The psychometric properties of the Dutch Naming Test II

Studied in an aphasic population

Name student: C.M. de Hilster (Carlijn)

Student number: 4197291

Course: Afstudeeronderzoek

Status: Final version

Date: 26th of June 2017

Number of words: 3672

Reference style: Vancouver Reference Style Guide

Course mentor: Dr. Lizet van Ewijk

Course lecturer: Dr. Rob Zwitserlood

Course examinator: Dr. Lianne Remijn

Organisation of internship: Lectoraat Hogeschool Utrecht

Intended journal: Brain and Language

Maximum words intended journal: 10.000

Clinical Language, Speech and Hearing Sciences, master Clinical Health Sciences, Faculty of Medicine, Utrecht University, the Netherlands

LIST OF ABBREVIATIONS AND RELEVANT DEFINITIONS

AoA	**Age of Acquisition**
BNT	**Boston Naming Test**
CVA	**Cerebrovascular accident**
DNT	**Dutch Naming Test**
IC	**Informed Consent**
ICC	**Interclass Correlation Coefficient**
K	**Kappa**
PALPA	**Psycholinguistic Assessments of Language Processing in Aphasia**
r	**Correlation Coefficient**
SAT	**Semantische Associatie Test**
SLT	**Speech Language Therapist**
WF	**Word Frequency**
WMO	**Medical Research Involving Human Subjects Act (in Dutch: Wet Medisch-wetenschappelijk Onderzoek met Mensen)**

Samenvatting

Achtergrond: Veel mensen met afasie hebben last van woordvindingsproblemen. Op dit moment is er in Nederland geen betrouwbaar en valide instrument beschikbaar om woordvindingsproblemen vast te stellen. Om deze reden werd in 2015 gestart met het ontwikkelen van een nieuwe test: de Dutch Naming Test (DNT). Na een pilotstudie in 2016, werd deze test aangepast en getest op gezonde volwassenen. In deze studie wordt de aangepaste versie van de test, de Dutch Naming Test II (DNT-II), voor het eerst getest bij personen met afasie.

Doel: Het doel is om vast te stellen of de DNT-II een betrouwbaar en valide instrument is om woordvindingspoblemen vast te stellen bij Nederlandse volwassenen met afasie.

Methoden en procedures: Om de validiteit en betrouwbaarheid van de DNT-II vast te stellen, werden 50 patiënten met afasie in de revalidatie- of chronische fase onderzocht. Participanten werden gerekruteerd vanuit ziekenhuizen, revalidatiecentra, verpleeghuizen en eerstelijnspraktijken gevestigd in heel Nederland. Bij de participanten werden de DNT-II, de Boston Benoemtaak, het verbale deel van de Semantische Associatie Test en PALPA-taak 14 afgenomen.

Resultaten: De DNT-II blijkt een goede gezichtsvaliditeit en constructvaliditeit te hebben. Test-hertest betrouwbaarheid en interbeoordelaarsbetrouwbaarheid werden beiden excellent bevonden. De DNT-II werd positief beoordeeld door logopedisten en patiënten.

Conclusie: Deze psychometrische studie laat zien dat de DNT-II een betrouwbaar, valide en gebruiksvriendelijk instrument is om woordvindingspoblemen vast te stellen bij personen met afasie.

Implicaties en aanbevelingen: De DNT-II is het eerste valide en betrouwbare instrument dat beschikbaar is om woordvindingspoblemen vast te stellen bij personen met afasie. Vervolgonderzoek zal zich moeten richten op het onderzoeken van de prestatie op de DNT-II van gezonde volwassenen en personen die een beroerte hebben doorgemaakt maar geen afasie hebben. Wanneer norm data beschikbaar is, kan de DNT-II worden gebruikt in de klinische praktijk.

Kernwoorden: Dutch Naming Test II (DNT-II), afasie, woordvindingspoblemen, volwassenen

Abstract

Background: Anomia is one of the most common problems in patients with aphasia. However, in the Netherlands, there is no valid and reliable instrument available to diagnose anomia properly. For this purpose, the development of the Dutch Naming Test (DNT) was started in 2015. After a pilot study, this test was adjusted and tested on healthy people in 2016. In this study the adapted version of this test, the Dutch Naming Test II (DNT-II), will be tested in an aphasic population.

Aims: This study aims to examine whether the DNT-II is a reliable and valid tool to diagnose naming difficulties in Dutch patients with aphasia.

Methods: The reliability and validity of the DNT-II was established by investigating 50 Dutch adults with aphasia in the rehabilitation or chronic phase. Subjects were included from hospitals, rehabilitation centres, nursing homes and private practices throughout the Netherlands. The subjects underwent the DNT-II, The Boston Naming Test, PALPA task 14 and the verbal part of the Semantic Association Test.

Results: The DNT-II was found to have a good face validity, a good construct validity, an excellent interrater reliability and an excellent test-retest reliability. The DNT-II was evaluated positively by speech language therapists and participants.

Conclusion: The DNT-II has shown to be a reliable, valid and user-friendly tool to diagnose anomia in patients with aphasia.

Implications of key findings: The DNT-II is the first valid and reliable tool to diagnose anomia in Dutch adults with aphasia. Further research into the performance of healthy adults and stroke patients without aphasia on the DNT-II should be conducted. When norm data is available, the test is ready to be used in the clinical practice.

Keywords: Dutch Naming Test II (DNT-II), aphasia, anomia, adults

Introduction

In the Netherlands, over 40.000 people are struck by stroke annually. Stroke, also referred to as cerebrovascular accident (CVA), can cause several physical and psychological disabilities, resulting in drastic changes in the life of the affected person (1). A frequent occurring consequence of stroke is aphasia, a communication disorder that impairs a person's ability to produce and comprehend language. The incidence of aphasia in the stroke population is estimated between 21 and 38% (2-4). One of the most striking problems of patients with aphasia, is the inability to retrieve words from the lexicon adequately, also known as naming difficulties or anomia (5).

Naming is a complex activity, which involves a multicomponent processing system (6). There are several language production models that address the different steps involved in naming. Although these models differ in many aspects, most of them agree that naming consists of four main processing steps: conceptual preparation, lexical selection, word form encoding, and articulation (6-8). The first step is a preverbal step, resulting in a preverbal message. The second step, a semantic step, involves mapping meaning onto an abstract representation of a word. In the third step, a phonological step, this abstract lexical representation is mapped onto its phonology. Finally, the word is articulated in the last step (9). This study focuses on the semantic and the phonological step (i.e. step two and three).

There is a variety of cognitive linguistic factors that contribute to the speed and accuracy of naming (10). Research has shown that the age of acquisition (AoA) and word frequency (WF) both have a fundamental influence on speed and accuracy of naming. Words acquired early in life and words that occur more frequently in a language are generally easier to name (11). The effect of AoA may be due to the organisation of the semantic system, such that early acquired words are richer, more accessible and more robust against brain damage (12). WF is thought to be associated with the phonological step of naming: the word form encoding (13-15). Based on these findings, AoA and WF both seem to influence a different step of the word retrieval process. However, recent studies emphasize the high correlation between AoA and WF, and question the isolated effect of WF (12).

Assessment of anomia typically involves picture naming. Research has shown that scores on picture naming tests can be related to the ability to retrieve nouns in everyday conversations for people with aphasia (16). In the Netherlands, the guideline for the diagnosis and treatment of aphasia (17) recommends using the Boston Naming test (BNT) for this purpose. The BNT is a picture naming test that was developed in the US in 1983. However, the diagnostic value of the BNT has been criticized, since evaluation studies of the BNT have shown that the test has poor psychometric properties, no adequate standardization, and

inadequate norms (18). Moreover, the Dutch version of the BNT is not tailored to the properties of the Dutch language, since it is a literal translation from the American version.

The lack of a proper diagnostic tool for naming difficulties in the Netherlands initiated the development of a new test: the Dutch Naming Test (DNT) (19). In this test, the word frequency of the items is adapted to the Dutch language. A pilot study in aphasic patients showed that the first version of the DNT was not sensitive enough to diagnose anomia properly (20). Therefore, in 2016, a new version of the DNT (the DNT-II) was developed by Jessurun (21). In the new version, several adjustments have been made, with the most important adjustment being the inclusion of the AoA as a predictor value. A pilot study with healthy adults showed that the DNT-II, including both variables AoA and WF as predictor values, has the potential to be a sensitive diagnostic tool to diagnose naming disorders in Dutch adults. In addition, the authors stated the hypothesis that the inclusion of both cognitive linguistic factors AoA and WF ensures that the DNT-II also has the ability to determine which step/steps of the word retrieval process is/are impaired (21).

In this study, the DNT-II will be tested in an aphasic population for the first time. The aim is to test whether the DNT-II is a valid and reliable tool to diagnose anomia in Dutch aphasic patients. Additionally, the hypothesis about the test's ability to distinguish between impairments in different steps of the word retrieval process will be tested.

Aim

The primary objective of the present study is to examine whether the DNT-II is a reliable and valid tool to diagnose naming difficulties in Dutch patients with aphasia in the rehabilitation or chronic phase. For this purpose, the face validity, construct validity, interrater reliability, and test-retest reliability will be determined.

The secondary objective of this study is to examine whether the DNT-II is able to distinguish between underlying semantic and phonological deficits.

Method

Study design

This study concerns a psychometric cross section research in which the face validity, construct validity, interrater reliability, and test-retest reliability of the newly developed DNT-II were determined.

Population & domain

Dutch adults diagnosed with aphasia in the rehabilitation or chronic phase were sampled. Patients in both phases were included, since reliability studies should be performed in a representative sample of the population the test is meant for (22).
In order to be eligible to participate in this study, subjects had to be: aged over 18 years; native Dutch speakers; diagnosed with aphasia (token test score =>7 or ScreeLing score <68); in the rehabilitation or chronic phase (more than two weeks after the occurrence of aphasia). Patients were excluded from participation in this study in case of serious or terminal comorbidity, visual problems resulting in the inability to recognise pictures, or the inability to give informed consent.

Data collection

Speech language therapists (SLTs) working with patients that are eligible for participation in this study, were contacted by the researcher in January 2017. A total of 27 SLTs cooperated in this study. Data collection took place between February and May 2017. Participants were recruited from hospitals, rehabilitation centres, nursing homes, and private SLT practices throughout the Netherlands.

Instruments

In order to conduct validity and reliability analysis, test administration consisted of the DNT-II, the BNT, the verbal part of the Semantische Associatie Test (SAT) (23) and the Psycholinguistic Assesments of Language Processing in Aphasia (PALPA) task 14 (24).
The BNT was used for construct validity analysis. The SAT-verbal and PALPA 14 were used to assess the ability of the DNT-II to distinguish between phonological and semantic underlying deficits. The SAT was used to assess semantic deficits. Psychometric research has shown that this test has the ability to detect semantic deficits adequately (23).
PALPA task 14 was used to assess phonologic deficits. This task was used since it is the only task that measures phonology without asking patients to speak, so test scores cannot be biased by semantic problems. It is important to note that the psychometric properties of this task were never determined, therefore conclusions based on the score on this test should be taken with caution (24).

The DNT-II

The DNT-II consists of 92 coloured pictures. These items can be divided into six categories by combining AoA (early and late) with word frequency (high, mid and low). Low frequency words have a maximum frequency count of five per million, mid frequency words have a frequency count between 5 and 50 per million, and high frequency words have a frequency count of at least 50 occurrences per million. Words were defined as early acquired when they were acquired at the age of five years or earlier. Words acquired at a later age than five years were defined as late acquired (21).

At the baseline of this study, the sequence of the items was not fixed and the scoring system consisted of only “correct” and “incorrect” scores. Prior to the first test administration, the sequence of items was defined. The items were sequenced based on their WF and AoA values, with an increasing difficulty. Prior research showed that AoA has a greater influence on naming latencies than WF in healthy adults (21). Therefore, the value of AoA was weighted heavier than the value of WF.

A pilot test administration showed that the used scoring system was not sufficient, since the major part of the aphasic patients’ responses could not be scored as “correct” or “incorrect”. Therefore, a new scoring system was developed in which also “partly correct” scores were included. The new scoring system is based on the scoring system of the BNT and consists of an ordinal 4-point scale (see appendix A).

Procedures

The tests were administered by the main researcher, one of the research assistants or the SLT in one or two sessions. The DNT-II items were displayed on a laptop. Participants were instructed to name the displayed picture in one word. The SAT-verbal, PALPA 14 and the BNT were administered following the instructions stated in the manuals of these tests. The administration of the DNT-II and the BNT was recorded on a voice recorder. The test administration of all the tests was standardized by a test protocol (see appendix B). SLTs were asked for their opinion about the DNT-II after the test administration, in case they were present during the administration of the DNT-II or administered the DNT-II themselves. If possible, patients were also asked for their opinion.

Data analysis

Data was analysed using IBM SPSS Statistics 24 (25).

Face validity

Face validity was estimated as the degree to which the DNT-II looks as though it is an adequate reflection of the construct to be measured. There are no standards regarding how it should be assessed, and it cannot be quantified by using statistical analyses (22).

Construct validity

The construct validity was determined using correlation analysis between the mean score on the DNT-II and the mean score on the BNT.

Test-retest reliability

The test-retest reliability, was determined by using the interclass correlation coefficient (ICC). An ICC ≥ 0.75 was found to be acceptable, according to the cut-off values of standard psychometric methods stated by Hilari and colleagues (26).

Interrater reliability

To evaluate the interrater reliability, the samples of ten participants on the DNT-II were scored independently by a second reviewer based on the audio recordings. The level of agreement between the two reviewers was determined by calculating the weighted Kappa. It was preferable to use Weighted Kappa over Cohen's Kappa, since it also takes the degree of disagreement into account, which is relevant when ratings are ordered (27). In general, a Weighted Kappa of 0.60 or higher is found to be acceptable (28).

Word Frequency and Age of Acquisition:

At first, correlation analyses for the specific WF and AoA-values and the mean score per item were conducted to assess WF and AoA effects on group level. After this, a Kruskal Wallis test was performed to assess the differences between the mean scores on items for the three different groups of word frequency (high, mid and low). When a significant difference was found, a Mann-Whitney U Test was performed to make pairwise comparisons between the subgroups. A Mann-Whitney U Test was also performed to assess differences between the mean scores on items and the two different groups of AoA (early and late). A Bonferroni correction was done to correct for alpha inflation.

The presence of WF and AoA effects were also assessed per participant. Therefore, a correlation analysis was performed for the WF- and AoA-values and the scores on the items for each participant individually. The outcome was dichotomized in the effect being present (i.e. a significant correlation) or absent (i.e. no significant correlation). The scores on the SAT-verbal and PALPA 14 were also dichotomized, with scores being either deviant (i.e. below the cut-off score as stated in the manual) or normal (i.e. not below the cut-off score). Finally, Pearson's chi square was used to test a possible relation between WF and AoA effects and semantic and phonological deficits.

Ethical issues

This study was conducted according to the principles of the Declaration of Helsinki (29) and in accordance with the Medical Research Involving Human Subjects Act (Wet Medisch Wetenschappelijk Onderzoek Met Mensen). Participants gave written informed consent

before participation. The information letter and informed consent form were adjusted for aphasic patients (see appendix B). Handling of personal data in compliance with the Dutch law for protecting personal data (Wet Bescherming Persoonsgegevens).

Results

Participants

A total of 50 patients participated in this study. The participants' baseline characteristics are shown in table 1. Due to unexpected transfers of participants or participant frustration, a few tests could not be administered. Eventually, all participants underwent the DNT-II, 49 underwent the BNT, 48 underwent the SAT-verbal and 48 underwent PALPA 14. Three participants underwent a second administration of the DNT-II seven days after the first administration to allow for test-retest analysis.

Table 1: Baseline characteristics of the participants

Participants (*n*= 50)	**n**	**Percentage**
Gender		
Female	31	62%
Male	19	38%
Handedness		
Right-handed	34	67%
Left-handed	6	14%
Ambidextrous	4	9%
Unknown	7	14%
Level of education		
Elementary school, household school[1]	14	28%
Lower vocational education[2]	3	6%
Middle school[3]	11	22%
Junior high school or middle vocational education[4]	6	12%
Senior high school or higher vocational education[5]	8	16%
University[6]	4	8%
Unknown	4	8%
Type of stroke		
Infarction	26	51%
Haemorrhage	14	28%
Both (infarction and haemorrhage)	1	2%
Tumour	1	2%
Unknown	9	8%
Clinical localization of stroke		
Left hemisphere	42	82%
Right hemisphere	2	4%
Stem	1	2%
Both hemispheres	1	2%
Unknown	5	10%
Age in years, mean (SD) [range]	71 (12) [38-93]	
Time post-onset in months, mean (SD) [range]	26 (36) [1-155]	

Note.
[1] Lagere school, huishoudschool (HH)
[2] Lager Beroepsonderwijs (LBO)
[3] Middelbaar Voortgezet onderwijs (VMBO)
[4] Hoger Voortgezet Onderwijs (HAVO), Middelbaar Beroepsonderwijs (MBO)
[5] Voorbereidend Wetenschappelijk Onderwijs (VWO), Hoger beroepsonderwijs (HBO)
[6] Wetenschappelijk onderwijs (WO)

Psychometric properties

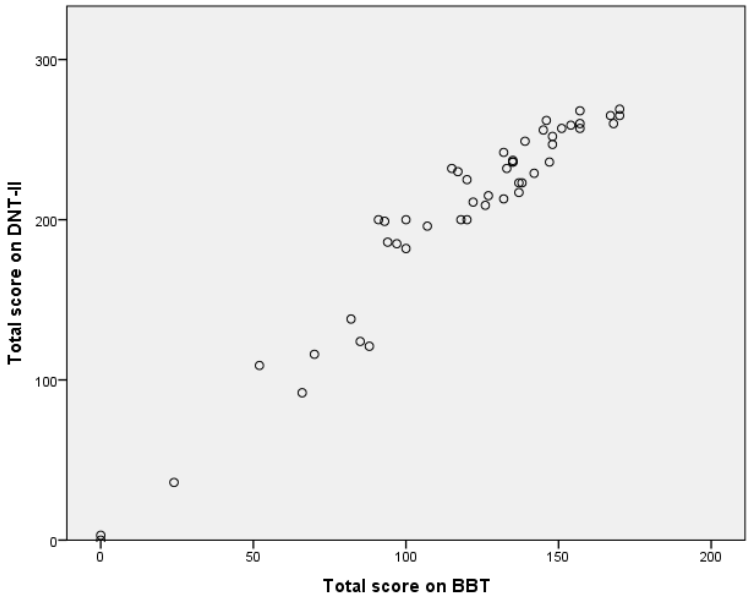
Face validity

The face validity was found to be good, since all test administers affirmed that the performance on the DNT-II looks as though it is an adequate reflection of the presence and severity of the word finding difficulties.

Construct validity

Total scores on the DNT-II correlated significantly with total scores on the BNT. The correlation coefficient (r= 0.965) shows a strong positive relation between those variables (see figure 1). These outcomes suggest a good construct validity.

Figure 1: Correlation between DNT-II and BNT scores; a scatterplot.



Test retest reliability

The interclass correlation coefficient between the single items of the participants that accomplished the DNT-II twice (n=3) was found to be excellent (ICC=0.873).

Interrater reliability

The weighted Kappa showed an almost perfect agreement between the two reviewers per item (weighted K=0.961).

Frequency and AoA effects on group level

Correlation analyses showed a weak positive relation between WF-values and the mean score per item (r=0.272), and a moderate negative relation between AoA-values and the mean scores per item (r=-0.534). Both correlations were significant ($p<0.05$).

The Mann-Whitney U Test showed a significant difference between the two AoA groups (p=0.000). The Kruskal Wallis test showed a significant difference in mean scores between the three different WF groups (p=0.007). Post hoc pairwise comparisons with the Mann-Whitney U Test showed a significant difference between the mean scores for the high frequency and low frequency subgroups of words. No significant differences were found between high frequency and mid frequency, and low frequency and mid frequency subgroups of words (see table 2).

Table 2: Influence of word frequency on mean ranks of mean scores

Word frequency	N	Mean rank	U-statistic	Z	p
HF-MF	33-30	36.18 – 27.40	357.000	-1.900	0.057
HF-LF	33-29	37.77 – 24.36	271.500	-2.921	0.003*
MF-LF	30-29	33.60 – 26.28	327.000	-1.638	0.101

Note.
Correlation is significant at the 0.01 level (2-tailed): 0.003 (Bonferroni correction)
HF=High word frequency (>50 occurrences per million); MF = Mid word frequency (5-50 occurrences per million); LF = Low word frequency (<50 occurrences per million).

Frequency and AoA effects on participant level

A significant correlation between WF-values and item scores was found for 8 participants (WF effect), whilst a significant correlation between AoA-values and item scores was found for 31 participants (AoA effect). Pearson's chi square test showed no statistically significant association between an AoA effect and a semantic underlying disorder (p=0.082). Also, no statistically significant association between a WF effect and a phonologic underlying disorder was found (p=0.173).

Experiences of SLTs and patients

A total of ten SLTs and ten patients were systematically asked for their experience with testing or being tested with the DNT-II. All patients indicated to prefer the DNT-II to the BNT, because the images were more up-to-date and could be distinguished more easily. Comparing the administration of the DNT-II and the BNT, there was no difference in experienced burden, despite the larger number of items of the DNT-II.
SLTs experienced the DNT-II as an improvement in diagnosing anomia in the clinical practice; the test was said to be user-friendly for the SLTs. The SLTs also felt that the use of digital, coloured pictures was profitable for the patients. Some items were found to be unclear or vague in patient and SLT opinions. Based on these experiences combined with the scores on these items, some recommendations for adjustments of the DNT-II were made.

Discussion

This study is the first validation study of DNT-II, which was developed because of the lack of a valid and reliable tool to diagnose naming difficulties in Dutch adults with aphasia. In this study, face validity, construct validity, interrater reliability and test-retest reliability were determined by studying 50 patients with aphasia. With all the measured psychometric properties turning out to be good to excellent, the DNT-II proved to be a valid and reliable tool to assess naming difficulties in Dutch aphasic patients in the rehabilitation or chronic phase.

The secondary aim of this study was to test the hypothesis whether the DNT-II is able to distinguish between underlying semantic and phonological deficits. Previous research stated that AoA and WF both seem to influence a different step of the word retrieval process. AoA is thought to be associated with the sematic step (12), while WF is thought to be associated with the phonological step (13-15). It was therefore hypothesized that patients with semantic disorders will show an AoA effect, while patients with phonological disorders will show a WF effect (21). Although AoA and WF effects were found on group and participant level, no significant association between these effects and an underlying semantic or phonological disorder was found. Thus, this study does not confirm the stated hypothesis. Therefore, no conclusion about the underlying deficit of anomia can be drawn based only on the results of the DNT-II.

The contradiction of the results of this study and the statements made in previous studies, may also be due to the nature of the tests that were used to determine semantic and/or phonological underlying disorders (i.e. PALPA 14 and SAT-verbal). Especially the conclusions based the score on PALPA 14 should be taken with caution, since this task was never assessed for its psychometric properties. The use of this test can therefore be seen as a limitation of this study. However, at this time, there is no other appropriate tool available that assesses phonological deficits in isolation.

The analysis of the mean scores on the items for the different groups of AoA- and WF-value, showed that both AoA and WF significantly affect the scores on the items. These findings correspond to prior research, which indicated that the AoA and WF both have a fundamental influence on the speed and accuracy of naming (11). However, recent studies emphasize the high correlation between AoA and WF, and question the isolated effect of WF (12). By taking both factors, AoA and WF, into account, this study was able to demonstrate an isolated effect of WF on the accuracy of naming. However, this effect was found to be weak. AoA has shown to be of greater influence than WF, which is in line with the findings of previous research in healthy adults (21). These findings are of great importance for the current clinical

practice, where the main focus is on the frequency of words. This study emphasizes the importance of including AoA as a factor in a picture naming test.

During this study, the scoring system was evaluated. The original scoring system was adapted directly after a pilot study with a test patient, since aphasic patients showed a wider range of answers that could not be rated as completely “correct” or “incorrect”. The new scoring system has proven to be manageable for SLTs and the interrater reliability has shown to be excellent. In the current scoring system, only responses of the healthy persons that participated in previous research are listed as examples for the subcategories of the scores. In this study, the responses of 50 patients with aphasia were collected. These responses can be used to extend the examples of the scoring system, in order to facilitate the scoring of the DNT-II for SLTs.

In this research, 50 participants with aphasia were studied, which meets the recommended minimum for validation studies (22). Studying a large group of patients with a large variety of age, etiology, level of education, gender, and post-onset time, results in a possible generalization to a wider population. The test has shown to be user- and patient-friendly, also for the elderly.

A limitation of the present study is that the discriminative power of the DNT-II was not examined, since subjects were only included if they were diagnosed with aphasia. Future research should compare the performance on the DNT-II of aphasic patients with the performance of healthy adults and stroke patients without aphasia. Another limitation of this study is that no item specific analyses were conducted. It is therefore recommended to test the items for name and image agreement in future research.

At last it must be stated that the group of patients used for test-retest analysis may have been too small. Based on the results of this limited group, the test-retest reliability seems to be very good. However, it is desirable to examine test-retest reliability in a larger group of patients to increase the reliability of these findings.

This study was the first to test the newly developed DNT-II in the population the test is designed for. It can be concluded that the DNT-II has the potential to be a valuable addition in diagnosing anomia. Due to the involvement of patients, SLTs, as well as several researchers, a wide range of experiences was acquired. These experiences from different points of view can be useful for further test development.

Conclusion

The DNT-II has shown to be a reliable, valid and user-friendly tool to diagnose anomia in patients with aphasia. The DNT-II is the first tool available for this purpose in the Netherlands. Therefore, it is a valuable addition to the current clinical practice, since an adequate diagnosis is an essential basis for therapy.

Recommendations

The next step in the development of the DNT-II is to study the performance of a large group of healthy adults and stroke patients without aphasia on the test, in order to determine the sensitivity and specificity of the test and assemble norm data. In addition, the scoring system should be extended based on the results of this study. It is also recommended to develop correction tables for age and level of education. Finally, the items that were found to be vague should be appraised critically of name and image agreement. After these steps are taken, a test manual can be developed, so that the newly developed DNT-II can be used in the clinical practice.

References

(1) Kwaliteitsinstituut voor de Gezondheidszorg CBO. (2008). Richtlijn beroerte. Geraadpleegd van www.diliguide.nl/document/230/file/pdf/

(2) Dickey L, Kagan A, Lindsay M, Fang J, Rowland A, Black S. Incidence and profile of inpatient stroke-induced aphasia in Ontario, Canada. Arch Phys Med Rehabil. 2010; 91: 196202.

(3) Engelter S, Gostynski M, Papa S, Frei M, Born C, Ajdacic-Gross V, Gutzwiller F, Lyrer PA. Epidemiology of aphasia attributable to first ischemic stroke. Stroke. 2006; 37: 13791384.

(4) Pedersen PM, Jorgensen HS, Nakayama H, Raascha HO, Olsen TS. Aphasia in acute stroke: incidence, determinants, and recovery. Annals of Neurology. 1995; 38: 659-666.

(5) Raymer A. Naming and word-retrieval problems. In L. L. LaPointe (Ed.), Aphasia and related language disorders. Third ed. New York: Thieme; 2005.

(6) Nozaria N, Kittredgea AK, Della GS, Schwartzb MF. Naming and repetition in aphasia: Steps, routes, and frequency effects. Journal of Memory and Language. 2010; 63(5): 541559.

(7) Levelt WJM. Models of word production. Trends in Cognitive Sciences. 1999; 3(6): 223232.

(8) Indefrey P, Levelt WJM. The Neural Correlates of Language Production. In M. Gazzaniga (Ed.), The new cognitive neurosciences. Cambridge, MA: MIT Press; 2000.

(9) Friedmann N, Biran M, Dotan D. Lexical retrieval and its breakdown in aphasia and developmental language impairment. In: Grohman KK. The Cambridge Handbook of Biolinguistics: Cambridge University Press; 2013.

(10) Bonin P, Chalard M, Méot A, Faylo M. The determinants of spoken and written picture naming latencties. The British Psychological Society. 2002; 93: 89-114.

(11) Meschyan G, Hernandez A. Age of acquisition and word frequency: Determinants of object-naming speed and accuracy. Memory & Cognition. 2002; 30(2): 262-269.

(12) Brysbaert M, Ellis AW. Aphasia and age-of-acquisition: Are early-learned words more resilient? Aphasiology. 2015; 30(11): 1240-1263.

(13) Nozari N, Kittredge A, Dell G, Schwartz M. Naming and repetition in aphasia: Steps, routes, and frequency effects. Journal of Memory and Language. 2010; 63(4): 541-559.

(14) Dell GS. Effects of frequency and vocabulary type on phonological speech errors. Language and Cognitive Processes. 1990; 5: 313–349.

(15) Jescheniak JD & Levelt WJM. Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. Journal of Experimental Psychology: Learning, Memory, and Cognition. 1994; 20: 824–843

(16) Herbert R, Hickin J, Howard D, Osborne F, Best W. Do picture-naming tests provide a valid assessment of lexical retrieval in conversation in aphasia? Aphasiology 2008; 22(2): 184-203.

(17) CBO. Richtlijn Diagnostiek en Behandeling van afasie. Rotterdam: Vereniging van Afasietherapeuten. 2015.

(18) Harry A, Crowe S. Is the Boston Naming Test still fit for purpose? Clinical Neuropsychology 2014; 28(3): 486-504.

(19) Wijngaarden G. The influence of Word Frequency on Accuracy and Latency in a Newly Developed Dutch Naming Test in Heathy Adult Participants [unpublished master's thesis]. Utrecht: Utrecht University; 2015.

(20) Broeder A. De Nederlandse Benoemtaak [unpublished master's thesis]. Groningen: The Faculty of Arts of Groningen; 2015.

(21) Jessurun MY. A picture naming test for Dutch adults: DNT-II [unpublished master's thesis]. Utrecht: Utrecht University; 2016.

(22) Vet de HW, Terwee CB, Mokkink LB & Knol DL. Measurement in Medicine. Cambridge: Cambridge University Press; 2011.

(23) Visch-Brink E, Stronks D, Denes G. Semantische Associatie Test. Amsterdam: Pearson; 2005.

(24) Bastiaanse R, Bosje M, Visch-Brink EG. PALPA, vertaling in, aanpassing aan en normering voor het Nederlands van de PALPA, Psycholinguistic Assesments of Language Processing in Aphasia. Hove: Lawrence Erlbaum Associates; 1995.

(25) De Vocht A. Basishandboek SPSS 24: IBM SPSS statistics. 2nd edition. Utrecht: Bijleveld Press; 2013.

(26) Hilari K, Byng S, Lamping D, and Smith S. Stroke and aphasia quality of life scale-39 (SAQOL-39): evaluation of acceptability, reliability and validity. Stroke 2003: 34(8): 1944-1950.

(27) Cohen J. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. Psychological Bulletin. 1968; 70: 213-220.

(28) Sim J, Wright CC. The Kappa Statistic in Reliability Studies: Use, Interpretation, and Sample Size Requirements. Physical Therapy. 2005; 85(3): 257-268.

(29) World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA. 2013; 310(20): 2191-2194.

Appendix A: Scoring system of the DNT-II

Juiste benoeming ***Score 3***	**Toelichting**	**Voorbeeld uiting en doelwoord**
3.1 juiste benoeming	- Enkelvoudige correcte benoemingen van het item - Doelwoord in zinsframe (vb. 1)* - Doelwoord gevolgd door „of" met nevenschikking, waarna een keuze wordt gemaakt (vb. 2) - Toevoeging van bovenschikking (vb. 3) - Toevoeging van onderschikking (vb. 4) - Toevoeging van synoniem (vb. 5) - Volledige beschrijving bij het item (vb.6) - Verkleinwoord van doelwoord (vb. 7)	1. „het is een …", „het lijkt wel een …", 2. "Koffer of tas, koffer" bij koffer 3. „zadel voor op een paard" bij zadel 4. "kegel, bowlingkegel ofzo" bij kegel 5. "bier of pils" bij bier 6. "wijnfles en wijnglas" bij wijn 7. "biertje" bij bier
3.2 Juiste benoeming met fonematische- of dysartrische afwijking	- Tweederde van de klanken van de benoeming zijn juist gerealiseerd	1. "islo" bij iglo
3.3 (Onnodig) specifieke benoeming	- Correcte, (onnodig) specifieke benoeming - Het doelwoord in deel van (specifieke) samenstelling (vb. 1 t/m 3) Neologistische specifieke benoemingen krijgen een score 2.7 gezien ze het doelwoord niet nader specificeren	1. "bowlingkegel" bij kegel

Goed passende benoeming *Score 2*	**Toelichting**	**Voorbeeld uiting en doelwoord**
2.1 lang aarzelen	- Stilte langer dan 10 sec. met juist resultaat (vb. 1) - niets zeggende frase langer dan 10 sec. met juist resultaat (vb. 2 en 3)	1. „… (>10 sec.) “vlag” bij vlag 2. „eh…emm (> 10 sec.) “vlag” bij vlag 3. „hoe heet dat nou ook al weer (> 10 sec.) “vlag” bij vlag
2.2 Zelfverbetering	- Zelfcorrectie met juist resultaat Als er na meerdere foutieve reacties een juiste reactie komt, wordt deze als zelfverbetering gescoord.	1. “tas, nee een koffer” bij koffer
2.3 Semantisch goed passende foutieve benoeming	- Goed passende nevenschikking Voor minder passende nevenschikkingen zie 1.3	1. “revolver” bij pistool 2. “Kanaal” bij rivier
2.4 Semantisch goed passende omschrijving	- Goede omschrijving van doelwoord zonder dat doelwoord benoemd wordt	1. “Plank om op te strijken” bij strijkplank
2.5 Semantisch goed passend neologisme	- Specifieke benoeming met een neologistisch karakter (vb. 1) - Goed passend neologisme (vb. 2) - Volledig doelwoord met „ding“ (vb. 3)	1. “waslijnknijper” bij wasknijper 2. “telrek” bij telraam 3. “Theepotding” bij theepot
2.6 Juiste benoeming met semantisch goed passende specificering		1. “computerscherm” bij computer
2.7 Semantische bovenschikking en/of tweede deel samenstelling	- Goed informatieve bovenschikking (vb. 1) - Juiste benoeming tweede deel samenstelling (vb. 2) Minder informatieve bovenschikkingen krijgen een 1.7	1. “vogel” bij pauw 2. “pen” bij vulpen
2.8 Synoniem in een vreemde taal		1. “locker” bij kluis
2.9 Goede benoeming in een zin gebruikt	- Toevoeging van een/meerdere woord(en) die geen boven-, onderschikking of synoniem van doelwoord zijn	1. “zachte bal” bij bal

Enigszins passende benoeming of omschrijving ***Score 1***	**Toelichting**	**Voorbeeld uiting en doelwoord**
1.3 Semantisch enigszins passende foutieve benoeming	- Minder passende nevenschikking Zie voor goed passende nevenschikking 2.3	1. “gooiding” bij boemerang 2. “sproeier” bij douche
1.4 Semantisch enigszins passende omschrijving	- Enigszins passende omschrijving van doelwoord zonder dat doelwoord benoemd wordt	1. “zo’n ding dat je weggooit” bij boemerang
1.5 Semantisch enigszins passend neologisme	- Het neologisme bevat elementen uit het semantisch veld van het doelwoord (vb. 1) - Deel van doelwoord met ding (vb. 2) - Neologistische beschrijving die niet het doelwoord bevat (vb. 3) - Contaminaties die het goede doelwoord bevatten, maar die door de toevoeging misleidend worden (vb. 4)	1. “flessendraaier” bij kurkentrekker 2. “theeding” bij theepot 3. “theebewaarder” bij theepot 4. “iglogletsjer” bij iglo
1.7 te hoge bovenschikking of eerste deel van een samenstelling	- Te algemene bovenschikking (vb. 1&2) - Juiste benoeming eerste deel samenstelling	1. “dier” bij pauw 2. “instrument” bij gitaar

Niet passende benoeming ***Score 0***	**Toelichting**	**Voorbeeld uiting en doelwoord**
0.1 Geen uiting of ontwijkende frase	- Geen reactie - Niet in de situatie passende reactie	
0.2 Perseveratie	- Benoeming die maximaal zeven items ervoor actueel was.	
0.3 Automatisme		
0.4 Fonematisch neologisme	Fonematische parafasieën/woorden die geen semantische verwijzing naar het doelwoord bevatten - Zinloze woordsamenstellingen - Fonematische parafasieën die toevallig een woord zijn	
0.5 Semantisch zwaar afwijkende benoeming	Woorden die niet in het semantisch veld (staan of tot de periferie ervan horen - Mogelijke perseveraties - Woorden met vormovereenkomst met doelwoord	
0.6 Semantisch slecht passende omschrijving	- Omschrijvingen die vrijwel geen informatie bevatten aangaande het doelwoord	
0.7 Semantisch slecht passend neologisme	Vage beschrijvingen - Eerste deel van samenstelling voorafgegaan of gevolgd door een vorm die niet van toepassing is	
0.8 Visuele fout (incl. benoeming stippellijnen of benoeming van een deel van het item)	- Reactie is gebaseerd op een verkeerde waarneming of onjuiste interpretatie van het item	

Scoringsvoorbeelden

Doelwoord	Respons	Score
Auto		
Baby	kind	2.7
Bal	strandbal	3.3
Banaan		
Band	Autoband	3.3
	wiel	2.3
Bar	Restaurant	2.7
	Café	2.7
	barkrukken	0.8
Bier	Glas bier	3.3
	Bierglas	2.3
Bijl	Hakbijl	3.3
	beitel	1.3
Boek		
Boemerang	Frisbee	1.3
	kapstok	0.8
	gooiding	1.3
Boot	Schip	2.7
	speedboot	2.6
Brief	Papier	2.7
	geschrift	2.3
	document	2.7
	schrift	1.7
	perkament	1.7
	tekst	2.7
Brug		
Bureau	kantoor	2.7
Caravan		
Computer	Computerscherm	2.3
	beeldscherm	2.3
Croissant		
Deur	Voordeur	3.3
	portaal	2.3
Douche	Douchekop	3.3
	sproeier	1.3
	kraan	1.7
Eekhoorn		
Eiland	Zee	0.8
Envelop	brief	2.3
Fiets		
Flessenopener	Opener	2.7
	bieropener	2.3
Geweer	Pistool	1.3
	jachtgeweer	3.3
Gieter		

Gitaar		
Glijbaan		
Hagelslag	Hagel	1.7
Hand		
Hoefijzer	Hoef	1.7
Hond		
Horloge		
Huis	Huismodel	2.3
Iglo	Gletsjer	0.6
IJsje	Softijs ijsco	2.3 2.3
Jurk		
Kaars		
Kapstok		
Kegel	Bowlingbal Bowlingkegel pion	0.6 3.3 1.7
Kerk	kasteel	
Klomp		
Kluis		
Koets		
Koffer		
Konijn		
Koning	Cape Mantel prins	0.8 0.8 1.3
Krant		
Kuiken	Eend eendje	2.7 2.7
Kurkentrekker	Opener flessendraaier	1.3 1.5
Lepel		
Lift		
Liniaal	Meetlat meter	2.3 1.3
Medaille		
Paard		
Paddenstoel		
Paperclip	pincet	0.5
Parachute		
Pauw	Vogel fazant	2.7 1.3
Pistool	Geweer revolver	1.3 2.3
Planeet	Saturnus jupiter aarde, aardbol	1.3 1.3 1.3

	heelal	1.7
	hemellichaam	2.3
Puzzel	Legpuzzel	3.3
Raam(kozijn)	Kozijn	2.7
	Deur	0.5
	venster	1.7
Ridder		
Rivier	Zee	1.3
	Water	1.7
	Meer	1.3
	kanaal	2.3
Rolstoel		
Schoenen		
Slab	Schort	1.3
	servet	1.3
Slot	Hangslot	3.3
Springtouw	Touwtje spring	2.5
Step		
Strijkplank	Strijkijzer	0.5
Taart	“Kaart”	3.2
Tafel	Bijzettafel	2.6
Taxi	Auto	2.7
	taxi-auto	3.3
Telraam	Telrek	2.5
	rekje	0.6
	kralenrijgplank	0.4
	kinderspeelgoed	0.8
Theepot	Theeding	1.5
	theekan	2.3
	ketel	1.3
Ton	Emmer	1.3
Trap	Trappenhuis	2.7
Tube	Verf	0.8
	gel	0.8
	doucheschuim	0.8
Vis		
Vlag		
Vogelhuisje		
Vork		
Vulpen	Pen	2.7
	balpen	1.3
Vuur	Wol	0.6
	vlammen	1.3
Washand		
Wasknijper	Klem	1.3
	Knijper	2.7
	waslijnknijper	2.5

Wijn	Rode wijn glas wijn wijnfles plus glas fles glas wijn	3.3 3.3 3.1 3.2
Zadel	Rijzadel paardrijzetel paardzitje	3.3 1.5 1.3
Zandbak		
zwaard	Sabel dolk	1.3 1.3

Appendix B: Research protocol

Onderwerp
Dutch Naming Test II

Datum
Februari 2017

INSTITUUT VOOR PARAMEDISCHE STUDIES
HOGESCHOOL UTRECHT

Postbus 85182
3508 AD Utrecht
Fax 088 481 59 36
www.hu.nl

Protocol testafname "Ontwikkeling van een benoemtest voor het diagnosticeren van woordvindingsstoornissen bij personen met afasie".

Stap 1: Proefpersonen worden geïncludeerd door de logopedist op locatie of door de student-onderzoeker/de onderzoeksassistenten.

Stap 2: Proefpersonen krijgen een informatiebrief en tekenen informed consent.

Stap 3: Alles wordt klaargemaakt voor de testafname. Onderzoeksruimte wordt opgeruimd, tafel wordt leeg gemaakt. Scoreformulieren en opname apparatuur worden klaargelegd.

Stap 4: De opname wordt gestart, de proefpersoon wordt respectievelijk gevraagd om zijn naam, leeftijd en opleidingsniveau.

Stap 5: Testen worden afgenomen op locatie door de behandelend logopedist, onderzoeker of onderzoeksassistent in de volgende volgorde:

1. DNT-II
2. SAT verbaal
3. BBT
4. PALPA 14

Indien gewenst wordt een korte pauze ingelast tussen de SAT verbaal en de BBT.

De testen worden in principe in zijn geheel afgenomen. De test afnemer kan er echter in uitzonderlijke gevallen voor kiezen om de test afname te staken. Hiervoor kunnen twee redenen zijn;
(1) de belastbaarheid van de proefpersoon kan te laag zijn om de testafname voort te zetten. Indien de proefpersoon aangeeft te moe te zijn voor voortzetting, wordt de testafname gestaakt en op een later tijdstip hervat.
(2) de talige problemen van de proefpersoon zijn dermate ernstig, dat voortzetting van de testafname gepaard gaat met onacceptabele frustratie. Indien een proefpersoon continue geen antwoorden geeft, ofwel blijvend persevereert op een voorgaand item, dan kan de testafnemer ervoor kiezen om de afname van de betreffende test te staken. De resterende testen worden vervolgens wel alsnog afgenomen.

Indien één van de volgende situaties zich voordoet én de proefpersoon zichtbaar gefrustreerd raakt, wordt de afname gestaakt:
- DNT-II: de proefpersoon geeft op 10 of meer achtereenvolgende items geen antwoord of persevereert op een voorgaand item.
- SAT verbaal: de proefpersoon geeft op 10 of meer achtereenvolgende items geen antwoord.
- BBT: de proefpersoon geeft op 10 of meer achtereenvolgende items geen antwoord of persevereert op een voorgaand item.
- PALPA 14: de proefpersoon geeft op 10 of meer achtereenvolgende items geen antwoord.

Stap 6: De testen worden gescoord door de test afnemer, voor zover dit nog niet gebeurd is tijdens de testafname. De behaalde scores hoeven niet te worden geïnterpreteerd, tenzij de behandelend logopedist dat wenst voor behandeldoeleinden.

Stap 7: De testformulieren en de opnames worden ingeleverd bij de student onderzoeker.

Stap 8: De student-onderzoeker of een onderzoeksassistent scoort de DNT-II op basis van de opname, zodat de interbeoordelaarsbetrouwbaarheid kan worden bepaald.

Stap 9: De student-onderzoeker codeert de persoonsgegevens en testgegevens en verwerkt alle gegevens in SPSS.

Contactgegevens student-onderzoeker:

Carlijn de Hilster
c.m.dehilster@students.uu.nl
06-51153274

Contactgegevens supervisor:

Dr. Lizet van Ewijk
lizet.vanewijk@hu.nl

Appendix C: Informed consent

dr. Lizet van Ewijk	Heidelberglaan 7
dr. Ingrid Cnossen	3584 CS Utrecht
Lectoraat Logopedie	088-481 54 45
Participatie door Communicatie	secretariaat.lectoraat.logopedie@hu.nl
Faculteit Gezondheidszorg	

Een test voor het meten van woordvindproblemen

Deelnemer:

Onderzoeker: **Carlijn de Hilster**

Masterstudent Logopediewetenschap

Universiteit Utrecht

06 511 532 74

c.m.dehilster@students.uu.nl

☺ **Bedankt** voor uw interesse in dit onderzoek!

Dit document geeft u meer informatie over:

- **Waarom** we dit onderzoek doen
- **Wat we van u vragen**, als u besluit mee te doen
- **Wat** het onderzoek **inhoudt**

<u>Waarom</u> <u>we dit onderzoek doen</u>

Uw **hersenen** zijn **beschadigd**,

Hierdoor heeft u **afasie**,
en **problemen** met de **taal**.

U heeft misschien **moeite**
met het **vinden** van **woorden**:

woordvindproblemen.

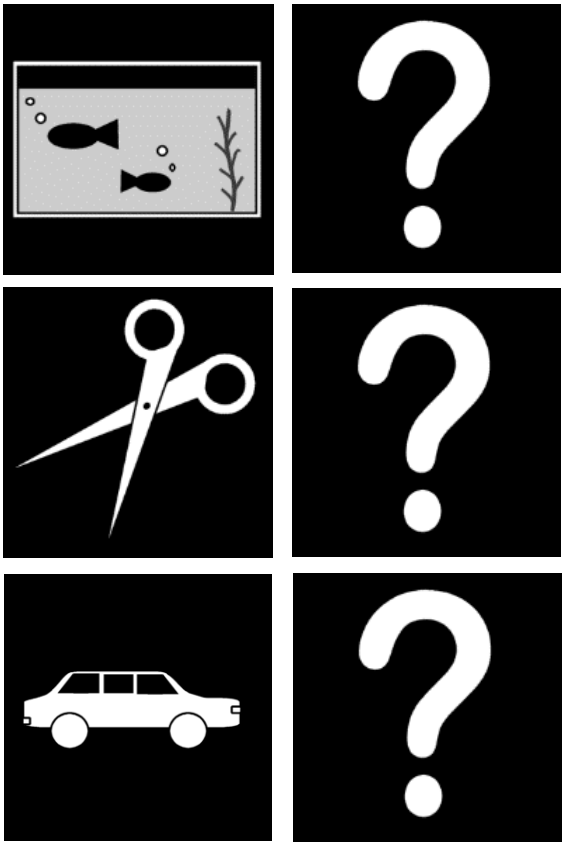
Hoe erg zijn uw **woordvindproblemen**?

We willen **onderzoek** doen:
Gebruiken we een **goede woordvindtest**?

Wat we van u vragen:

Het **beantwoorden** van **vragen**,

Het **benoemen** van **afbeeldingen**.



Onderzoekster **Carlijn de Hilster**:

Wat het **onderzoek inhoudt**

Het **onderzoek** duurt
ongeveer **15 tot 30 minuten**.

Als u **moe** wordt, **stopt** u even.
U kunt **later verdergaan**.

Waar?

De **onderzoeker bezoekt** u op de **afdeling logopedie.**

Voor- en nadelen

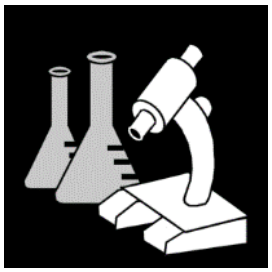
Voordelen

Er zijn **geen risico's** verbonden
aan dit **onderzoek**.

Door mee te doen aan het onderzoek:

- **Helpt u mee** met het **ontwikkelen** van een **nieuwe test.**

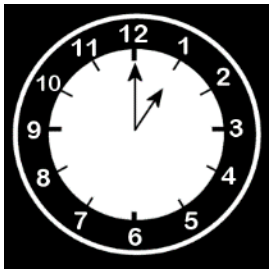
- **Helpt u mee** met het beter **begrijpen** van **woordvindproblemen** bij mensen met **afasie**.



- **Draagt u bij** aan **wetenschappelijk onderzoek**.

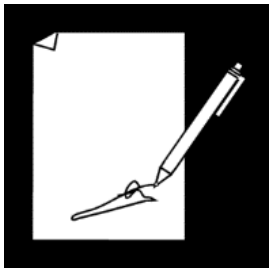
<u>Nadelen</u>

Het **afnemen** van het **onderzoek** neemt **tijd** in beslag.



U **steekt** hier **tijd** in.

<u>Toestemming</u>



Als u besluit **mee te doen**,
wordt u **gevraagd**
een **formulier** te **ondertekenen**.

Met uw **handtekening**:

- **Bevestigt** u dat u **voldoende informatie** heeft **ontvangen**

- Geeft u **toestemming** om het **onderzoek** bij u **uit te voeren**

U kunt **altijd stoppen** met het **onderzoek**.

Hier hoeft u **geen reden** voor te geven.

Stoppen heeft **geen invloed** op uw verdere **behandeling**.

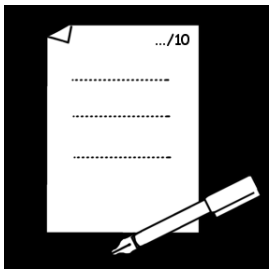
Wie is er nog meer **betrokken** bij het **onderzoek**?

Met uw **toestemming** vragen we
meer informatie over uw **afasie**
aan uw **behandelaar**.

Wat gebeurt er met mijn **gegevens**?

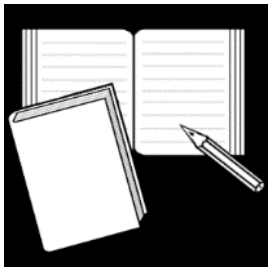
Uw **gegevens** worden
vertrouwelijk behandeld.
Uw **naam** zal **niet gebruikt** worden.

Als u wilt, kunnen we
uw behandelaar vertellen
hoe u het gedaan heeft.



Resultaten

De **resultaten** van het onderzoek worden **opgeschreven**.



Ze worden **bewaard** aan de **Hogeschool Utrecht**.

Toestemmingsformulier

(informed consent)

De **informatie** op de vorige bladzijden is aan mij **uitgelegd**.

☐ JA

☐ NEE

Ik heb een **kopie gekregen** van dit **document**.

☐ JA

☐ NEE

Ik ga **akkoord** met **deelname** aan dit **onderzoek**.

☐ JA

☐ NEE

Naam:_____ Datum:_____

Handtekening:_____