval of words from a verbal knowledge store. This retrieval process is essentially the ability to pull out the correct word at will (M. Lezak, Howieson, Bigler, & Tranel, 2012; Rohrer et al., 2008). It involves three main phases: perceptual analysis of the stimulus, name activation, and response generation (Carla J. Johnson, 1996; DeLeon et al., 2007; Paivio, Clark, Digdon, & Bons, 1989). In the first phase, perceptual analysis of the stimulus, the object is visually identified and recognized. In the second phase, name activation, the lexical-semantic representation of the object is activated in memory. The third phase, response generation, involves

doi:10.6342/NTU201600919

phonological activation and subsequent motor execution of the word (i.e., planning and implementing the articulation of the word). In summary, naming depends on not only linguistic functions, but also working memory and executive control (Duffau, Moritz-Gasser, & Mandonnet, 2014).

In the first visual recognition phase, the inferior longitudinal fascicle, located within the basal occipito-temporal region, links the visual cortex with the postero-inferior temporal area (Duffau et al., 2014). It is well documented that the conceptual representations making up word meanings reside in widespread areas of the perisylvian regions, including the anterior inferior and middle temporal cortex and angular gyrus (Roelofs, 2014). The meaning mapping onto articulation relies primarily on a ventral pathway underpinned by the left uncinate fasciculus and the tracts passing through the left extreme capsule (Ueno, Saito, Rogers, & Ralph, 2011). Another part of the arcuate fasciculus, connecting the left middle temporal gyrus (MTG) and the Broca's area, wherein lexical-semantic representations are mapped onto output phonemes, is referred to as the lexical-semantic or MTG pathway, which is crucial for conceptually driven speech production in conversation and picture naming (Glasser & Rilling, 2008; Schwartz, Faseyitan, Kim, & Coslett, 2012).

Naming impairment, or anomia, is mostly seen in patients with language dysfunction, due to stroke, epilepsy, and neurodegenerative diseases, such as

doi:10.6342/NTU201600919

Huntington's disease, Alzheimer's disease and the semantic variant of fronto-temporal dementia (Berthier & Pulvermüller, 2011; Mesulam et al., 2009).

On one short naming screening test, forty-five percent of the stroke patients scored below the cut-off point (Riepe, Riss, Bittner, & Huber, 2003). About a third of stroke survivors have aphasia in the early stage post stroke, and fifteen percent remain aphasic (Engelter et al., 2006). Naming difficulty is a common deficit in different types of aphasia, and it is a residual symptom as well. Thus, a naming test should be the first step in aphasia assessment (Nicholas, Obler, Albert, & Helm-Estabrooks, 1985).

Naming impairment may result from dysfunction of visual perception, semantic processing, or word retrieval. The errors of Huntington's disease patients are due to the inability to retrieve information with relatively intact semantic knowledge (Azambuja, 2012), while Alzheimer's disease (AD) patients show significantly more semantic superordinate errors (i.e., when patients use the general class to represent the target item) (Lin et al., 2014). Mildly demented patients with Alzheimer's disease often have impaired naming performance, and their naming ability declined during the progression of the disease (Bayles & Tomoeda, 1983). Studies suggest that semantic degradation is the major contributing factor in the naming difficulty in AD patients (Lin et al., 2014). Same in AD patients, anomia in the semantic variant of fronto-temporal patients is due to a breakdown in both verbal and nonverbal semantic knowledge about objects and

doi:10.6342/NTU201600919

words (Murre, Graham, & Hodges, 2001). Naming requires the integrity of semantic concepts, and dysfunction may be a marker of primary semantic memory impairment rather than otherwise cognitive decline (Verma & Howard, 2012). Aside from the affected temporal lobe, the reduced integrity of the arcuate fasciculus is also related to the poor naming performance of temporal lobe epilepsy patients (Hamberger, 2015).

Monitoring naming performance at different times is used as a tool to assess the degree of deterioration in neurodegenerative disorders and the degree of recovery following brain lesions (Franzen, 2000). Furthermore, naming is a predominant neuropsychological test for aphasia assessment. Therefore, naming tests with good reliability and validity are crucial in neuropsychological assessment.

Many kinds of stimuli are used as the naming materials (e.g., real objects, photographs, and pictures/drawings). Research has shown that picture/drawing materials can be too difficult for the low-educated due to their lack of two-dimensional representation learning (Reis, Guerreiro, & Castro-Caldas, 1994). Therefore, real objects can be used to reduce discrepancies due to education differences. In order to evaluate patients with anomia of different education levels, Hua, Chang, and Chen (1997) adopted (a) the Visual Naming subtest of the Multilingual Aphasia Examination (MAE; Benton & Hamsher, 1978), and (b) the Object Naming subtest of the Neurosensory Center Comprehensive Examination for Aphasia (NCCEA; Spreen &

doi:10.6342/NTU201600919

Benton, 1969) to composite the Battery of Naming Tests.

The battery of Naming Tests has three clear benefits. One is its short administration time, which is suitable for patients with attention deficits and poor physical condition. Another is its simple and clear instructions, which facilitate the cooperation of the participants in responding. The other benefit is its ability to probe into different categories and levels of verbal knowledge. Emery (2000) hypothesized that language forms learned last deteriorate first, suggesting that naming ability is affected by both frequency of item use and age of acquisition. Therefore, it is essential to have items of different levels and categories.

Little research has been conducted on the psychometric properties of the MAE and NCCEA. The test-retest reliability of the Boston Naming Test is known to be good varying from 2 weeks to 11 months (r = .91- .92) (Flanagan & Jackson, 1997). However, the inter-rater reliability data of the MAE and NCCEA have not been reported (Harry & Crowe, 2014). A validity study of the Visual Naming Test showed a significant correlation to the Verbal Comprehension index of the WAIS-R, but a negligible correlation to the Perceptual Organization index (Axelrod, Ricker, & Cherry, 1994). The criterion-related validity of the Boston Naming Test revealed excellent agreement (r = .86) (Axelrod et al., 1994). Jordan, Ozanne, and Murdoch (1988) used the NCCEA to assess the language function of 20 children and adolescents with traumatic brain

doi:10.6342/NTU201600919

injury (TBI). They found that participants with moderate to severe TBI had impaired naming performance two years post injury.

With regard to demographic factors, studies of aphasia test batteries, such as the Boston Diagnostic Aphasia Examination (BDAE; Goodglass & Kaplan, 1983) have also been conducted. Naming function is generally well preserved in aging with only subtle decline in the 7^{th} and 8^{th} decades of age (Zec, Markwell, Burkett, & Larsen, 2005). In the study by Zec and co-workers (2005), the age effect was significantly, but the interaction of age and education appeared to be a better predictor of naming performance (Welch, Doineau, Johnson, & King, 1996). More variability was found in higher age groups and in those with less education. Naming performance remained stable in the higher education group (above 12^{th} grade) until 80 years, whereas those participants who had not completed high school demonstrated a decrement in naming ability at 70 years. Education was found to account for 13% of the variance of the scores of the Visual Naming Test (Axelrod et al., 1994). Research has also shown that age-related difficulties are associated with brain atrophy in linguistic areas and that white matter integrity changes in regions related to language production (Rolheiser, Stamatakis, & Tyler, 2011; Tyler, Stamatakis, Post, Randall, & Marslen-Wilson, 2005).

Findings on the gender effect on naming performance have been inconsistent. Studies have demonstrated the presence of a gender-by-category interaction in naming

doi:10.6342/NTU201600919

abilities (Laws, 1999). Particularly, males performed better on naming non-living things; females, on naming living-things. Gender-related familiarity differences might account for such a phenomenon.

The naming subtests of the MAE and the NCCEA are standardized tests with clinical utility. Furthermore, the naming subtests have several advantages, including convenience of use, short administration time, and ease of administration. However, no representative Taiwanese normative data have been generated for the Battery of Naming Tests, which is composed of the subtests of the MAE and the NCCEA. There was also limited research conducted on the psychometric properties. Therefore, the present study aimed to construct a representative Taiwanese norm for the Battery of Naming Tests and to examine the demographic effects and psychometric properties.

doi:10.6342/NTU201600919

Method

Participants

The 322 volunteers (138 males, 184 females) were recruited for participation in this study from the communities in Taiwan. According to data from the 2013 Population Statistical Yearbook (Department of Household Registration, 2014), participants were stratified based on age, education level and geographical region (see Table 1).

Inclusion/ exclusion criteria. In order to participate in the study, individuals must meet the following criteria: 1) above 16 years old; 2) speaking Chinese; 3) capable of living independently and engaging activities in society. Exclusion criteria are: 1) no current or past history of alcohol or drug abuse, brain injury, stroke, dementia, endocrine disorders (particularly of the thyroid gland), or any other systemic diseases, neurological disorder, or psychiatric illness as determined by screening interview; 2) no current cognitive impairment as determined by the Mini-Mental State Examination (MMSE) score lower than 24/30 or 18/30 for individuals with less than 2 years of education (Wang, 2007). However, in order to avoid recruiting “hyper-normal” subjects, individuals with mild hypertension or other systemic problems under medical control were not excluded.

Measurements

The Battery of Naming Tests (Hua et al., 1997) was adapted from (a) the Visual

doi:10.6342/NTU201600919

Naming subtest of the Multilingual Aphasia Examination (MAE; Benton & Hamsher, 1978), and (b) the Object Naming subtest of the NCCEA (Spreen & Benton, 1969). The scholars modified items thimble and tweezers with pin and chopsticks respectively of the Object Naming subtest for cultural differences.

The Visual Naming Test (VNT) of the MAE was a validated measure of naming performance. It was composed of 30 items presented on 10 pictures. Within each picture, items were presented in order of increasing difficulty (e.g., letter, stamp and postmark). Each correct name was scored 2 points. The total maximum score was 60 points.

The Object Naming Test (ONT) of the NCCEA was also a test which required subjects to say the names of the presented items. The test consisted of 16 objects, given in two sections. If subject succeed in naming all objects of the first section, the full points was given. If subject failed in anyone of object naming in the first section, object pointing of the first section and proceeding to the naming of the second section were suggested. Each correct name was scored 1 point. The total maximum score was 16 points.

Neuropsychological assessments

Although visuospatial and executive function are involved in the process of naming, they were not the primary components compare to verbal abilities. Thus, tests related to visuo-perceptual and executive functions were not included as reference tests.

doi:10.6342/NTU201600919

Verbal abilities were assessed by the Vocabulary and Similarity subtests of the Wechsler Adult Intelligence Scale-Third Edition (WAIS-III) Taiwan Version (H. Y. Chen & Chen, 2002).

Procedure

The study was approved by the Research Ethics Committee of National Taiwan University (NTU-REC No.: 201401HS026) (See Appendix C). Information regarding participants' age, education, medical history, current health status, and medication regimen was obtained through a semi-structured interview. The MMSE was administered to participants as a means of screening cognitive functioning. The battery of Naming Tests was presented to all participants. Thirty participants, selected by stratified random sampling, additionally completed other neuropsychological tests for validity examination. A subgroup of 30 participants, who were selected by convenience sampling with a match of demographic variables, was retested in at least 2 months in order to establish test-retest reliability and determine potential practice effects.

Statistical Analysis

All statistical tests were performed in SPSS version 20.0. The demographic characteristics of samples were presented by descriptive statistics. The Kolmogorov-Smirnov test was used to examine if the raw scores were normally distributed, and the test results revealed negative for all raw scores. Thus, non-parametric statistical

doi:10.6342/NTU201600919

procedures were applied as follows. Mann-Whitney U test and Kruskal-Wallis one-way analysis of variance by ranks tests was used to compare all the scores among different groups. Dunn's test was employed to specify the pairwise comparison of groups.

Scatter plots and curve estimates were used to determine the most appropriate regression model to assess the relationship between the demographic factors and the raw scores. Multiple regression analysis and stepwise regression were used to examine the potential contribution of demographic variables to the scores. To establish normative data, the regression coefficients were used to derive corrected equations, and then to generate adjusted scores and their percentile ranks.

With regard to the psychometric properties of the Battery of Naming Tests, Spearman's rho was employed to calculate the test-retest reliability. Furthermore, Wilcoxon signed rank test was performed to clarify the practice effect. To examine construct validity, correlation analyses were conducted to investigate the relationship between the scores and the neuropsychological tests mentioned in an earlier section, and the effect of participant's education on the scores was also analyzed. Furthermore, patient group and demographic-compatible normal controls were compared by nonparametric statistical procedures to provide further evidence of construct validity.

doi:10.6342/NTU201600919

Results

Demographic Characteristics

The demographic characteristics of the participants are presented in Table 2. The mean age of the sample was 45.21 years (ranging from 16 to 86; *SD* = 17.94). The mean year of education was 12.91 years (ranging from 0 to 19; *SD* = 3.54). There was no significant difference in age (t = -.927, p = .355) and level of education (t = .923, p = .357) between male and female participants.

Demographic Effects

Descriptive statistics of raw scores by age, gender, and education level are presented in Table 3. Generally, mean raw scores were higher in younger age groups and in higher education levels. Due to the ceiling effect, however, these trends were vanished in the Object Naming Test. Mann-Whitney U test showed significant difference between males and females in scores of the Visual Naming Test (Z = -2.16, p = .03), but not in the Object Naming Test (Z = -.816, p = .41).

Spearman's rank correlation coefficients between raw scores and demographic variables (i.e., gender, age and education) are presented in Table 4. The scores of the Visual Naming Test were significantly correlated with each demographic variables. Age has weak correlations (r_s = -.264, $p < .01$), while education has moderate correlation (r_s = .565, $p < .01$). Gender only had very weak correlations (r_s = .120, $p < .05$). On the

doi:10.6342/NTU201600919

contrary, the scores of the Object Naming Test was not correlated with any demographic variables.

To determine the contribution of the demographic variables to the test scores, scatter plots and curve estimates were used to access the most appropriate type of regression functions, and then test scores were entered into multiple regression analyses (see Table 5). It revealed that the relations between test scores and demographic variables were linear. The result of regression analyses are presented in Table 6. Education accounted for the most proposition of variance of the test scores (36% of variance), followed by gender (1.2% of variance).

Results of between-group comparison calculated by Kruskal-Wallis one-way analysis of variance by ranks was presented in Table 7. Significant differences among different education levels in the Visual Naming Test were found. Post-hoc pairwise comparisons showed that above 13 years group had significant higher scores than the other education groups (the 0-6, 7-9, and 10-12 year groups); moreover, the 10-12 year group performed better than the 0-6 year group (see Table 8).

Psychometric Properties

Reliability. The demographic characteristics of the sample for test-retest reliability are presented in Table 9. The average interval time was 93.33 days (ranging from 61 to 141 days). As shown in Table 10, the test-retest reliability of the Visual

doi:10.6342/NTU201600919

Naming Test was moderate (r_s = .60). Meanwhile, the test-retest reliability of the Object Naming Test was non-significant due to prominent ceiling effect (see Table 15). In addition, no significant practice effect was found in the test scores by Wilcoxon signed rank test.

Construct validity. The demographic characteristics of the samples for validity were presented in Table 11, and the correlation coefficients were listed in Table 12. The scores of the Visual Naming Test had strong correlations with the scores of the Vocabulary and the Similarities subtests of the WAIS-III. It also had moderate correlation with the Token Test, the Sentence Repetition Test, and the Reading Comprehension Test, and mild correlation with the Aural Comprehension Test. However, the Object Naming Test scores were not significantly correlated with any measures.

Both age and education effects were also adopted to verify the construct validity of the Battery of Naming Tests. The results of ANOVA test and regression analyses revealed noticeable education effects on the scores of the Visual Naming Test (see Table 5, 6, 7 and Figure 1-1). However, the results showed no significant age effects on all test scores (see Table 5, 6, 7).

The demographic characteristics of the clinical patient groups and the healthy control group are listed in Tables 16 and 20. The 5^{th} percentile score of the healthy

doi:10.6342/NTU201600919

control group was used as the cut-off point, and the percentage of participants below the cut-off point are shown in Table 17. On the Visual Naming Test, 3.51% of the participants in the healthy control group, 20.00% in the very mild AD group, and 30.00% in the mild AD group had test scores below the cut-off point. On the Object Naming Test, 3.51% of participants in the healthy control group, 11.11% in the very mild AD group, and 42.86% in the mild AD group had test scores below the cut-off point. The data analysis with the Kruskal-Wallis one-way analysis of variance by ranks revealed significant differences in the test scores on the Visual Naming Test and on the Object Naming Test among patient groups with different AD severities (see Table 18). Dunn's test for the post-hoc pairwise comparisons showed that the scores of the healthy control group were significantly higher than those of both patient groups (the very mild and mild AD groups) on the Visual Naming Test, and significantly higher than those of the mild AD group on the Object Naming Test (see Table 19).

Furthermore, the two naming test scores of all patients with global aphasia, Wernicke's aphasia, conduction aphasia and transcortical sensory aphasia were below the cut-off point (see Table 21).

Normative Data

In the present study, regression coefficients were used to rule-out demographic effects on the performances of the naming tests through establishing a demographic-

doi:10.6342/NTU201600919

correlated equation for raw scores. Corrected equations are listed in Table 13. Scores of the Object Naming Test needed no corrections since there was no gender, age and education effects on the performance. All adjusted scores were ranked by percentile, and a referential table was presented in Table 14. The fifth percentile as cut-off point was recommended by M. D. Lezak (2004): if a participant's score is in the 5^{th} or less percentile, the performance is regarded as defective.

doi:10.6342/NTU201600919

Discussion

The present study examined the demographic effects and psychometric properties of the Battery of Naming Tests. This study also aimed to establish demographically corrected normative data based on a large-scale representative Taiwanese sample for future research and clinical practice.

Demographic Effects on the Battery of Naming Tests

Effect of education. The results of the present study revealed that education had a prominent effect on the performance of the Visual Naming Test (VNT). This finding is consistent with previous research (Axelrod et al., 1994; Schum & Sivan, 1997). In the multiple regression analyses, education accounted for 36% of the variance in VNT scores. However, there was no education effect on the performance of the Object Naming Test, due to the prominent ceiling effect (see Table 15). Hence, the Object Naming Test could be used to reduce discrepancies due to variation in level of education, especially for low-educated patients.

In consideration of the time constraints of a clinical setting, some suggestions for test selection are provided based on the present findings of a differential education effect on the naming tests. As mentioned previously, the Visual Naming Test requires two-dimensional representation, which is prohibitively difficult for the low-educated (Reis, Guerreiro, & Castro-Caldas, 1994). It may also increase the possibility of refusal

doi:10.6342/NTU201600919

to complete the test. Therefore, the clinicians should choose to administer the Object Naming Test first to patients with a lower education level. As for patients with a higher educational level, the ceiling of the Object Naming Test diminishes its usefulness to detect their naming impairment (Spreen & Risser, 2002). Thus it is recommended that the Visual Naming Test, having items of different levels and categories, be administered first to patients with a higher education level.

Effect of age. Age had a significant effect in the present study, but it was less influential than the education effect on performance of the Visual Naming Test. Age had weak correlation with test scores (r_s = -.264). Nonetheless, the age effect disappeared in the explained variance calculation. This finding is consistent with that of a previous report that naming performance is stable from the early adult years to the ninth decade (Schum & Sivan, 1997).

Effect of gender. An unexpected gender effect was found in the Visual Naming Test. Even though male participants performed significantly better than female participants, the correlation between gender and score was very weak (r_s = .120), and the mean difference between males and females was only 1.52 points.

Most previous studies have reported conflicting results of the gender effect on naming tests because of the gender-by-category interaction. Ample evidence suggests that males outperform females on naming non-living things, while females outperform

doi:10.6342/NTU201600919

males on naming living things (Laws, 1999). In addition, the Visual Naming Test has four more items on naming non-living things than on naming living ones. However, the gender effect of the present study was quite small. The absence of a significant gender effect in the present study was possibly due to cultural factors. There seems to be no such gender-related familiarity difference in Taiwanese culture. Future investigation on this issue is thus necessary.

Psychometric Properties

Reliability. As Morgan, Gliner, and Harmon (2006) addressed, a reliable measurement is expected to have a coefficient between .7 and 1.0. In the present study, the results showed that the test-retest reliability coefficient of the Visual Naming Test (VNT) was moderate (r_s = .60) with an average interval of 93 days. After the raw data were carefully examined and an outlier (a participant with an 8-points discrepancy between two tests) was eliminated, the test-retest reliability coefficient was close to high (r_s = .69).

The reliability coefficient of the Object Naming Test could not be calculated but appeared stable because most participants earned the maximum scores in both test sessions (see Table 15). The Neurosensory Center Comprehensive Examination for Aphasia was used to derive patient groups reflecting both type and magnitude of disorder, and several subtests have ceiling effects (Crockett, 1977). Therefore, the

doi:10.6342/NTU201600919

ceiling effect of the Object Naming Test in the present study was expected. Furthermore, considering the lack of significant differences between the test-retest scores, it appeared that the practice effect of the naming tests was not remarkable.

Previous research revealed that naming tests had high test-retest reliability (*r* = .91- .92) (Flanagan & Jackson, 1997). The variation in reliability coefficients between the previous and the present studies might be due to differences in the retest intervals and in the characteristics of the participants. First, the retest intervals in Flanagan and Jackson (1997) were 7 to 17 days, which is much shorter than the interval in the present study. In addition, participants in that study were sampled from a more homogenous group with a smaller standard deviation on age (7.4) and education (1.75). In our study, we recruited participants with a wider range of ages and levels of education (see Table 9).

Considering the common clinical practice in Taiwan, the follow-up interval in this study was set at two to four months. However, the test-reliability was lower in the present study than that in the study by Flanagan and Jackson (1997). Lower test-retest reliability increases the standard error of the measure and also decreases the stability of the measure. To achieve higher test-retest reliability, Morgan and co-workers (2006) suggested that a shorter test-retest interval could prevent potential factors from interfering with the participant's test performance. Thus clinically, a shorter follow-up

doi:10.6342/NTU201600919

interval appears to be imperative in naming performance assessment. However, repeating such assessments with a shorter interval might also produce a practice effect that could cause overestimation. Therefore, further research is needed to determine a proper test-retest interval so as to promote test stability and minimize the practice effect.

In regard to the characteristics of the participants, our study used a large scale representative sample with a wide range of ages and levels of education (see Table 2). Meanwhile, the participants' standard deviations in age and education in the study by Flanagan and Jackson (1997) were only 7.4 and 1.74 respectively. Such figures indicate a possible sampling bias in that study. Clarification of this issue awaits further study.

Construct validity. Our study revealed that the score of the Visual Naming Test was significantly correlated with the scores of the Vocabulary and the Similarities subtests of the WAIS-III. This strong correlation may indicate that the naming tests are related to the activation of lexical-semantic knowledge, for the Vocabulary and Similarities subtests involve verbal concept formation and verbal expression, both of which are based on semantic knowledge (Hawkins et al., 1993).

In a study by Hua et al. (1997), the Visual Naming Test was related to the verbal comprehension factors of the Token Test, Sentence Repetition, and Reading and Aural Comprehension Test. In these tests, participants need to follow verbal instructions (words, phrases, and sentences), retrieve lexical-semantic knowledge with auditory or

doi:10.6342/NTU201600919

visual cues, and use immediate memory (Hua et al., 1997). Our study also found moderate correlation with the Token Test, the Sentence Repetition Test, and the Reading Comprehension Test, and mild correlation with the Aural Comprehension Test.

Construct validity: education effect. The items on the Visual Naming Test were found to have varying levels of difficulty. Some items, such as "hexagon" and "peninsula", involve geometry and geography. Individuals with higher education are generally more familiar with the difficult test items (Hawkins et al., 1993). As Axelrod et al. (1994) and Benton, Hamsher, & Sivan (1994) proposed, education has an impact on naming performance. The prominent education effects on the performance scores of the Visual Naming Test in our study confirmed that the Visual Naming Test has construct validity.

Construct validity: age effect. As mentioned before, naming requires two mental stages: activation of lexical-semantic knowledge and phonology activation (Carla J. Johnson, 1996; DeLeon et al., 2007; Paivio, Clark, Digdon, & Bons, 1989). The ability to retrieve the phonology of words appears to decline in older adults (Cotelli, Manenti, Brambilla, Zanetti, & Miniussi, 2012). This temporary inability to access the word phonology, referred to as the tip-of-the-tongue phenomenon, increases with aging. There is also evidence of age-related atrophy in linguistic areas and changes in white matter integrity (Stamatakis, Shafto, Williams, Tam, & Tyler, 2011).

doi:10.6342/NTU201600919

However, among the mental abilities involved in naming, the essential component is lexical-semantic knowledge. This knowledge, also known as crystalized knowledge, is relatively well preserved with advancing age. As Zec et al. (2005) suggested, naming performance is generally well preserved in aging, declining only slightly in the 7th and 8th decades of age. The present study also revealed no significant age effects on all test scores, with only a slight decline in naming performance in the 65-74 age group (see Figure 1-3).

Construct validity: clinical patients. The patient study also confirmed the adequate construct validity of the Visual Naming Test and the Object Naming Test. As expected, patients with mild AD performed significantly more poorly than normal controls on the Visual Naming Test. On the Object Naming Test, the percentage of participants below the cut-off point rose according to the severity of AD; it was lowest in the healthy controls, higher in the very mild AD group, and highest in the mild AD group. The results support the findings by Bayles and Tomoeda (1983) that naming performance declines during the progression of the disease. In Alzheimer's disease, semantic knowledge degrades as the pathology encroaches upon the temporal, frontal, and parietal association cortex, which are crucial for activating names (Hodges & Patterson, 1995). Furthermore, the defective naming performance was evident in patients with different types of aphasia. The finding support Nicholas and co-workers'

doi:10.6342/NTU201600919

claim (1985) that anomia is a common symptom of different types of aphasia and that the naming test is the first step in aphasia assessment.

Norm Appropriacy

Mehrens and Lehmann (1987) proposed that an appropriate norm possesses three attributes: representativeness, recency, and relevance. First, the present study included 322 participants stratified by age, gender, education, and demographic location. Except for the relatively limited sample size in the 0-6 and 7-9 education groups and the slightly higher number of female participants, our overall sample generally matched the population distribution of Taiwan. Therefore, the representativeness of the samples was adequate. Second, the norm data were collected from March 2014 to August 2015, so it appears to have good recency. Third, the normative data demonstrate good relevance, for our sample was appropriate for comparison of adults over 16 years old in Taiwan. In sum, this study provides appropriate normative data for clinical and research purposes.

doi:10.6342/NTU201600919

Conclusion

The present study indicates that the Battery of Naming Tests has adequate construct validity and moderate test-retest reliability. The present normative data have good representativeness, recency, and relevance. Our study confirmed that education significantly influences the Visual Naming Test performance. Regarding the education effect, meaningful differences were found between the two highest groups (i.e., the 10-12 years and above 13 years groups) and the lower groups. Due to the above findings on demographic variables, a demographically-corrected set of normative data for the Visual Naming Test is necessary. Meanwhile, the scores of the Object Naming Test need no correction for demographic variables since there was no demographic effects on the performance. Furthermore, due to the ceiling effect of the Object Naming Test, a cut-off point (the 5^{th} percentile) was recommended for clinical screening of naming impairment.

The limitations of the present study should be noted. The sample size of the lower education groups (i.e., 0-6 and 7-9 education group) was smaller than that of the expected ones. Moreover, the average number of education years in the above 75 age group was higher (mean = 11.76) than that of the expected one. These sampling biases should be considered while interpreting the tests scores. Future study with more low-educated participants is needed.

doi:10.6342/NTU201600919

A caveat should also be considered when performing clinical follow-up evaluations of a patient's naming test performance. Unlike the high test-retest reliability of the study by Flanagan & Jackson (1997), which reported that naming tests had high test-retest reliability, our study showed the Visual Naming Test to have a moderate test-retest reliability. The issue of whether a sampling bias in that study was a primary contributing factor to the discrepancy remains unclear and awaits further investigation.

Furthermore, future investigations on the feasibility of a qualitative scoring system, especially for performance error patterns in the present naming tests, will be necessary in order to build a link to neurocognitive rehabilitation (Gleichgerrcht, 2015; Lin et al., 2014) and promote understanding of the neuropsychological mechanisms involved (Randolph, Lansing, Ivnik, Cullum, & Hermann, 1999).

doi:10.6342/NTU201600919

References

Axelrod, B. N., Ricker, J. H., & Cherry, S. A. (1994). concurrent validity of the MAE visual naming test. *Archives of Clinical Neuropsychology, 9*(4), 317-321.

Azambuja, M. J., Radanovic, M., Haddad, M. S., Adda, C. C., Barbosa, E. R., & Mansur, L. L. (2012). Language impairment in Huntington's disease. *Arquivos de neuro-psiquiatria, 70*(6), 410-415.

Bayles, K. A., & Tomoeda, C. K. (1983). Confrontation naming impairment in dementia. *Brain and language*, *19*(1), 98-114.

Benton, A. L. (1968). Development of a Multilingual Aphasia Battery. *Journal of the neurological Sciences, 9,* 39-48.

Benton, A. L., & Hamsher, K. D. (1978). *Multilingual Aphasia Examination*. Iowa City: University of Iowa Hospital, Department of Neurology.

Benton, A. L., & Hamsher, K. D., & Sivan, A. B. (1994). *Multilingual aphasia examination: Manual of instructions.* Iowa City, IA: AJA Associates.

Berthier, M. L., & Pulvermüller, F. (2011). Neuroscience insights improve neurorehabilitation of poststroke aphasia. *Nature Reviews Neurology, 7*(2), 86-97.

Carla J. Johnson, A. P., James M. Clark. (1996). cognitive components of picture naming. *Psychological Bulletin, 120*(1), 113-139.

doi:10.6342/NTU201600919

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences (2nd ed.) Hillsdale, New Jersey: L: Erlbaum.

Cotelli, M., Manenti, R., Brambilla, M., Zanetti, O., & Miniussi, C. (2012). Naming ability changes in physiological and pathological aging. *Frontiers in Neuroscience, 6*, 120. doi: 10.3389/fnins.2012.00120

Crockett, D. J. (1977). A comparison of empirically derived groups of aphasic patients on the Neurosensory Center Comprehensive Examination for Aphasia. *Journal of clinical psychology*, *33*(1), 194-198.

DeLeon, J., Gottesman, R. F., Kleinman, J. T., Newhart, M., Davis, C., Heidler-Gary, J., . . . Hillis, A. E. (2007). Neural regions essential for distinct cognitive processes underlying picture naming. *Brain, 130*(5), 1408-1422.

Department of Household Registration, M. O. I. (2014). *2013 Population Statistical Yearbook of Interior*. Taipei City, Taiwan: Department of Household Registration, Ministry of the Interior. Retrieved from http://sowf.moi.gov.tw/stat/year/list/html

Duffau, H., Moritz-Gasser, S., & Mandonnet, E. (2014). A re-examination of neural basis of language processing: proposal of a dynamic hodotopical model from data provided by brain stimulation mapping during picture naming. *Brain and Language, 131*, 1-10. doi: 10.1016/j.bandl.2013.05.011

doi:10.6342/NTU201600919

Emery, V. O. B. (2000). Language impairment in dementia of the Alzheimer type: a hierarchical decline? *The International Journal of Psychiatry in Medicine, 30*(2), 145-164.

Engelter, S. T., Gostynski, M., Papa, S., Frei, M., Born, C., Ajdacic-Gross, V., . . . Lyrer, P. A. (2006). Epidemiology of aphasia attributable to first ischemic stroke incidence, severity, fluency, etiology, and thrombolysis. *Stroke, 37*(6), 1379-1384.

Flanagan, J. L., & Jackson, S. T. (1997). Test-retest reliability of three aphasia tests: Performance of non-brain-damaged older adults. *Journal of Communication Disorders, 30*(1), 33-43.

Franzen, M. D. (2000). Tests of Verbal Functions. *Reliability and validity in neuropsychological assessment* (pp. 255-265) America: Springer.

Gainotti, G. (2010). The influence of anatomical locus of lesion and of gender-related familiarity factors in category-specific semantic disorders for animals, fruits and vegetables: a review of single-case studies. *Cortex, 46*(9), 1072-1087.

Glasser, M. F., & Rilling, J. K. (2008). DTI tractography of the human brain's language pathways. *Cerebral Cortex, 18*(11), 2471-2482. doi: 10.1093/cercor/bhn011

Gleichgerrcht, E., Fridriksson, J., & Bonilha, L. (2015). Neuroanatomical foundations

doi:10.6342/NTU201600919

of naming impairments across different neurologic conditions. *Neurology, 85*(3), 284-292.

Goodglass, H., & Kaplan, E. (1983). *Boston diagnostic aphasia examination booklet*. Philadelphia: Lea & Febiger.

Hamberger, M. J. (2015). Object naming in epilepsy and epilepsy surgery. *Epilepsy and Behavior, 46*, 27-33. doi: 10.1016/j.yebeh.2014.12.019

Harry, A., & Crowe, S. F. (2014). Is the Boston Naming Test still fit for purpose? *The Clinical Neuropsychologist, 28*(3), 486-504. doi: 10.1080/13854046.2014.892155

Hawkins, K. A., Sledge, W. H., Orleans, J. F., Quinlan, D. M., Rakfeldt, J., & Huffman, R. E. (1993). Normative implications of the relationship between reading vocabulary and Boston Naming Test performance. *Archives of Clinical Neuropsychology, 8*(6), 525-537.

Hodges, J. R., & Patterson, K. (1995). Is semantic memory consistently impaired early in the course of Alzheimer's disease? Neuroanatomical and diagnostic implications. *Neuropsychologia*, *33*(4), 441-459.

Hua, M. S., Chang, S. H., & Chen, S. T. (1997). Factor Structure and Age Effects With an Aphasia Test Battery in Normal Taiwanese Adults. *Neuropsychology, 11*(1), 156-162.

doi:10.6342/NTU201600919

Jordan, F. M., Ozanne, A. E., & Murdoch, B. E. (1988). Long-term speech and language disorders subsequent to closed head injury in children. *Brain Injury, 2*(3), 179-185.

Laws, K. R. (1999). Gender affects naming latencies for living and nonliving things: Implications for familiarity. *Cortex, 35*(5), 729-733.

Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and brain sciences, 22*(1), 1-38.

Lezak, M., Howieson, D., Bigler, E., & Tranel, D. (2012). *Neuropsychological Assessmentt* (5th ed.). New York: Oxford University Press.

Lezak, M. D. (2004). *Neuropsychological assessment*: Oxford University Press, USA.

Lin, C. Y., Chen, T. B., Lin, K. N., Yeh, Y. C., Chen, W. T., Wang, K. S., & Wang, P. N. (2014). Confrontation naming errors in Alzheimer's disease. *Dementia and Geriatric Cognitive Disorder, 37*(1-2), 86-94. doi: 10.1159/000354359

Mehrens, W. A., & Lehmann, I. J. (1987). *Using standardized tests in education (4th ed.).* New York, NY, US: Longman/Addison Wesley Longman.

Mesulam, M., Rogalski, E., Wieneke, C., Cobia, D., Rademaker, A., Thompson, C., & Weintraub, S. (2009). Neurology of anomia in the semantic variant of primary progressive aphasia. *Brain, 132*(9), 2553-2565. doi: 10.1093/brain/awp138

Morgan, G. A., Gliner, J. A., & Harmon, R. J. (2006). *Understanding and evaluating*

doi:10.6342/NTU201600919

research in applied and clinical settings: Psychology Press.

Murre, J. M. J., Graham, K. S., & Hodges, J. R. (2001). Semantic dementia: relevance to connectionist models of long-term memory. *Brain, 124,* 647-675.

Nicholas, M., Obler, L. K., Albert, M. L., & Helm-Estabrooks, N. (1985). Empty speech in Alzheimer's disease and fluent aphasia. *Journal of Speech, Language, and Hearing Research, 28*(3), 405-410.

Paivio, A., Clark, J. M., Digdon, N., & Bons, T. (1989). Referential processing: Reciprocity and correlates of naming and imaging. *Memory & Cognition, 17*(2), 163-174.

Randolph, C., Lansing, A. E., Ivnik, R. J., Cullum, C. M., & Hermann, B. P. (1999). Determinants of confrontation naming performance. *Archives of Clinical Neuropsychology, 14*(6), 489-496.

Reis, A., Guerreiro, M., & Castro-Caldas, A. (1994). Influence of educational level of non brain-damaged subjects on visual naming capacities. *Journal of Clinical and Experimental Neuropsychology, 16*(6), 939-942. doi: 10.1080/01688639408402705

Riepe, M. W., Riss, S., Bittner, D., & Huber, R. (2003). Screening for cognitive impairment in patients with acute stroke. *Dementia and Geriatric Cognitive Disorder, 17*(1-2), 49-53.

doi:10.6342/NTU201600919

Roelofs, A. (2014). A dorsal-pathway account of aphasic language production: the WEAVER++/ARC model. *Cortex, 59*, 33-48. doi: 10.1016/j.cortex.2014.07.001

Rohrer, J. D., Knight, W. D., Warren, J. E., Fox, N. C., Rossor, M. N., & Warren, J. D. (2008). Word-finding difficulty: a clinical analysis of the progressive aphasias. *Brain, 131*(1), 8-38. doi: 10.1093/brain/awm251

Rolheiser, T., Stamatakis, E. A., & Tyler, L. K. (2011). Dynamic processing in the human language system: synergy between the arcuate fascicle and extreme capsule. *The Journal of Neuroscience, 31*(47), 16949-16957. doi: 10.1523/JNEUROSCI.2725-11.2011

Schum, R. L., & Sivan, A. B. (1997). Verbal abilities in healthy elderly adults. *Applied neuropsychology, 4*(2), 130-134.

Schwartz, M. F., Faseyitan, O., Kim, J., & Coslett, H. B. (2012). The dorsal stream contribution to phonological retrieval in object naming. *Brain, 135*(12), 3799-3814. doi: 10.1093/brain/aws300

Spreen, O., & Benton, A. L. (1969). *Neurosensory Center Comprehensive Examination for Aphasia: Manual of instructions*. Victoria, BC: University of Victoria.

Spreen, O., & Risser, A. H. (2002). *Assessment of aphasia*. New York: Oxford

doi:10.6342/NTU201600919

University Press.

Stamatakis, E. A., Shafto, M. A., Williams, G., Tam, P., & Tyler, L. K. (2011). White matter changes and word finding failures with increasing age. *PLoS One, 6*(1), e14496. doi: 10.1371/journal.pone.0014496

Tyler, L. K., Stamatakis, E. A., Post, B., Randall, B., & Marslen-Wilson, W. (2005). Temporal and frontal systems in speech comprehension: an fMRI study of past tense processing. *Neuropsychologia, 43*(13), 1963-1974. doi: 10.1016/j.neuropsychologia.2005.03.008

Ueno, T., Saito, S., Rogers, T. T., & Ralph, M. A. L. (2011). Lichtheim 2: synthesizing aphasia and the neural basis of language in a neurocomputational model of the dual dorsal-ventral language pathways. *Neuron, 72*(2), 385-396. doi: 10.1016/j.neuron.2011.09.013

Verma, M., & Howard, R. J. (2012). Semantic memory and language dysfunction in early Alzheimer's disease: a review. *International Journal of Geriatric Psychiatry, 27*(12), 1209-1217. doi: 10.1002/gps.3766

Welch, L. W., Doineau, D., Johnson, S., & King, D. (1996). Educational and Gender Normative Data for the Boston Naming Test in a Group of Older Adults. *Brain and Language, 53,* 260-266.

Zec, R. F., Markwell, S. J., Burkett, N. R., & Larsen, D. L. (2005). A longitudinal

doi:10.6342/NTU201600919

study of confrontation naming in the normal elderly. *Journal of the International Neuropsychological Society, 11*, 716-726.

doi:10.6342/NTU201600919

Tables

Table 1

The Expected and Actual Sample Size of the Enrolled Participants

	Age (years)								
	16-24	25-34	35-44	45-54	55-64	65-74	≥75	Total	Exp. N
Gender									
Male	23	25	24	23	23	16	4	138	148
Female	33	25	24	40	25	24	13	184	152
Education (years)									
0-6	0	0	0	7	7	16	4	34	45
7-9	2	2	3	7	6	3	1	24	40
10-12	17	16	20	24	17	7	3	104	96
≥13	37	32	25	25	18	14	9	160	119
Area									
Northern	25	28	26	22	22	13	10	147	138
Western	15	8	9	22	9	15	3	81	75
Southern	15	12	13	19	17	10	4	90	82
Eastern	1	2	0	0	0	2	0	5	5
Total	56	50	48	63	48	40	17	322	
Exp. N	45	58	56	57	44	23	17		300

Note. Exp. N: expected sample size

doi:10.6342/NTU201600919

Table 2

Descriptive Statistics of Demographic Characteristic

	N	Age [a]	Education [b]
Total	322	45.21 (17.94)	12.91 (3.54)
Age (years)			
16-24	56	20.80 (2.32)	13.93 (2.34)
25-34	50	28.74 (2.98)	14.82 (2.70)
35-44	48	39.02 (2.90)	13.67 (2.38)
45-54	63	50.00 (2.88)	12.54 (3.46)
55-64	48	59.33 (2.92)	11.81 (3.95)
65-74	40	68.80 (2.84)	10.60 (4.49)
≥75	17	78.35 (3.61)	11.76 (4.02)
Gender			
Male	138	44.14 (17.34)	13.12 (3.39)
Female	184	46.01 (18.35)	12.76 (3.64)
Education (years)			
0-6	34	64.56 (8.88)	5.59 (1.40)
7-9	24	50.21 (15.34)	9.00 (.00)
10-12	104	43.62 (16.68)	11.83 (.53)
≥13	160	41.38 (17.85)	15.76 (1.39)

Note. [a] mean (standard deviation); [b] percentage

doi:10.6342/NTU201600919

Table 3

Descriptive Statistics of Raw scores by Age, Gender, and Education

	Visual Naming		Object Naming	
	M (*SD*)	*Md* (*R*)	*M* (*SD*)	*Md* (*R*)
Total	53.84 (5.42)	56.00 (32-60)	15.69 (.62)	16.00 (13-16)
Age (years)				
16-24	55.57 (4.80)	57.00 (40-60)	15.66 (.55)	16.00 (14-16)
25-34	54.98 (4.70)	56.00 (44-60)	15.74 (.63)	16.00 (13-16)
35-44	54.58 (4.15)	55.00 (42-60)	15.60 (.57)	16.00 (14-16)
45-54	53.87 (5.74)	56.00 (32-60)	15.62 (.71)	16.00 (13-16)
55-64	53.25 (5.77)	55.00 (40-60)	15.92 (.28)	16.00 (15-16)
65-74	50.75 (5.39)	52.00 (36-60)	15.70 (.72)	16.00 (13-16)
≥75	51.53 (7.02)	52.00 (42-60)	15.47 (.94)	16.00 (13-16)
Gender				
Male	54.71 (4.73)	56.00 (40-60)	15.67 (.63)	16.00 (13-16)
Female	53.19 (5.82)	54.00 (32-60)	15.71 (.62)	16.00 (13-16)
Education (years)				
0-6	46.65 (6.09)	46.00 (32-58)	15.71 (.72)	16.00 (13-16)
7-9	50.25 (4.54)	49.00 (42-58)	15.71 (.62)	16.00 (14-16)
10-12	52.77 (4.96)	54.00 (40-60)	15.61 (.66)	16.00 (13-16)
≥13	56.61 (3.36)	58.00 (42-60)	15.74 (.58)	16.00 (13-16)

Note. M: mean*; SD:* standard deviation*; Md:* median*; R:* range

doi:10.6342/NTU201600919

Table 4

Spearman's rank correlation between the raw scores and the demographic variables

	Gender	Age	Education
Visual Naming	.120*	-.264**	.565**
Object Naming	-.046	.083	.047

Note. Education and age were continuous variables; *$p < .05$; ** $p < .01$.
Rule of thumb of the magnitude of correlation according to Cohen (1988): a correlation coefficient of .10 is thought as a weak association; a correlation coefficient of .30 is considered as a moderate correlation; and a correlation coefficient of .50 or larger is thought as a strong correlation

Table 5

Summary of Multiple Regression Model

	Gender	Age	Education
	b	*b*	*b*
Visual Naming	1.159	-	.876
Object Naming	-	-	-

Note. Education and age were continuous variables, *b*: unstandardized coefficient.

Table 6

Explained Variance in Stepwise Multiple Regression Analyses

	Gender	Age	Education
	ΔR^2	ΔR^2	ΔR^2
Visual Naming	1.2%	-	36.1%

Note. Education and age were continuous variables, ΔR^2: R square Change.

doi:10.6342/NTU201600919

Table 7

Comparison between Different Age and Education Groups by Kruskal-Wallis One-Way Analysis of Variance by Ranks

	Age group χ^2_6[a]	Education group χ^2_3[b]
Visual Naming	5.24[a,c]	101.92[b,c]**

Note. Education and age were discrete variables; $^{**}p < .005$; [a] partial out education effect; [b] partial out age effect; [c] partial out gender effect.

Table 8

Pairwise Comparisons between Different Age and Education Groups by Post Hoc Test with Bonferroni Correlations

	Education group[a,b]
Visual Naming	C >A; D > A, B, C

Note. [a] partial out gender effect; [b] partial out age effect.

A: 0-6 years; B: 7-9 years; C: 10-12 years; D: ≥13 years.

doi:10.6342/NTU201600919

Table 9

Demographic Characteristics of Samples for the Test-retest Reliability of the Battery of Naming Tests

N	Age[a]	Female (%)	Education[a] (year)	Interval time[b] (day)
30	42.43 (16.16)	17 (56.70)	13.13 (3.28)	93.33 [61-141]

Note. [a] mean (standard deviation); [b] mean [range].

Table 10

The Test-retest Reliability of the Battery of Naming Tests

	1st time[a]	2nd time[a]	Reliability coefficients[b]	Practice effect[c]
Visual Naming	55.07 (4.42)	55.80 (3.73)	.60**	Z = -1.26
Object Naming	15.83 (.38)	15.80 (.41)	.00	Z = - .33

Note. [a] mean (standard deviation); [b] Spearman’ rho; [c] Wilcoxon signed rank test.

* $p < .05$; ** $p < .01$.

doi:10.6342/NTU201600919

Table 11

Demographic Characteristics of Samples for the validity of the Battery of Naming Tests

N	Age[a]	Female (%)	Education[a]
30	42.10 (15.22)	17 (56.7)	13.17 (2.47)

Note. [a] mean (standard deviation).

Table 12

Spearman's rho between the Visual Naming Test and the Reference Tests

		Visual Naming Test
Function	Criterion	
Lexical- semantic knowledge	Vocabulary subtest	.640**
	Similarities subtest	.687**
	Token Test	.656**
	Aural Comprehension Test	.364*
Phonological activation	Sentence Repetition Test	.667**
	Reading Comprehension Test	.498**

Note. * p < .05; ** p < .01.

Table 13

Corrected Equations of Raw Scores of the Visual Naming Test

Adjusted scores	Equation
Adj_Visual Naming Test	= raw score $- .921 \times (\text{education} - 12.91)$

Note. Education: education in years

doi:10.6342/NTU201600919

Table 14

Referential Table for Percentile Rank of Test Scores of the Battery of Naming Tests

Percentile	Visual Naming	Object Naming
1	41.19	13.00
2	43.76	14.00
3	44.84	14.00
4	44.84	14.00
5	45.62	14.00
6	46.36	14.00
7	46.84	15.00
8	46.84	15.00
9	47.15	15.00
10	47.83	15.00
11	48.68	15.00
12	48.84	15.00
13	48.84	15.00
14	49.15	15.00
15	49.31	15.00
16	49.51	15.00
17	49.60	15.00
18	50.08	15.00
19	50.36	15.00
20	50.38	15.00
21	50.84	15.00
22	51.00	15.00
23	51.00	15.00
24	51.15	16.00
25	51.15	16.00
26	51.15	16.00
27	51.19	16.00
28	51.31	16.00
29	51.51	16.00
30	51.60	16.00
31	52.36	16.00
32	52.50	16.00
33	52.84	16.00
34	52.84	16.00
35	52.84	16.00
36	52.84	16.00
37	52.90	16.00
38	53.00	16.00
39	53.12	16.00
40	53.15	16.00

(continued)

doi:10.6342/NTU201600919

Percentile	Visual Naming	Object Naming
41	53.15	16.00
42	53.15	16.00
43	53.15	16.00
44	53.17	16.00
45	53.31	16.00
46	53.77	16.00
47	54.31	16.00
48	54.39	16.00
49	54.84	16.00
50	54.84	16.00
51	54.84	16.00
52	54.84	16.00
53	54.84	16.00
54	55.00	16.00
55	55.00	16.00
56	55.15	16.00
57	55.15	16.00
58	55.15	16.00
59	55.15	16.00
60	55.15	16.00
61	55.15	16.00
62	55.15	16.00
63	55.31	16.00
64	55.52	16.00
65	55.75	16.00
66	56.10	16.00
67	56.29	16.00
68	56.57	16.00
69	56.84	16.00
70	56.84	16.00
71	56.84	16.00
72	56.84	16.00
73	56.84	16.00
74	56.84	16.00
75	57.00	16.00
76	57.15	16.00
77	57.15	16.00
78	57.15	16.00
79	57.15	16.00
80	57.15	16.00

(continued)

doi:10.6342/NTU201600919

Percentile	Visual Naming	Object Naming
81	57.15	16.00
82	57.15	16.00
83	57.60	16.00
84	57.97	16.00
85	58.08	16.00
86	58.08	16.00
87	58.36	16.00
88	58.84	16.00
89	58.84	16.00
90	58.84	16.00
91	58.99	16.00
92	59.00	16.00
93	59.23	16.00
94	59.92	16.00
95	60.63	16.00
96	60.84	16.00
97	60.84	16.00
98	61.25	16.00
99	62.23	16.00

Note. The cut-off point is the 5th percentile.

doi:10.6342/NTU201600919

Table 15

Percentage of Participants Who Gained Full Score

	Whole sample	Reliability sample session 1	Reliability sample session 2	Validity sample
Visual Naming	17.4%	16.7%	20.0%	23.3%
Object Naming	76.4%	83.3%	80.0%	80.0%

Table 16

Demographic Characteristics of Samples for the Construct Validity of the Battery of Naming Tests

	N	Age[a]	% of female	Education[a]
Mild AD	12	76.50 (8.67)	50.00	13.17 (3.74)
Very mild AD	40	74.05 (9.76)	47.50	12.78 (3.87)
Normal controls	57	71.65 (5.36)	52.63	10.95 (4.35)

Note. [a] mean (standard deviation).

Mild AD: mild stage Alzheimer’s disease (CDR = 1).

Very mild AD: very mild stage Alzheimer’s disease (CDR = 0.5).

doi:10.6342/NTU201600919

Table 17

The Cut-off Point and the Percentage below the Cut-off Point of the Battery of Naming Tests

The cut-off point[a]	
Visual Naming Test	Object Naming Test
42.96	14.00

The percentage below the cut-off point			
	Healthy control	Very mild AD	Mild AD
Visual Naming Test	3.51% (2/57)[b]	20.00% (7/35)[b]	30.00% (3/10)[b]
Object Naming Test	3.51% (2/57)[b]	11.11% (2/18)[b]	42.86% (3/7)[b]

Note. [a] the cut-off point came from the 5th percentile score of the normal control group [b](number of participants whose performance were below cut-off point/ total number of participants)

Table 18

Comparison between Different Severity Groups by the Kruskal-Wallis One-Way Analysis of Variance by Ranks

	Severity group χ^2_2
Visual Naming Test	16.40**
Object Naming Test	6.60*

Note. * $p < .05$; ** $p < .01$.

Table 19

Pairwise Comparisons between Different Severity Groups by Post Hoc Test with Bonferroni Correlations

	Severity group
Visual Naming Test	A > B, C
Object Naming Test	A > C

Note. A: healthy control; B: very mild stage Alzheimer's disease; C: mild stage Alzheimer's disease.

doi:10.6342/NTU201600919

Table 20

Demographic Characteristic of Patient Groups for the Construct Validity of the Battery of Naming Tests

	N	Age[a]	% of female	Education[a]
Global aphasia	7	64.57 (15.87)	14.29	8.29 (4.07)
TSA	2	66.00 (16.97)	0.00	10.00 (5.66)
Conduction aphasia	1	75.00	0.00	12.00
Wernicke's aphasia	1	71.00	0.00	16.00

Note. [a] mean (standard deviation); TSA: transcortical sensory aphasia

Table 21

Percentage of Defective Performance of Patient Groups on the Battery of Naming Tests

	Global aphasia	Wernicke's aphasia	Conduction aphasia	Transcortical sensory aphasia
Visual Naming Test	Unavailable	100%	Unavailable	100%
Object Naming Test	100%	100%	100%	100%

Note. If the patient's adjusted score is at or below the fifth percentile, the performance is regarded as defective.

doi:10.6342/NTU201600919

Figures

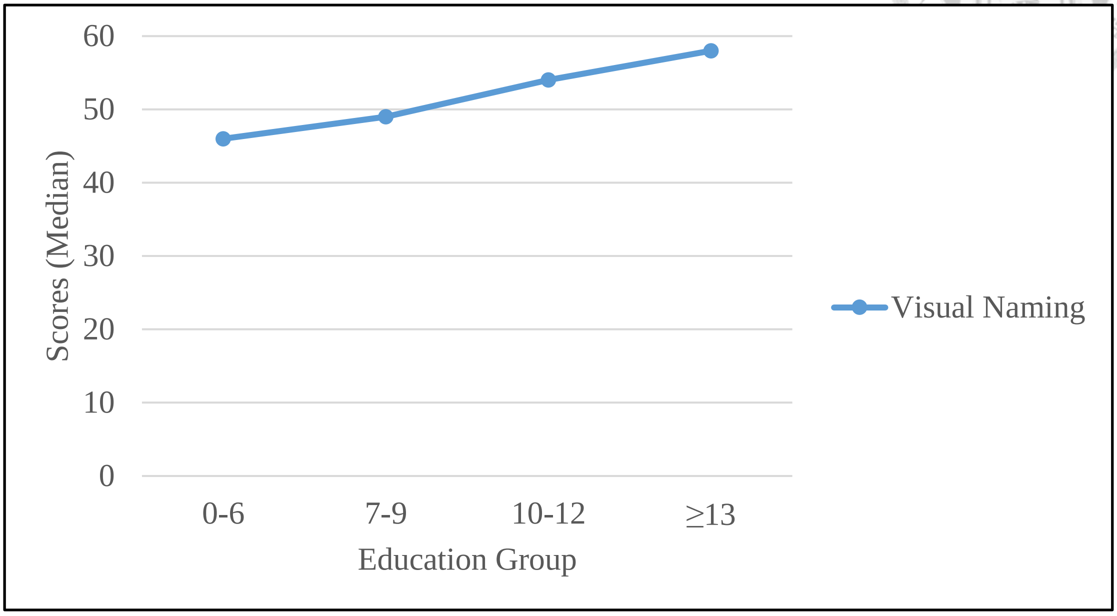

Figure 1-1 Scores (median) of the Visual Naming Test in different education groups

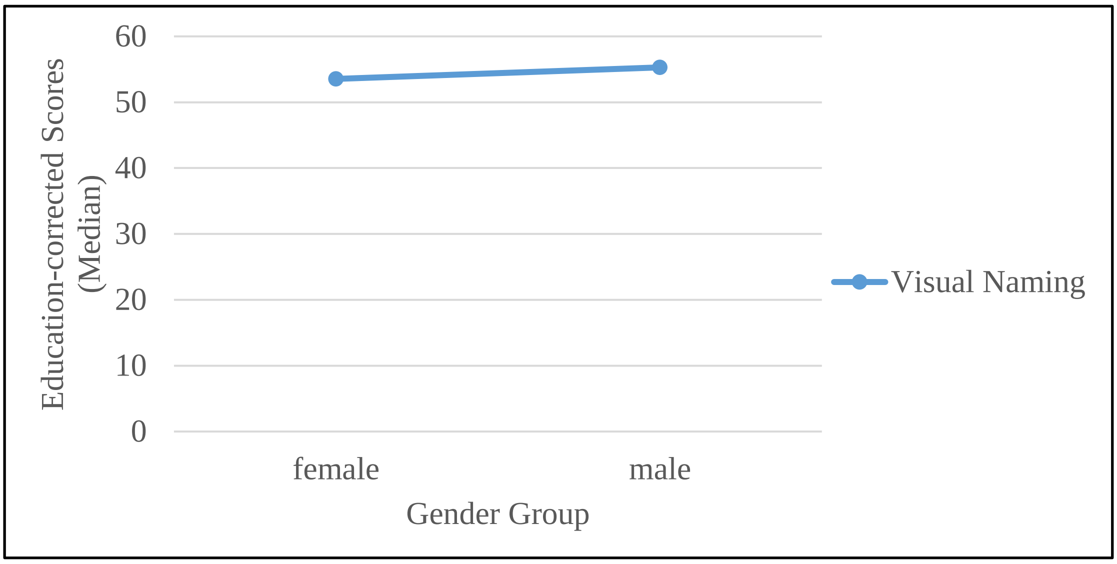

Figure 1-2 Education-corrected scores (median) of the Visual Naming Test in different gender groups

doi:10.6342/NTU201600919

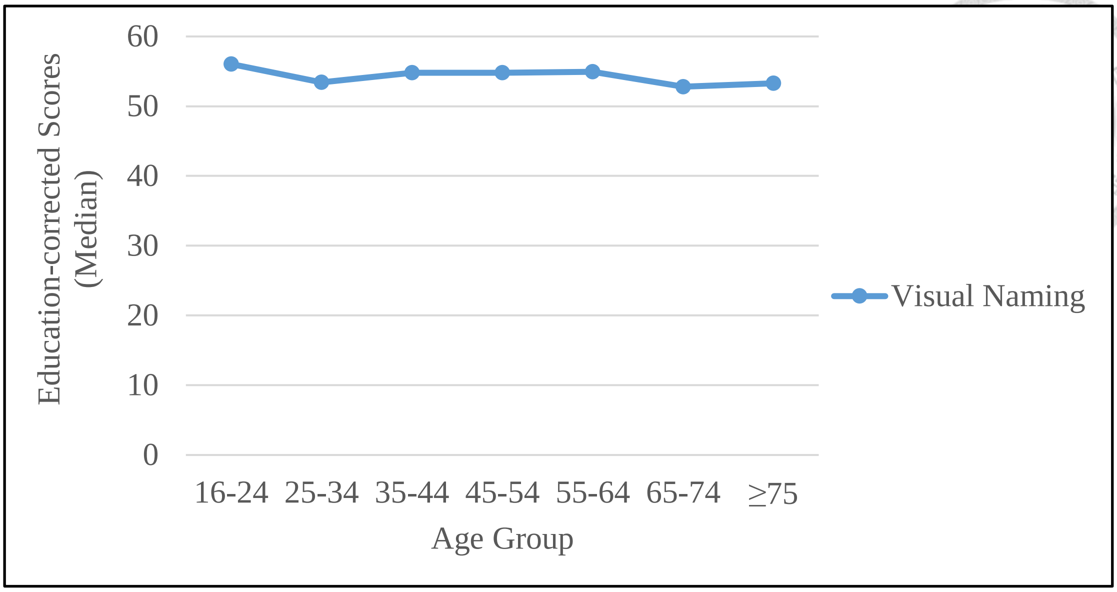

Figure 1-3 Education-corrected scores (median) of the Visual Naming Test in different age groups

doi:10.6342/NTU201600919

Appendix

Appendix A: Informed Consent Form

參與研究同意書

國立台灣大學心理學研究所

您好：

首先，感謝您參與本次的研究。本研究名稱為「臺大失語症檢查之臺灣健康成人常模研究」。此研究之主要目的在於建立臺灣地區臺大失語症檢查之常模，並探討臺灣地區人口學變項與臺大失語症檢查表現之相關性。同時評估臺大失語症檢查在臺灣健康成人使用上之信度與效度。

語言障礙常見於中風患者、腦傷患者以及神經退化性疾病(如阿茲海默症)患者身上，因此一套具備良好信效度的失語症檢查工具對臨床上的診斷、復健以及告知家屬相關資訊，是非常有幫助的。臨床上判斷一個人是否有語言障礙，必須要將他的表現和一般正常人做對照、比較才可以得知。而本研究便是要為「臺大失語症檢查」，收集正常健康成人語言測驗表現的資料。由於中文的閱讀與書寫可能與視覺和空間能力有關，因此本研究也會需要您做兩個視空間能力的作業。此外，因為性別、年齡、教育程度等因素有可能會與測驗表現有關，所以本次研究也會請您提供這些資訊。本研究希望能提供健康成人語言表現水準的資料，幫助臨床實務工作者更有效地評估與追蹤病人語言功能狀況。

參與者條件：16-84 歲之臺灣健康成人適合參與本研究；若有未矯正之聽覺/視覺障礙、物質濫用(包括正在戒毒或戒酒)、腦傷病史、精神疾患病史、中樞神經系統疾病病史、不識字、MMSE 分數低於 24 分(或少於兩年教育程度且分數低於 18 分)等上述情況者則不適合參與本研究。

本研究過程會透過問卷填寫以及神經心理測驗評估的方式完成。參與本研究沒有特定的禁忌與限制，對個人並無侵入性危險，不會涉及任何風險，也不提供補償與保險。研究主要分成以下兩個部分，大約需一個小時的時間完成所有內容：

1. 施測前：主試者簡介研究目的，研究參與者閱讀並填寫同意書、基本資料、與疾病史等相關問卷。
2. 正式施測：由主試者對參與者進行神經心理功能測驗。主試者會展示一些圖片、文字或物體的刺激，並請參與者依照主試者的指導語作出反應。

基於我們對於您個人權益的尊重，本研究對您有以下的承諾：

※您有權隨時停止作答，並且不會因此而受到懲罰。

※您有權要回您的資料，並且撤銷被納入分析。

※退出研究不會引起任何不愉快，或影響研究者對您的評價，更不會損及您的任何權益。

※您有權在本研究結束後知道研究結果。

※如有任何疑問，您可隨時經由參與者聯中所附之聯絡方式與研究人員進行聯絡。

doi:10.6342/NTU201600919

本研究之結果，僅供學術使用，將不做其他用途，也無衍生商業利益。研究計畫主持人將依法把任何可辨識您身分之紀錄與您個人隱私之資料視同機密處理，絕對不會公開。將來發表研究結果時，您的身份將被充分保密。凡簽署了知情同意書，即表示您同意各項原始紀錄可直接受監測者、稽核者、研究倫理委員會及主管機關檢閱，以確保研究過程與數據，符合相關法律和各種規範要求；上述人員承諾絕對維繫您身分之機密性。請您放心填答，謝謝您的合作！

本研究計畫由國立臺灣大學行為與科學研究倫理委員會審查通過，委員會係依規範運作，並通過中央目的事業主管機構查核認證之審查組織。凡研究參與者於研究過程中自認權利受到影響、傷害，可直接與國立臺灣大學研究倫理中心聯絡，電話為 02-33669956、02-33669980

如您同意參與本研究，請於簽上您的姓名與聯絡方式，謝謝您。

本人已詳細閱讀本同意書，並同意參與研究。

參與者簽名：__________ 聯絡方式：__________日期：_____年____月____日

再次感謝您。

實驗/主試者：程柏雅、鄭德雯、馬子倫、吳瑋庭、林依儒
國立臺灣大學心理學研究所/臨床心理學組/碩士班研究生
聯絡地址：國立台灣大學心理學系南館 118 室
研究計畫聯絡人：程柏雅、馬子倫
聯絡電話/手機：02-23695438/09********/09********
指導老師：花茂棽
國立臺灣大學心理學系教授
聯絡地址：國立臺灣大學心理學系南館 118 室
聯絡電話/手機：02-23695438

研究者簽名：_________ 日期：_____年____月____日

doi:10.6342/NTU201600919

Appendix B: Ethical Review Approvals

國立臺灣大學
行為與社會科學研究倫理委員會

Research Ethics Committee
National Taiwan University
No. 1, Sec. 4, Roosevelt Rd., Taipei, Taiwan 10617, R.O.C
Phone: 3366-9956 Fax: 2362-9082

審查核可證明

核可日期：2014 年 3 月 10 日

倫委會案號：201401HS026
計畫名稱：臺大失語症檢查之臺灣健康成人常模研究
校/院/系/計畫主持人：國立臺灣大學/理學院/心理學系暨研究所/花茂棽教授
計畫文件版本日期：【2014 年 3 月 3 日】

上述計畫業經 2014 年 3 月 10 日國立臺灣大學行為與社會科學研究倫理委員會同意，符合研究倫理規範。本委員會的運作符合國立臺灣大學行為與社會科學研究倫理準則與規範及政府相關法律規章。

本案需經研究經費補助單位核准同意後，該計畫始得執行。

本審查核可證明之有效期限為 1 年（自 2014 年 3 月 10 日起至 2015 年 3 月 9 日止），計畫主持人最遲應於本核可證明到期前的 6 週，提出持續審查申請表，本案需經持續審查，方可繼續執行。

在計畫執行期間，若有計畫變更或嚴重不良反應事件，計畫主持人須依國內及國立臺灣大學相關法令規定通報本委員會。

行為與社會科學研究倫理委員會主任委員 謝世忠

Ethical Review Approval
National Taiwan University

Date of approval: March 10, 2014

NTU-REC No.：201401HS026
Title of protocol：National Taiwan University Aphasia Examination: Norms for Healthy Taiwanese Adults
University/College/Department/Principle Investigator：National Taiwan University / College of Science / Department of Psychology / Professor Mau-Sun Hua
Version date of documents：【March 3, 2014】

The protocol has been approved by Research Ethics Committee of National Taiwan University and has been classified as expedited on March 10, 2014. The committee is organized under, and operates in accordance with, Social and Behavioral Research Ethical Principles and Regulations of National Taiwan University and governmental laws and regulations.

Approval by funding agency is mandatory before project implementation.

The duration of this approval is one year （from March 10, 2014 to March 9, 2015）. Continuing Review Application should be submit to Research Ethics Committee no later than six weeks before current approval expired.

The investigator is required to report protocol amendment and Serious Adverse Events in accordance with the National Taiwan University and governmental laws and regulations.

Chairperson Shih-chung Hsieh
Research Ethics Committee

doi:10.6342/NTU201600919

Appendix C: Administration and Scoring

視覺命名測驗—施測程序與計分

施測程序

指導語:「這叫做什麼?(同時指圖片)這個部分叫什麼?」

計分

每題 2 分,共 30 題,滿分 60 分。

題項

1. 雞	16. 鼓
2. 雞冠	17. 鼓棒
3. 雞爪	18. 鼓面
4. 雞翅	19. 海島
5. 膝蓋	20. 海灣
6. 小腿	21. 半島
7. 腳跟	22. 信封
8. 腳背	23. 郵票
9. 十字	24. 郵戳
10. 橢圓形	25. 茶盤
11. 半圓形	26. 茶壺
12. 六角形	27. 茶杯
13. 圓形	28. 電話
14. 三角形	29. 號碼盤
15. 長方形	30. 筷子

doi:10.6342/NTU201600919

物體命名測驗—施測程序與計分

施測程序

指導語：「這叫做什麼？(同時指物品)」

計分

每個項目 1 分，每個分測驗有 8 題，但只要施測 A+B 或 C+D，所以滿分 16 分。如果受試者 A(或 C)分測驗全對，就不用施測 B(或 D)分測驗，B(或 D)分測驗以全對(8 分)計。

題項

A 組	B 組
梳子	刀子
戒子	叉子
鑰匙	奶瓶
茶杯	鞋帶
煙灰缸	刷子
鈕扣	水壺
鎖	開罐器
別針	筷子

C 組	D 組
手槍	手錶
盤子	眼鏡
燈泡	湯匙
螺絲起子	牙刷
海綿	火柴盒
尺	印章
飯匙	原子筆
彈簧	黑板擦

doi:10.6342/NTU201600919