Quality, Quantity? Performance of Irish people with & without aphasia on the Boston Naming Test

Author: Rachel Kennedy
Student Number: 11007281
Supervisor: Professor Sue Franklin
Journal Style: Aphasiology
Word Count: 7,462

Submitted in part fulfilment of the MSc speech and language therapy (professional qualification), department of clinical therapies
Acknowledgements:

My sincerest thanks to all the people who were so kind, generous and patient in giving their time to participate in this study.

My deepest gratitude to all the speech and language therapists who took the time to review their caseloads, gather the details of suitable participants for the study and coordinate with the researchers.

Endless thanks to the friends and family who were there to problem solve, brainstorm, discuss, debate and commiserate with me for the duration of this study and its write up. Particularly deep thanks to JM, AH & HK.

Thanks also very deservedly to the three other student researchers with whom I worked to gather the data for the study; in particular Claire Johnstone, my partner in analysis and late night data rechecks.

To my supervisor Sue Franklin, greatest thanks for your patience, guidance, time, advice and encouragement throughout.
Table of Contents:

Abstract ... 1, p.1

Introduction to the Study ...2, p.2

Models of Word Production ...2.1, p.3
The Components of Successful Naming... 2.2, p.3
Theoretical Operations of Successful Naming ...2.3, p.3
Individual Factors Affecting Word Retrieval ...2.4, p.4
Level of Education ...2.4.A, p.4
Cognitive Ability ...2.4.B, p.4
Culture ...2.4.C, p.4
Ageing ...2.4.D, p.5
Naming Errors ...2.5, p.5
BNT & Research ...2.6, p.7
BNT & Methodology ...2.7, p.7
Summary ...2.8, p.8

Aims of the Current Study ...3, p.9

Method ...4, p.9

Participants ...4.1, p.9
Materials ...4.2, p.12
Procedure ...4.3, p.12
Codification of the Data... 4.4, p.13
Inter Rater Reliability ... 4.5, p. 14
Correct Responses & Error Types ...4.6, p.14

Results ...5, p.15

Section One: Score & Normality ...5.1, p.15
Score & Gender, Age & Years of Education ...5.2, p.16
Score Range ...5.3, p.17
Mean Scores ...5.4, p.18
Section Two: Error Types ...5.5, p.19

Discussion...6, p.22

Summary ...6.1, p.22
Naming Agreement & Cultural Appropriateness ...6.2, p.23
External Factors Influencing Naming...6.3, p.23
Gender ...6.3.A, p.23
Age ...6.3.B, p.23
Education ...6.3.C, p.23
Comparison of Scores Across Groups ...6.4, p.23
Limitations of the Study ...7, p.26
Clock Drawing Test ...7.1, p.26
Sample Sizes ...7.2, p.27
PWA & PWOA ...7.3, p.27
Error Analysis ...7.4, p.27
Participant Factors ...7.5, p.28

Conclusions... 8, p.28

Recommendations ...9, p.29

References ...10, p.30

Tables

1 Demographic Data for PWOA ...p.10
2 Demographic Data for PWA ...p.11
3 Criteria for Inclusion or Exclusion in Study ...p.11
4 Scoring Systems Used ...p.13
5 Defined Error Types ...p.16
6 Overlap Group Error Analyses: Overall Errors & PWA Errors ...p.21
7 Main Findings of The Study ...p.23
8 BNT – Original Adult Norm Data ...p.25

Figures

1 PWOA – Total Correct on First Response ...p.15
2 PWA – Total Correct on First Response ...p.16
3 Overlap of mean scores across groups ...p.17
4 Frequencies of PWA & PWOA obtaining each score across groups ...p.17
5 Receiver Operating Curve Analysis ...p.20
6 Mean Number of Error Types Made Across Groups ...p.21

Appendices ...p.40

A Information about the Research Project for Participants ...p.40
B Aphasia Friendly Information about the Research ...p.42
C Clock Drawing Test Scoring Instructions ...p.46
D Test Forms ...p.47
E Error Types (Explanation of Classification System) ...p.50
1. Abstract

**Background:** Word Finding Difficulty (WFD) is a near-ubiquitous component of aphasia, and may be mild to severe in impact. WFD is typically assessed with confrontation naming assessments, such as the Boston Naming Test (BNT, Kaplan et al 1983). Recent research indicates that error score alone may not detect WFD in people with aphasia (Grima & Franklin, 2012) and that the sensitivity of the BNT is increased once standardized for specific populations (Casas et al, 2008).

**Aims:** This study compared the performance of people with aphasia (PWA) and without aphasia (PWOA) on the BNT in Ireland. Aims included determination of sensitivity and cultural appropriateness; querying whether score could differentiate PWA from PWOA; whether age and education could explain score overlap between groups, and whether a qualitative error profile could differentiate PWA from PWOA.

**Methods:** The BNT was administered to aphasic (N=35) and non-aphasic Irish older adults (N=67). BNT scores, error types and demographics for the two groups were compared.

**Outcomes & Results:** As reflected in earlier studies (James, 2012) mean scores of PWA and PWOA were significantly different. Even when culturally inappropriate items were removed from data analysis, the score ranges of the two groups overlapped. This overlap was not explicable in terms of age or education. Comparison of error types revealed two kinds unique to PWA, and several significantly more common in PWA.

**Conclusions:** Research supports earlier evidence that BNT score alone may not be sufficient to diagnose WFD due to aphasia. Atypical error types do not accurately differentiate PWA from PWOA either, but may be key to more informed diagnosis of WFD in PWA.
2. Introduction to the Study

In our conversations and interactions every day, we employ a complex, high-speed system of lexical access and processing that works to create fluent, meaningful speech. This system is integral to effective communication, participation (Gollan et al 2007, Carragher et al 2012) and individual identity (Pearl et al, 2011).

Damage to this system has profound effects (Herbert et al, 2008). Anomia is a frequent residual symptom of aphasia (Nickels 2002, Laine & Martin 2012). People with aphasia (PWA) often pinpoint it as their principle source of frustration and impedance (RCSLT, 2005). Anomia has been described as a more troublesome, frequent and persistent form of the occasional WFD, or lexical retrieval difficulty (LRD) experienced by all people without aphasia (PWOA).

The common nature of LRD in PWA means that its remediation is a high priority treatment goal in speech & language therapy. Accurate, sensitive assessments best inform precisely targeted, effective therapies.

The Boston Naming Test (BNT) is the most frequently used assessment tool for the identification of Anomia, or Lexical Retrieval Difficulties (LRD). It is geared towards detection of mild naming deficits. It is used worldwide, and generally supported by available research as a valid tool (Connor et al 2004) though not without psychometric shortcomings (Jackson & Tompkins, 1991). It has not yet been formally standardized for an Irish population.

The BNT uses error score, and score cut-off points to determine a diagnosis of LRD. However, Scores alone do not always successfully differentiate between PWA and PWOA (Grima, 2011). Research on error type has been suggested as a viable tool for differential diagnosis (Budd et al, 2010). Additionally, demographic factors such as age and education have been identified as
influencing naming for PWA and PWOA, and are important to consider in assessment of LRD.

2.1 Models of Word Production

It is generally accepted that successful naming is an outcome of a network of interconnecting ‘processes,’ or ‘levels’ all functioning effectively (Dell et al 1997, Levelt 2001). Research on how different levels in these networks may be impaired (explaining different forms of LRD) has informed research about how normal naming may occur (Wilshire 2008, Nickels 2002).

2.2 The Components of Successful Naming

The five levels of processing of this system generally agreed upon are as follows: Concepts, Lexical Semantic, Lexical Phonology, Phonological Assembly and Articulatory Programming (Grima 2011, Nickels 2002, Nickels 2006).

These components provide valuable information about why different naming errors may occur, via levels of breakdown at specific points (Nickels 2002). Naming ability and lexical access go beyond models and neurological networks, however, and are influenced by many more factors (Nickels & Howard, 1995).

It is possible to divide the research between explorations of factors that may have an effect due to the speaker versus an effect due to what is being named (Grima, 2011). These are defined as individual versus material categories. Material factors have been widely discussed in the literature. For the purposes of relevance and brevity, individual factors are briefly delineated as follows.

2.3 Theoretical Operations of Successful Naming

Some theories of naming posit that the connections between levels are unidirectional and discrete (Levelt et al, 1999) and others state that connections operate multidirectionally and simultaneously (Dell et al, 1997). An intermediate
position has also been proposed called the *restricted interaction account* (RIA). (Rapp & Goldrick, 2000) that constitutes components of both ideas.

There is some debate as to whether significant difference exists between the operations of retrieval at single word versus sentence level (Mayer & Murray, 2003). Some studies on aphasics performing retrieval tasks in speech versus confrontational naming tasks have found the former is easier (Pashek & Tompkins 2002), others the latter (Wilshire & McCarthy, 2002). Context, individual factors, aphasia subtype and severity of aphasia may account for the heterogeneity of findings.

### 2.4 Individual Factors Affecting Word Retrieval

The following particular factors are shown to give rise to a variation in performance on naming tasks as consistently demonstrated in the literature (Barker-Collo 2001, Tallberg 2005, Ross & Lichtenburg 1998).

#### 2.4.A Level of Education

(Tallberg et al 2005, Patricacou et al 2007) may be the best predictor of naming proficiency (Mansur et al 2006, Tombaugh & Hubley 1997). Higher levels of education purportedly eliminate the negative effects of ageing on naming (Goral et al 2007). Studies that have found no effects from education have frequently been criticized, for example, through deliberate use of test items high in frequency and low in age of acquisition (Tsang & Lee, 2003) or through bias to more educated samples (Van Gorp et al, 1986).

#### 2.4.B Global Cognitive Ability

(Tallberg 2005, Steinberg et al 2005) is usually measured as IQ and is significantly correlated with BNT performance; better IQ correlates to better performance. Higher functioning older adults also reported only mild decline with age on BNT.

#### 2.4.C Culture

(Casas et al, 2008, Barker-Collo, 2001) has been shown to influence performance on the BNT. For example, Australian participants may have much less familiarity with words like beaver and pretzel than US
participants. Urban participants may be less familiar with farming equipment. The number of languages a subject uses may also impact (Allegri et al 1997, Himmanen 2005).

2.4. D Ageing produces several changes in a variety of functions. It has been shown that ageing relates to physical changes in the brain (Good et al 2001, Salat et al 2009). Difficulties with word retrieval and “tip of the tongue” states are found to be more common in the elderly (Belke & Meyer, 2007). Some studies have linked ageing to increased difficulty with word retrieval (Wierenga et al 2008, Figueredo Balthazar et al 2008). It is a very common complaint of healthy ageing individuals.

2.5 Naming Errors

All people have occasional difficulty naming. Errors in naming have been hypothetically explained with a number of word-retrieval models, which have been discussed and applied to aphasic patients with varying degrees of success (Nickels, 2000, Capitani & Laiacona, 2004).

Different symptoms have been seen in individuals with similar lesion sites, and varying levels of breakdown has been seen in individuals who present with superficially similar aphasias (Howard and Gatehouse 2006, Laganaro et al 2006). Differences may be due to minute variation in precise site lesion, time post onset of damage, individual differences in cognition (Budd et al 2010), experience, background and premorbid abilities.

Some studies have focused on the processing level that may be associated with particular errors (Howard & Gatehouse, 2006). Error types and points of breakdown in the word finding chain can be compared with parallel research in deep dyslexia (Riley & Johnson, 2010). Others analyse the influence of the target word on the type of errors produced (Nickels & Howard, 1995).
Naming errors seem to be much more common in aphasic than in normal speech. Research on whether there are errors unique to aphasics, or whether it is simply a matter of error quantity, is small. Some studies (Grima 2012, James 2012) suggest that aphasics produce naming errors that are unique to their condition, such as unsuccessful conduite d’approche – the initial retrieval of part of the sound of the word without eventual retrieval of the full word.

The types of errors produced by people with aphasia can be classified into subtypes. For example, nine types are used by Cuetos et al (2002); semantic errors, formal errors (phonologically related real words), mixed semantic and formal errors (errors semantically and phonologically related to the real word), visual errors, neologisms, unrelated errors, no-response errors, circumlocutions (including conduite d’approche) and ‘other’. Many of these error types are also seen in normal populations. Other classifications of error types have been used that are both more and less specific in their definitions.

An interactive theory of word retrieval puts forward that the naming errors of aphasics differ in quantity but not in quality (Dell et al, 1997). Schwartz et al (2006) take the opposite stance, stating the likelihood of some errors is much higher in aphasics, with a continuum of likelihood from low to high severity.

Mixed errors are consistent with an interactive theory of word retrieval. Modular models, such as sole production of semantic or phonological errors, better explain other errors aphasics have been observed to make.

Functional imaging studies of anomic performance (Fridriksson et al 2009) found physiological differentiation in neurological activity correlating to breakdown diagnosis. This is evidence that different levels of breakdown are a physical reality and that LRD is deserving of accurate differential diagnosis.
2.6 BNT & Research

The BNT has been used for researching and diagnostics in many different populations to date: studying aphasic errors, LRD in adults and children, dementia (Hall et al, 2012) and progressive neurological conditions.

The BNT has been standardized in many countries (Patricacou et al 2007, Casanova et al 2009, Tallberg 2005, Mansur et al 2006), on different socioeconomic groups (Ross & Lichtenburg, 1998) and bilingual populations (Gollan et al 2006, Sachs et al 2012). Standardization for specific populations provides accurate normative information that can guide clinicians in diagnosing and decision-making to provide best possible outcomes for their clients (Zec et al, 2007).

Some studies have found significant substandard scores in normal populations, which were attributed to generational and cultural difference (Barker-Collo, 2007). Common errors from studies have also been attributed to these differences (Pedraza et al, 2009). For this reason the BNT has often been adapted and some of its pictures swapped or resequenced (Cruice et al, 2000, Tallberg, 2005).

2.7 BNT & Methodology

As mentioned, some research projects have used revised protocols for administration. A study may note the use of lenient scoring with detailed explanation and rationale (Tallberg, 2005) or descriptions of scoring may be vague (Barker-Collo, 2007).

Further to that, sample sizes have often been small in relation to the population that they represent, or not fully representative of a population. The current sample used to standardize the BNT in the US is based on 119 participants (Kaplan et al, 2001). Studies will also frequently exclude participants in normative studies who have a history of depression, or low cognitive functioning
because these are known to impair performance (Belke & Meyer 2007, Cruice et al 2000).

The BNT has not escaped criticism of its layout & design (Himmanen et al 2005, Pedraza et al 2011). Its psychometric validity is low compared to recommended standards (Friberg 2010). More research that explicitly states the particulars of method, scoring and decisions in analysis would be valuable in better informing guidelines about future use of the BNT as an assessment and research tool.

The BNT is best used as one part of a clinical assessment of WFD. It has less use as sole source of assessment for WFD. Analysis of conversation and spontaneous speech, for example, should form part of a portfolio of assessment (Rohrer et al 2008, Berndt et al 2002).

2.8 Summary

LRD can have a devastating effect on an individual’s quality of life. Some individuals may feel these effects in their day-to-day experience, but demonstrate a score on the BNT that is within the ‘normal’ range. Conversely, some people without aphasia achieve scores that appear to suggest the presence of a naming impairment.

Further to that, it has been shown that outside America (the population for whom the BNT was originally standardized), scores from samples of people without aphasia overlap with scores with a range suggestive of LRD, which may be due to a collection of individual differences psychometrically unaccounted for.

Score alone cannot demonstrate impairment. Error types may tell an assessing clinician more about a person with aphasia, and may help to distinguish normal from impaired naming better than score. Country and culture specific standardization may achieve higher accuracy for assessment using the BNT.
3. Aims of the Current Study

- To analyze the performance of aphasic Irish speakers on the BNT.
- To examine cultural appropriateness of the BNT for Irish PWA & PWOA.
- To test the hypothesis that BNT score alone differentiates Irish PWA from PWOA.
- To explore the hypothesis that value exists in a qualitative error profile.
- To explore the hypothesis that some error types are unique to PWA.

4. Method

4.1 Participants.

PWOA: a sample totaling 51 people was used from earlier research. The purpose of collecting data from people without aphasia was to ascertain whether the BNT had face validity for an Irish sample. The study tested whether the pictures used in the BNT were recognised, familiar and associated with the intended concept. It was also to ascertain the typical naming abilities of healthy Irish older adults.

This initial sample of 51 people was recruited using poster advertisements placed in active retirement clubs, work places and social clubs. Recruitment and participation for this study took place in 2011 & 2012, subsequent to ethical approval from the University of Limerick. Participants met once with the researcher and gave their consent to participate immediately before their participation, due to the non-invasive nature of the research.

A limitation of the earlier sample was that volunteers were mostly people with high levels of education. In order to try and remediate this sample bias, more
PWOA with lower levels of education were sought under the same ethical approval, which was still valid at the time of this research (Jacobsen, 2011). This was an anticipated challenge; firstly, people who have lower educational status are less likely to respond positively to requests for research participation (Patel et al, 2003). Advertising explicitly for people with less education, and/or advertising in areas with lower rates of education were vetoed due to uncertainty of response and limited time.

Therefore, the researchers recruited participants by word of mouth through friends, colleagues and acquaintance. 18 additional PWOA were recruited this way. Of the 18 additional PWOA, two were later excluded for analysis due to corrupted sound files. 67 PWOA therefore constituted the final sample.

Table 1: Demographic Data for PWOA

<table>
<thead>
<tr>
<th>Males</th>
<th>Females</th>
</tr>
</thead>
<tbody>
<tr>
<td>29</td>
<td>38</td>
</tr>
<tr>
<td>Mean Age of all PWOA</td>
<td>Age Range of all PWOA</td>
</tr>
<tr>
<td>65.39</td>
<td>50 – 83</td>
</tr>
<tr>
<td>Mean Years Education of all PWOA</td>
<td>Mean Range Years Edu all PWOA</td>
</tr>
<tr>
<td>14.46</td>
<td>7 – 27</td>
</tr>
</tbody>
</table>

**PWA:** The purpose of collecting data from Irish people with aphasia was to compare the scores of a group with confirmed acquired LRD with the Irish norms for scores of naming ability and also for familiarity, recognition and association with the BNT items obtained from the PWOA sample.

Ethics clearance for the research on PWA was obtained through the University of Limerick for four hospital sites in Dublin.

Recruitment was via judgmental sampling – a group of referring speech and language therapists at the hospital sites provided the researchers with referrals for clients with confirmed word finding difficulties.
A total of 37 PWA were recruited in this way. 2 were later excluded for analysis; both had their assessments discontinued, one due to fatigue and evidence of significant visual impairment, and the other due to personal time commitments. 35 PWA therefore constituted the final sample.

All those who agreed to participate were presented with information about the research, verbally and in written form (Appendix A). An aphasia-friendly written form was also available for the PWA (Appendix B).

**Table 2: Demographic Data for PWA**

<table>
<thead>
<tr>
<th></th>
<th>Males</th>
<th>Females</th>
</tr>
</thead>
<tbody>
<tr>
<td>21</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>Mean Age of all PWA</td>
<td>74.29</td>
<td>Age Range of all PWA</td>
</tr>
<tr>
<td>Mean Years Education of all PWA</td>
<td>11.69</td>
<td>Mean Range Years Edu all PWA</td>
</tr>
</tbody>
</table>

**Table 3: Criteria for Inclusion or Exclusion in Study**

<table>
<thead>
<tr>
<th>Exclusionary Criteria For PWA &amp; PWOA:</th>
<th>Exclusionary Criteria, PWOA Specific:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Failure of the Clock Drawing Test (Indicating any/or other Cognitive Impairments)</td>
<td>Any Individual less than 3 months post CVA</td>
</tr>
<tr>
<td></td>
<td>Inability to correctly answer more than 5% of picture naming task (SLT judged)</td>
</tr>
<tr>
<td></td>
<td>Any Individual aged under 18</td>
</tr>
<tr>
<td></td>
<td>Severe Dysarthria/Dyspraxia that would complicate naming assessment</td>
</tr>
</tbody>
</table>

**4.2 Materials**

The second edition of the BNT was used (Kaplan et al, 2001). Test forms were used from the BNT to score item response, together with forms that established appropriate demographic information, and a version of the Clock Drawing Test (CDT, Shulman 2000, Appendix C). The BNT assessment was audio recorded (with participant permission) by all researchers, and responses were coded and
transcribed in detail after the assessment with the recording. Transcriptions and error descriptions were inter-rated and verified with the same audio recordings.

4.3 Procedure

Four researchers shared the task of meeting all PWOA in their homes on one occasion. PWA were also shared and assessed once, either at the site from which referred, or at their homes. Due to time constraints for both participants and researchers, and to the non-invasive nature of the research, consent forms (Appendix D) were explained and signed immediately before the assessment was carried out. Participants understood that they could cease the assessment at any time.

Researchers then administered the CDT, followed by the 60-item BNT. Correct answers were coded with a tick. All other answers were transcribed verbatim. Stimulus cues were used with marginally increased lenience in comparison with standard suggested protocol, being provided if target items were misperceived, and/or if the client became discouraged or unresponsive during the assessment. Spontaneous, immediate self-corrections were counted as correct. These were classified as occurring after only one error, within three seconds of initial attempt at naming. The initial attempt error was also recorded.

Phonemic cues were given after a client either stated that they didn't know what an item was, or had failed to provide the correct answer after twenty seconds elapsed. At the end of the BNT assessment, the researcher discussed the outcome of the assessment at the participant’s request. Assessment results were also sent on to the referring SLT if requested.

All sixty items of the BNT were used during testing, for the purposes of obtaining error and score data on the three items recommended for removal by earlier research on the BNT in Ireland – ‘Beaver’, ‘Pretzel’ and ‘Dominoes’.
Based on both earlier research and analysis of data for this research, alternative responses deemed culturally acceptable for some items were as follows: ‘Garden Seat’ for Bench (first response N = 70) and ‘Brush’ for Broom (first response N = 62). Subordinated items with a common word stem were also accepted as correct, such as ‘coat hanger’ for ‘Hanger’ and ‘tennis racquet’ for ‘Racquet.’ These decisions were made due to the similar scoring decisions made by other studies that aimed to standardize the BNT for different populations (Tallberg, 2005).

Table 4: Scoring Systems Used

| Out of 60, including self-corrections as correct, |
| Out of 60, not including self-corrections as correct, |
| Out of 57 (excluding 3 above items), including self-corrections as correct. |

4.4 Codification of the Data

Error types were coded on the basis of the verbatim answer given by each participant, in response to all 60 items on the BNT. Coding was designed to correlate with an earlier analysis of the BNT performance of older Irish adults (James, 2012; See Table 5 & Appendix E). Focus was on the first response given, with self-correction errors also noted in the count – these were necessarily preceded by another type of error.

This data was then placed in two data sets. The first data set recorded number of error types by category. The second data set contained data on the first response by each participant to each item. As this data set relied strictly on first response only, data on self-correction was not recorded.

4.5 Inter-rater reliability

In order to establish inter-rater reliability, each researcher shared with the group one of each of the audio recordings from PWA and PWOA that they had collected. A member of the group took these samples and transcribed and coded the error types themselves, before comparing their transcriptions and error codes with the data of the original researcher. In this way, 4 of the 16 PWOA
(25%) and 4 of the 35 PWA (11.4%) were inter-rated. 86.6% agreement on error type classification was found for PWA, and 93.3% for PWOA, suggesting a reliable degree of agreement (Shaughnessy et al, 2009).

4.6 Correct Responses & Error Types
Items 19 (Pretzel), 29 (Beaver) and 35 (Dominoes) had been excluded from earlier research focused on standardizing the BNT on an Irish population, on the basis of disproportionately low correct answers, as seen in Figs 1 & 2. In this study, of all participants, (n = 102), 20 gave correct answers for pretzel (18 PWOA; 2 PWA). 25 were correct for Beaver (20 PWOA, 5 PWA). 43 were correct for Dominoes (35 PWOA, 8 PWA). In line with previous research (James 2012), they were excluded from score analysis.

Items 55 (Sphinx) and 59 (Protractor) also received very low amounts of correct responses from PWOA relative to responses to other items. However, as they were amongst items having lowest frequency and higher expected retrieval difficulty, they acted as good discriminators and were not excluded from analysis.

![PWOA - Correct on First Response](image)

**Fig. 1.** Items 19, 29 & 35 subsequently removed.
Fig. 2. Items 19, 29 & 35 subsequently removed.

Table 5: Defined Error Types

<table>
<thead>
<tr>
<th></th>
<th>Error Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>No Response</td>
</tr>
<tr>
<td>2</td>
<td>Don’t Know</td>
</tr>
<tr>
<td>3</td>
<td>Self Correct</td>
</tr>
<tr>
<td>4</td>
<td>Coordinate</td>
</tr>
<tr>
<td>5</td>
<td>Superordinate</td>
</tr>
<tr>
<td>6</td>
<td>Subordinate</td>
</tr>
<tr>
<td>7</td>
<td>Semantic Association</td>
</tr>
<tr>
<td>8</td>
<td>Circumlocution</td>
</tr>
<tr>
<td>9</td>
<td>Unrelated</td>
</tr>
<tr>
<td>10</td>
<td>Multiword</td>
</tr>
<tr>
<td>11</td>
<td>Phonological Real Word</td>
</tr>
<tr>
<td>12</td>
<td>Phonological Non Word</td>
</tr>
<tr>
<td>13</td>
<td>Neologism</td>
</tr>
<tr>
<td>14</td>
<td>Conduite D’Approche – Achieve</td>
</tr>
<tr>
<td>15</td>
<td>Conduite D’Approche – Lost</td>
</tr>
<tr>
<td>16</td>
<td>Morphological</td>
</tr>
<tr>
<td>17</td>
<td>Perseveration</td>
</tr>
<tr>
<td>18</td>
<td>Visual Misperception</td>
</tr>
<tr>
<td>19</td>
<td>Unrecognised</td>
</tr>
<tr>
<td>20</td>
<td>Gesture</td>
</tr>
<tr>
<td>21</td>
<td>Compound</td>
</tr>
<tr>
<td>22</td>
<td>Mixed</td>
</tr>
<tr>
<td>23</td>
<td>Forgot</td>
</tr>
</tbody>
</table>

5: Results

5.1, Section 1: Score: Normality

Scores for this group were made out of a total of 57, after the exclusion of Beaver, Pretzel and Dominoes. Score analysis was therefore conducted on the 57 possible responses for the group.
To determine if the scores were normally distributed across all the participants (n=102), a Kolmogorov-Smirnov test was conducted on scores correct including self-corrections. Results D (102) p = .005 suggested that the scores were not normally distributed.

To determine normality of distribution of the correct scores between groups, Kolmogorov-Smirnov tests were conducted. Results indicated that distribution for PWOA was not normal D (67) p = .012. A Shapiro-Wilk test (more appropriate for smaller sample sizes) also suggested data for PWOA was not normal, D (67) = .865, p = .000. Distribution of PWA scores was normal D (35) p = .964. A Shapiro-Wilk test also suggested data for PWA was normal, D (35) = .983, p = .865.

**5.2 Score: Gender, Age & Years of Education**

Multiple linear regression analyses were run to determine whether significant relationships existed between age, gender or years of education in PWA & PWOA as two separate groups, using scores that did and did not include self-correction. This method of analysis was chosen due to the lack of normality in the data. Results show all p values for differences between PWA and PWOA < .05: with age and education as co-variants in all four analyses, p = 1.0. With gender, for PWA p = .884 on scores excluding self-correction, and p=.949 including. For PWOA p = .995 excluding self-corrections and p = .379 including.

Non-significance remained true when the demographic variables were assessed cumulatively in a univariate analysis of variance. This analysis showed that for the whole group, the only factor to make a difference was the status of participant (p = .000); that is to say, having or not having a diagnosis of aphasia and LRD was the only factor impacting on total score.
5.3 Score Range

Range of scores obtained showed a distinct overlap, as expected on the basis of earlier research. PWOA obtained minimum scores of 30, and maximum scores of 57. PWA obtained minimum scores of 2, and maximum scores of 56.

Fig 3. Overlap of mean scores across groups

Fig 4. Frequencies of PWA & PWOA obtaining each score across groups.
As seen, Figure 4 represents degree of overlap of scores across groups, and Figure 5 shows the frequencies of scores across groups. These highlight the presence of three outliers in the PWA group, who scored 49, 52 and 56 respectively. Each group did receive greater numbers of scores at either end of the ranges, yet a central overlap remained that totaled 27 participants; 14 aphasics and 13 non-aphasics, in score range of 30 – 43.

5.4 Mean Scores

An Independent Samples T-Test was conducted to compare mean scores for PWA (mean = 27.57, SD = 12.953) and PWOA (mean = 48.76, SD = 6.948). Equal Variance not assumed, t (44.471) = 9.024, p = .001 (two-tailed). The size of difference between group mean scores (mean difference = 21.190, 95% CI: 16.459 – 25.921) was large (ETA squared = 0.537).

Due to the difference in sample size between the two groups, a Mann-Whitney U-test was performed and confirmed a significance difference between the 2 groups (U (102) = 195.000, Z = -6.897, p < .001).

In order to establish the most sensitive and specific possible cut off score distinguishing Irish PWA from PWOA, a Receiver Operating Characteristic (ROC) analysis was carried out on the data. The area under the curve was .917 (95% confidence interval .851 to .983), which does not overlap 0.5, p< .001. This showed that the BNT has relatively good strength in differentiating the two groups (Fawcett, 2006). Results showed that a score of 40.5 was the imperfect but optimal score to distinguish PWA from PWOA, with a sensitivity (true positive rate if scored below) of 86.6% and specificity (true negative rate if scored above) of 85.7%. The score with highest true positive (i.e. most chance of confirming PWA) was 35, with 95% sensitivity; score with highest true negative (i.e. most chance of showing PWOA) was 48, with 92% specificity.
35 PWA produced a total of 1197 errors, or 1120 excluding ‘self correction’ subcategory. Range was between 1 – 20 error subtypes per individual.

67 PWOA produced a total of 899 errors, or 692 excluding ‘self correction’ subcategory. Range was between 0 – 10 error subtypes per individual. Note that 5 PWOA made no errors whatsoever.

A univariate analysis of variance was run to ascertain difference in the mean numbers of error subtypes from PWA & PWOA. Raw data showed PWOA made a mean number of 4.9 subtypes, SD 2.8. PWA made a mean number of 12.1 subtypes, SD 3.7. When corrected for number of errors, PWA used a mean of 8.778 subtypes, 95% CI 7.824 – 9.733; PWOA used a mean of 6.698, 95% CI 6.082 – 7.314. Statistical significance was achieved; p = 0.02.
Fig. 6: Mean number of error types made across groups

Mean number of each type of error for each group is shown in Figure 6. As can be seen, the groups differed both in number of error types produced, and frequency of error types. Two error types were only seen in PWA in the whole group. These were neologisms, and unsuccessful conduite d’approche. The uniqueness of these errors in PWA is consistent with earlier research (James, 2011).

14/35 PWA produced Neologisms: 8 erring once, 5 erring twice and 1 erring three times.
7/35 PWA produced Unsuccessful conduite d’approche: 3 erring once, 1 twice, 1 three times, 1 four times and 1 thirty times.
Frequencies of nine error types were statistically significantly greater in PWA than PWOA. In order to ascertain statistically significant differences, a cross-tabulated analysis was run comparing the frequency of error subtypes used by PWA and PWOA. Due to the difference in size of each group, equal variances were not assumed. Results showed that PWA made significantly more (p<0.05) of the following:

‘Don’t Know’ t (36.057) = 4.763, p < .001; ‘Semantic Association’ t (46.436) = 4.670, p < .001; ‘Circumlocution’ t (41.899) = 5.512, p < .001; ‘Multiword’ t (34.628) = 4.339, p < .001; ‘Mixed’ t (41.336) = 3.298, p = .002; ‘Phonologically Related Non Word’ t (35.939) = 2.783, p = .009; ‘Visual’ t (50.159) = 2.550, p = .014; ‘No Response’ t (34.015) = 2.557, p = .015; ‘Unrelated’ t (39.557) = 2.440, p = .019.

Error types unique to PWA were infrequent enough in the total sample that they did not effectively distinguish PWA from PWOA.

An analysis of range of use of each error subtype by each PWA and PWOA was performed by hand, which showed that overlap of range was still existent between PWA & PWOA.

Of the 14 PWA whose scores overlap with controls, 6 made one of the errors unique to PWA. Errors of these PWA were analyzed in terms of both unique & higher frequency errors for PWA. All of this group’s errors were accounted for by at least 58%, and at most 100% of the PWA’s 2 unique & 9 higher frequency errors.
Table 7: Overlap Group Error Analysis: Overall Errors & PWA Errors

<table>
<thead>
<tr>
<th>Participant</th>
<th>Overall Error No</th>
<th>No. PWA unique &amp; higher frequency errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 – MD</td>
<td>23</td>
<td>14</td>
</tr>
<tr>
<td>4 – NB</td>
<td>29</td>
<td>23</td>
</tr>
<tr>
<td>12 – MT</td>
<td>26</td>
<td>16</td>
</tr>
<tr>
<td>13 – BN</td>
<td>28</td>
<td>19</td>
</tr>
<tr>
<td>14 – PA</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>17 – WS</td>
<td>17</td>
<td>16</td>
</tr>
<tr>
<td>18 – TF</td>
<td>27</td>
<td>25</td>
</tr>
<tr>
<td>21 – DOB</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>22 – TMcD</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>24 – CD</td>
<td>26</td>
<td>26</td>
</tr>
<tr>
<td>25 – RK</td>
<td>28</td>
<td>28</td>
</tr>
<tr>
<td>26 – MBO</td>
<td>19</td>
<td>18</td>
</tr>
<tr>
<td>28 – TD</td>
<td>26</td>
<td>22</td>
</tr>
<tr>
<td>30 – PME</td>
<td>28</td>
<td>19</td>
</tr>
</tbody>
</table>

PWOA were more frequent self-correctors, with the whole group producing a total of 207 self-corrections, or an average of 3.08 self-corrections per participant (compared with 77 & an average of 2.2 from PWA). This did not reach statistical significance, however: t (1.475) = -1.890, p = .064.

Discussion

6.1 Summary

Table 8: Main Findings of the Study

| Some Items of the BNT are inappropriate for the Irish population |
| This study found no correlation between score and age, education or gender |
| BNT detects score difference across groups, although ranges overlap between. |
| Some errors are unique, & some errors are significantly more common, for PWA |

6.2 Naming Agreement & Cultural Appropriateness

Previous studies have found the need to remove or replace test items (Cruice et al, 2000; Patricacou et al 2007, Emily 2011). Other studies have reordered test items in terms of matching frequency and difficulty from the populations being tested (Tallberg, 2005). PWOA in this study did not meet naming agreement
expected of three items relative to their position in the BNT; high numbers of visual misperceptions and ‘don’t know’ answers were given for these items. They were considered culturally inappropriate and removed.

6.3 External Factors Influencing Naming

6.3.A Gender: As much of the literature does support the idea that gender has no effect on word finding, except in cultures where opportunities for education are less equal (Grima & Franklin, 2012), no effect of gender was expected.

6.3.B Age: The literature supports a weak effect of age on naming. No effect was found in this study. This may be due to several reasons; small sample size is one. A minority of significantly older and younger participants prevented sensitive stratification. The ‘snapshot’ effect of a cross sectional study may limit the ability to detect true differences in naming due to age. Future research might benefit from increased sensitivity of age stratification, and a wider, more populated age range of participants.

6.3.C Education: That this study found no effect of education – a well researched and known factor impacting naming – could be due to the following. More aphasic participants had less years of education compared to the PWOA group, who had high average years of education even after additional PWOA with lower YOE were sought & added to the sample. In this sense, a larger sample size & more sensitive stratification could be of more use in future research. Cognitive ability could not be accounted for, and may have been an unknown factor that contributed to naming.

6.4 Comparison of Scores Across Groups

When culturally inappropriate items were included, some PWOA obtained scores well below what would have been expected for a non-aphasic individual, and some PWA obtained scores much higher than might be expected for those having
a diagnosis of LRD. This was seen in results, where three outliers in the PWA group were seen.

Significance differences in mean score were evident and expected, permitting the tentative establishment of a ‘cut off score’ for the group at 40.5. The sensitivity and specificity of this score is imperfect, due to the range overlap and outliers. This score is significantly lower than the suggested scores indicative of normal performance given from the standardization of the BNT (see Table 8).

### Table 8: BNT Original Adult Norms Data (Kaplan et al, 2001)

<table>
<thead>
<tr>
<th>Age Group</th>
<th>N</th>
<th>Education Mean</th>
<th>Education SD</th>
<th>BNT Score Mean</th>
<th>BNT Score SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>18 – 39</td>
<td>21</td>
<td>15.1</td>
<td>2.3</td>
<td>55.8</td>
<td>3.8</td>
</tr>
<tr>
<td>40 – 49</td>
<td>11</td>
<td>15.1</td>
<td>2.5</td>
<td>56.8</td>
<td>3.0</td>
</tr>
<tr>
<td>50 – 59</td>
<td>49</td>
<td>13.5</td>
<td>2.1</td>
<td>55.2</td>
<td>4.0</td>
</tr>
<tr>
<td>60 – 69</td>
<td>56</td>
<td>13.2</td>
<td>2.3</td>
<td>53.3</td>
<td>4.6</td>
</tr>
<tr>
<td>70 - 79</td>
<td>41</td>
<td>13.9</td>
<td>3.0</td>
<td>48.9</td>
<td>6.3</td>
</tr>
</tbody>
</table>

If age and education did not explain score outcomes in the group as a whole, it would not explain the cause of the overlap in score; this is a finding of some earlier research (Grima & Franklin, 2012).

### 6.5 Error Types

Findings here seem to reflect arguments presented as opposite possibilities in earlier research; the PWA made much more of the same error types as PWOA (Nickels, 2001) and also made errors no PWOA made (Cuetos et al, 2002). PWA had a broader range of error subtypes than PWOA.

PWA had a higher mean number of types of errors per person than had PWOA. Since PWA made more errors overall than PWOA, this is not surprising. When the mean number of types errors was adjusted for the number of errors, PWA still had a greater mean number of types of error than PWOA.
PWA made 21 types of error; PWOA made 19 types of error. This shows that error types were common to both groups with the exception of only two errors. In terms of frequency, PWA made circumlocutory errors the most (223); followed by ‘don’t know’ responses (192), semantic associations (124) and co-ordinate (96). Non-responses (91), superordinate (66), multiword (59) & unsuccessful conduite d’approche (42) were the concurrently next most common.

PWOA made the most self corrections (207); co-ordinate (248), circumlocution (147), semantic association (84), superordinate (87), & don’t know (72) were the concurrently next most common.

As can be seen, there a ‘top five’ error types were common to both groups; the qualifying difference between groups being that PWA made significantly more of the error types shared with PWOA. High numbers of ‘don’t know’ and ‘no response’ errors in both groups may be attributable to non-familiarity with some of the items in the BNT. PWA who made atypical errors all had scores under 40.

As can be seen in table six, unique errors and errors of higher frequency in PWA accounted for 85 – 100% of the 9 of the 14 of PWA overlap group’s errors. 6 of the 14 PWOA made the PWA-unique errors. This shows that error analysis could confirm presence of LRD in just under half of the PWA in the overlap group, reducing overlap with PWOA in the group by 43%.

The three outlying PWA were not found to produce any errors preferentially.

6.6 Quality, Quantity, or Both? Conclusions

The results of this study show that neither quantitative nor qualitative analysis will perfectly differentiate normal from abnormal LRD. However, the imperfect differentiation made possible with a scoring system can be supported and informed with additional, qualitative notes on error type.

(1) After controlling for cultural and demographic factors, there may be people
without aphasia who experience LRD to a similar degree to PWA due to factors other than adult acquired aphasia; these may include long term difficulties with expressive language such as an undiagnosed language disorder or delay in childhood; undiagnosed neurological injury or trauma (undiagnosed acquired aphasia), or simple differences in the range of what constitutes normal naming.

(2) Similarly, people with aphasia may experience LRD that is not specific to noun finding, or find that their issue with LRD is supported by the visual cue the BNT supplies. This could also explain the overlap that occurs in total score range, and in error subtype use range.

The existence of a single test that can detect presence or absence of one symptom of aphasia with total, complete efficacy has not yet been found possible. In the health sciences and in medicine, such tests are rare beyond the most basic of diagnostic tasks (Lalkhen & McCluskey, 2008) and where the subject is undergoing tests for neurological conditions it is not uncommon to see two different cut off scores posited for optimum sensitivity and specificity, such as Addenbrooke’s Cognitive Examination, (Mioshi et al, 2006).

That a perfect differentiation of PWA & PWOA was not achieved using the BNT qualitatively, quantitatively or both can therefore be reasonably understood and accepted. Clinicians should not diagnose LRD on the basis of one test alone. Diagnosis is a matter of clinical judgment, structured and supported by diagnostic criteria and assessment tools using all the information available to them from formal and informal assessment in plural contexts.

7. Limitations of the Study

7.1 Clock Drawing Test (CDT)
The CDT was administered to all participants as a crude means of ascertaining healthy cognitive functioning. What had not been anticipated, however, was that this test was difficult to complete for many of the PWA, due to unilateral weakness or paralysis, impaired hand function, and hemianopoeias. Amongst
PWA & PWOA, some participants told assessors that they had an incomplete understanding of how to tell the time from a twelve-hour clock. More often than not, this meant that the CDT assessment was not a useful barometer of cognitive functioning and assessors had to use their own clinical judgment in determining suitability of participants for inclusion in the data.

7.2 Sample Sizes
The sample of PWA collected (35) was slightly more than half the size of the PWOA group (67). This smaller sample size presents some limitations with regards to sample power and accuracy of results. The present study accommodated for this where possible by assuming unequal variance and using analyses designed for comparison of disparate sample sizes. However, future studies where numbers of PWA and PWOA are larger and more equal would be ideal.

7.3 PWA & PWOA
The sample of PWA skewed towards lower years of education, which may have complicated analysis of external effects on naming. Similarly, the majority of PWOA skewed towards higher years of education. Recruitment of PWOA with less years of education involved a less desirable method of participant selection – word of mouth, acquaintance based recruitment, which has several limitations including skew towards healthier, motivated & active volunteers that would not represent an accurate reflection of all Irish people from this background.

7.4 Error Analysis
Earlier research has also shown particular types of mixed error may be unique to aphasics (James, 2012). This research did not complete an analysis of mixed error subtypes, due to the time constraints of the project. Future research that includes detailed data collection concerning mixed errors for every participant would be useful.

Mixed errors were common to aphasics and non-aphasics in this research; however, mixed errors were much more common in the aphasic group, with 15
aphasics producing between one and four mixed errors compared with 6 non aphasics producing one to three mixed errors.

Error classifications were highly specific in this study, and therefore numerous. For example, ‘don’t know’ and ‘no response’ errors were coded separately; these are commonly coded as one error type alone. Future studies may benefit from stratification of errors, layering from highly specific categories into sub groups.

The study classed self-corrections as an error subtype. Technically, self-corrections count as correct answers immediately subsequent to an incorrect answer; as such, they were anomalous amongst the subtypes, especially since first answer errors were also recorded. Future studies might account for this, and code self-corrections separately from error classifications.

7.5 Participant Factors
Both PWA & PWOA in this test were subject to extraneous factors influencing their naming. In particular, several participants lack of confidence in ability to supply the correct answer. One participant explained she resorted to talking around the item, or describing it, despite knowing its name, as a safety tactic. It is an important component of assessment to thoroughly inform the client of the purpose, nature and manner of the task in hand; however, it is also important to stay aware of client anxieties that may remain despite this and influence score and error profile in a potentially unrepresentative way.

8. Conclusions

- BNT Items 19 ‘Pretzel,’ 29 ‘Beaver’ & 35 ‘Dominoes’ should be removed when testing in Ireland.
- PWA & PWOA differ significantly in mean score on the BNT. PWA perform significantly worse.
- Error score cut-offs may serve as a guideline for diagnosis of LRD in Irish PWA, if used alongside additional data.
Participants whose scores overlapped could be partially differentiated through the use of a qualitative error profile analysis.

Certain error types are unique to PWA. Certain error types are made with much greater frequency by PWA.

Use of all information available is necessary to identify LRD in aphasia.

9. Recommendations

A study involving larger samples of Irish PWA & PWOA, more sensitively stratified in terms of demographic variables, would be of use in adding to or exploring further the findings from this study. Examination of other stimulus factors in naming, such as age of acquisition, would also be of potential use in predicting BNT performance in the Irish population.

More research looking at the use of qualitative error analysis in assessing PWA would be merited; in particular, this study could not perform in-depth analysis of mixed type errors, which is an area that could potentially yield more information about models of lexical retrieval.

Further research that looks at patterns of response in other confrontation naming tests, naming in connected speech and also any further research aiming to standardize other speech & language clinical assessments in Ireland, would be welcomed.
10. References


Laine, M. & Martin, N. (2012) Cognitive neuropsychology has been, is, and will be significant to Aphasiology. *Aphasiology*, 26(11), 1362 – 1376


Appendix A – Information about the Research Project For Participants

Quality, Quantity? The Performance of Irish People on the Boston Naming Test

Participant Information Sheet

Student Researcher: Rachel Kennedy, M.Sc Speech & Language Therapy Student
Contact Details: rachelsltkennedy@gmail.com
Research Supervisor: Professor Sue Franklin

- Why is this study happening? People who have a stroke can sometimes have problems with their language afterwards. The name of this condition is aphasia. One thing that a lot of people with aphasia experience is what is called ‘word finding difficulty,’ or ‘anomia.’ This is just like something that happens to all of us, now and then – the experience of knowing you know the name or word for a person or a thing, but cannot recall that name. The difference with anomia is that it happens much more often and causes much more trouble for people with aphasia.

- Okay. So why are you asking people who don't have aphasia to work with you? The Boston Naming Test is the most popular test in the world for diagnosing anomia in people with aphasia. People take the test by looking at sixty different drawings, and naming each drawing. At the moment, it uses a total score to diagnose anomia – judging how many pictures people got right or wrong. We’re not sure this is the best way to use this test in Ireland, so we want to check a couple of things:

  - How well do Irish people recognize all the pictures in the Boston Naming Test?
  - Are these pictures fair to use with Irish people?
  - What score do Irish people with healthy naming abilities get on the test?
  - If a normal person makes a mistake naming (and we all do!) is that mistake similar or different to the kind of mistake somebody with anomia would make?

- What do I need to do? I will meet with you at a time, date and place that suits you best – hopefully somewhere quiet and relaxed. I’ll ask you a couple of questions – these are simple, about your age, your life, the time you spent in school and what kind of work you did after school. After that, I will do a short, quick drawing test with you, and then we’ll go through the Boston Naming Test. I will be interested in the answers you give to each picture, and will be recording you and taking some notes when you’re going through these pictures. Remember that this isn’t a test of you – it’s a test of the test, and it’s a test to find out what is normal all over Ireland, so that we know about people from a range of different careers, experiences and backgrounds.
- **Do I have to take part?** There's absolutely no obligation to take part. If anything you read here makes you uncomfortable, you can ask me, or simply tell me you no longer want to take part. You can withdraw from this research at any time – whether before, during, or after your participation.

- **How do I help by taking part?** You will help us understand how normal, healthy people from Ireland tend to do on the Boston Naming Test, which will then help us understand better how people who’ve had strokes in Ireland might do with the test. You’ll help us improve the way we diagnose people with anomia in Ireland. This means we might be able to improve our assessments, which could lead to us improving the kinds of therapy we give for people with anomia.

- **What disadvantages are there for me?** You’ll need to give approximately an hour of your time – at most – to the research.

- **Is this confidential?** Yes. All information is gathered confidentially. You’re one of many who will be helping us, and nobody’s name is attached to any information.

- **Will you tell me how I’ve done on the test?** If you’d like to hear about your score, I am absolutely happy to tell you about that. What’s important to remember is that this isn’t a test of you. This is testing the test, and discovering what’s normal. Your results represent what’s likely to be normal for people in Ireland. I will also absolutely answer any and all questions you may have before, during and after the research, at any time.

- **How do I find out about the study?** When this research is finished, I will submit it to the University of Limerick in order to get my professional qualification as a speech and language therapist. If you would like to see or hear about the results of the test, contact me via email (address above) after September 2013.

- **Thankyou,** very much, for taking the time to read this sheet and for your consideration in participating. Any help you are able to give will be greatly appreciated.

This research has received approval from the University of Limerick Research Ethics Committee. If you have any concerns at all with regards to this study, please contact: Chairman, Education & Health Sciences, Research Ethics Committee, EHS Faculty Office, University of Limerick, Tel: (061) 234101. Email” ehsresearchethics@ul.ie
Appendix B – Aphasia Friendly Information about the Research

The student researcher working with you is called Rachel Kennedy.

She is a speech & language therapy MSc student at the University of Limerick.

Her supervisor is professor Sue Franklin.

This is some information about the research we are doing.

And how you can help!
What is the research project about?
The research is about picture naming. We are especially interested in the types of errors people make.

What will happen?

You will attend be seen by an SLT student.
It will take 60-90 minutes.

You will:

Draw a clock

Name some pictures "owl"

(Just like in your SLT sessions)

We will record your naming
What will the researchers do with the information?

They will **write reports**

They will **give** them to the **University of Limerick**

This is part of a **Masters of Science Degree**

This report will go into a **journal** for other SLTs to **read**

The results will help SLTs to assess **naming** difficulties

---

Confidentiality and Privacy

All **information** is kept **private and safe**

Your **name** will not be used
Taking Part in the Programme

You are **free** to decide if you want to **take part**

If you **agree** you need to **sign** a **consent form**

You are **free** to **stop** at any time

You **do not** have to give a **reason** for stopping

---

Other Information

You can contact

Sue Franklin, Professor of Speech and Language Therapy

061 234274

**Thank you for reading this!**

If you have any concerns about this study and want to contact someone independent you may contact:

Chairman Education and Health Sciences Research Committee

EHS Faculty Office

University of Limerick

Tel 061 234101

Email: ehsresearchethics@ul.ie
Appendix C – Clock Drawing Test Scoring Instructions

**SHULMAN scoring system**

0 No reasonable representation of a clock
   —No attempts at all
   —No semblance of a clock at all
   —Writes a word or name

1: Severe level of disorganisation as described in 2

2 Moderate visuo-spatial disorganisation of times such that accurate denotation of 10 after 11 is impossible
   —Moderately poor spacing
   —Omits numbers
   —Perseveration
   —Repeats circle or continues on past 12 to 13, 14, 15 etc.
   —Right-left reversal
   —Numbers drawn counterclockwise
   —Dysgraphia
   —Unable to write numbers accurately

3: Inaccurate representation of 10 after 11 when visuospatial organization is perfect or shows only minor deviations
   —Minute hand points to 10
   —Writes ‘10 after 11’
   —Unable to make any denotation of time

4 Minor visuospatial errors
   —A mildly impaired spacing of times
   —Draws times outside circle
   —Turns page while writing numbers so that some numbers appear upside down
   —Draws in lines (spokes) to orient spacing

5: Perfect clock

(Shulman 2000, p 554)
Appendix D – Test Forms; Participant Information

Today's Date: ____________

Participant (circle one) - PWA / PWOA

Participant Initials & Number: ________________

Date of Birth: ________________

Age Starting Education: ________________

Age Leaving Formal Education: ________________

Highest Level Qualification: ________________

Profession(s): ________________

Any Languages Spoken Beside English? ________________

Where Participant is PWA:

Date of most recent CVA: (mm/yy): ________________

Previous CVA? (Circle one) – Y/N

Date(s) of previous CVA(s): (mm/yy): ________________
Appendix D: Test forms; Clock Drawing Test

On the circle below, pencil in the numbers of a twelve-hour clock.

Set the hands of the clock, so that the time reads:
Ten minutes past Eleven.

Participant Initials & No: __________
Appendix D: Test Forms; Scoring Form for the Boston Naming Test

Participant Initials & Number: ____________

<table>
<thead>
<tr>
<th>Target</th>
<th>Response</th>
<th>Stimulus Cue</th>
<th>Phonemic Cue</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Bed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Tree</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Pencil</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. House</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Whistle</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Scissors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Comb</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Flower</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Saw</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Toothbrush</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Helicopter</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. Broom</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. Octopus</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14. Mushroom</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. Hanger</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16. Wheelchair</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17. Camel</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18. Mask</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19. Pretzel</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20. Bench</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21. Racquet</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22. Snail</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23. Volcano</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24. Seahorse</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25. Dart</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26. Canoe</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27. Globe</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28. Wreath</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29. Beaver</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30. Harmonica</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31. Rhinoceros</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>32. Acorn</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>33. Igloo</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>34. Stilts</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35. Dominoes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>36. Cactus</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>37. Escalator</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>38. Harp</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
</tr>
<tr>
<td>39.</td>
<td>Hammock</td>
<td></td>
<td></td>
</tr>
<tr>
<td>40.</td>
<td>Knocker</td>
<td></td>
<td></td>
</tr>
<tr>
<td>41.</td>
<td>Pelican</td>
<td></td>
<td></td>
</tr>
<tr>
<td>42.</td>
<td>Stethoscope</td>
<td></td>
<td></td>
</tr>
<tr>
<td>43.</td>
<td>Pyramid</td>
<td></td>
<td></td>
</tr>
<tr>
<td>44.</td>
<td>Muzzle</td>
<td></td>
<td></td>
</tr>
<tr>
<td>45.</td>
<td>Unicorn</td>
<td></td>
<td></td>
</tr>
<tr>
<td>46.</td>
<td>Funnel</td>
<td></td>
<td></td>
</tr>
<tr>
<td>47.</td>
<td>Accordion</td>
<td></td>
<td></td>
</tr>
<tr>
<td>48.</td>
<td>Noose</td>
<td></td>
<td></td>
</tr>
<tr>
<td>49.</td>
<td>Asparagus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50.</td>
<td>Compass</td>
<td></td>
<td></td>
</tr>
<tr>
<td>51.</td>
<td>Latch</td>
<td></td>
<td></td>
</tr>
<tr>
<td>52.</td>
<td>Tripod</td>
<td></td>
<td></td>
</tr>
<tr>
<td>53.</td>
<td>Scroll</td>
<td></td>
<td></td>
</tr>
<tr>
<td>54.</td>
<td>Tongs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>55.</td>
<td>Sphinx</td>
<td></td>
<td></td>
</tr>
<tr>
<td>56.</td>
<td>Yoke</td>
<td></td>
<td></td>
</tr>
<tr>
<td>57.</td>
<td>Trellis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>58.</td>
<td>Palette</td>
<td></td>
<td></td>
</tr>
<tr>
<td>59.</td>
<td>Protractor</td>
<td></td>
<td></td>
</tr>
<tr>
<td>60.</td>
<td>Abacus</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Acceptable alternative answers for the Irish Population:**
Item 12: Brush
Item 15: Clothes Hanger
Item 20: Garden Seat
Item 21: Tennis Racquet
Item 24: Horsefish
Item 30: Mouth Organ
Item 40: Door Knocker

**Additional Notes:**
Appendix E – Error Types (Adapted Explanation of Classification System from: James, 2012)

<table>
<thead>
<tr>
<th>Error Type</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Don’t Know</td>
<td>Pt states no knowledge of item or item name</td>
<td>Item: Abacus Response: I’ve no idea</td>
</tr>
<tr>
<td>No Response Forget</td>
<td>After 20 seconds, Pt has not said anything</td>
<td>Item: Yoke Response: …. (&gt;20 sec)</td>
</tr>
<tr>
<td></td>
<td>Pt states they know but cannot remember the name of an item</td>
<td>Item: Compass Response: It’s not coming back to me, I know it</td>
</tr>
<tr>
<td>Self-Correction*</td>
<td>Pt immediately corrects an incorrect response</td>
<td>Item: Canoe Response: Boat – Canoe (&lt; 3 sec)</td>
</tr>
<tr>
<td>Co-Ordinate</td>
<td>A response in same semantic category</td>
<td>Item: Rhinoceros Response: Hippo</td>
</tr>
<tr>
<td>Sub-Ordinate</td>
<td>A response more specific than the target</td>
<td>Item: Volcano Response: Mount Vesuvius</td>
</tr>
<tr>
<td>Super-Ordinate</td>
<td>A response broader than desired to describe target</td>
<td>Item: Beaver Response: Rodent</td>
</tr>
<tr>
<td>Semantic Association</td>
<td>A one word response semantically associated to item</td>
<td>Item: Compass Response: Maths</td>
</tr>
<tr>
<td>Circumlocution</td>
<td>Multiword response describing target specifically</td>
<td>Item: Pyramid Response: It’s in Egypt</td>
</tr>
<tr>
<td>Unrelated</td>
<td>Real word unrelated to item</td>
<td>Item: Sphinx Response: Needle</td>
</tr>
<tr>
<td>Multiword</td>
<td>Multiple words not specifically describing the item</td>
<td>Item: Tongs Response: Maria has one</td>
</tr>
<tr>
<td>Phonologically related real word</td>
<td>Real word bearing phonological similarity to item</td>
<td>Item: Stilts Response: Sticks</td>
</tr>
<tr>
<td>Phonologically related non word</td>
<td>Illegal (distorted) word phonologically similar to item</td>
<td>Item: Unicorn Response: /judikəl/</td>
</tr>
<tr>
<td>Neologism</td>
<td>Non-word &lt;50% target phonemes</td>
<td>Item: Stethoscope Response: /aps/</td>
</tr>
<tr>
<td>Successful Conduite D’Approche</td>
<td>Repeated approximations/parts of a word resulting in correct word</td>
<td>Item: Camel Response: /kam...ka...kam...kamal/</td>
</tr>
<tr>
<td>Unsuccessful Conduite D’Approche</td>
<td>As above, except with no attainment of correct word</td>
<td>Item: Helicopter Response: /halkɔpterə...halkɔpterə...hελικσαπα:/</td>
</tr>
<tr>
<td>Morphological Alteration</td>
<td>Alteration of correct structure</td>
<td>Item: Pyramid Response: Pyramids</td>
</tr>
<tr>
<td>Perseveration</td>
<td>Repetition of earlier item/used word</td>
<td>Item: Sphinx Response: Pretzel</td>
</tr>
<tr>
<td>Visual Misperception</td>
<td>Misunderstanding of what an item is</td>
<td>Item: Pretzel Response: Snake</td>
</tr>
<tr>
<td>Unrecognised</td>
<td>Non-recognition of correct naming of an item</td>
<td>Item: Beaver Response: Beaver, no, Rat</td>
</tr>
<tr>
<td>Gesture</td>
<td>Non-verbal response</td>
<td>Item: Comb</td>
</tr>
<tr>
<td>------------------</td>
<td>---------------------</td>
<td>-------------</td>
</tr>
<tr>
<td></td>
<td>Response: Pt mimics</td>
<td>Response: combing hair</td>
</tr>
<tr>
<td>Compound Word Error</td>
<td>Error that is compound specific</td>
<td>Item: Briefcase</td>
</tr>
<tr>
<td></td>
<td>Response: Riskcase</td>
<td>Response:</td>
</tr>
<tr>
<td>Mixed Error</td>
<td>Combination of 2+ errors</td>
<td>Item: Trellis</td>
</tr>
<tr>
<td></td>
<td>Response: Lattice (Semantic &amp; Phon Real Word)</td>
<td>Response:</td>
</tr>
</tbody>
</table>