The development of an online neuropsychological test battery

The Amsterdam Cognition Scan

Feenstra, H.E.M.

Chapter 6

An English version of the Amsterdam Cognition Scan – Development and validation for North American populations

Heleen E. M. Feenstra
James C. Root
Tim A. Ahles
Jaap M. J. Murre
Ivar E. Vermeulen
Sanne B. Schagen

ABSTRACT

Aim

The current study assessed the usability and validity of the English version of the newly developed online neuropsychological test battery, the Amsterdam Cognition Scan (ACS), for North American populations in the oncology setting.

Method

The Dutch ACS (ACS-NL) was translated and used as a basis to create the ACS-EN. Concurrent validity was studied in 35 cancer patients (54% female; mean age 57.1 (SD=10.6) years) who completed both the ACS-EN and an equivalent traditional neuropsychological test battery. Correlations were used to assess consistency between online and traditional tests. Results were compared to those of a previously performed Dutch validation study. Usability was assessed based on online debriefing and technical reports.

Results

Concurrent validity was observed to be moderately-high to high (r/ρ=.51 to .81). Based on a matched (age and education) sample, validity results were found to be similar between the ACS-EN and the ACS-NL (r/ρ=.32 to .78). Usability of the ACS-EN was good: 34 out of 35 participants were able to complete the online assessment without supervision; instructions and practice were rated as sufficiently clear.

Conclusions

Results indicate that the ACS-EN has good usability and that its tests provide valid measures of their cognitive measurement constructs in the current sample of North American non-CNS cancer patients. To attain a fully employable research tool, norm data as well as additional studies on the influence of computer familiarity and reliability are required.

INTRODUCTION

The ACS was developed as an online tool for unsupervised cognitive testing to facilitate large-scale data collection for oncology research. The original version is in Dutch. To enable international testing and data-pooling, we translated the ACS into English, German, French, Spanish, and Swedish. As a first step towards an international ACS, a fully functional English version has been created for the North American population. From now on we will refer to this version as "ACS-EN". Because the ACS-EN has been translated and has a different target population than the original Dutch ACS—from now on referred to as "ACS-NL"—validation is required prior to application in research. The primary focus of the current study was to assess the usability and validity of the ACS-EN for use in North American populations. The study was carried out with a US sample, data was collected at Memorial Sloan Kettering Cancer Center (MSKCC).

METHOD

Development

To create an American English version of the ACS, first, all its elements (neuropsychological tests, instructions, feedback, animation videos, buttons and subtitles) were translated from Dutch to English. This was done by the research team with input from both Dutch and American English native speaking neuropsychologists. For the two questionnaires that are included in the ACS, official English versions were used (Smets et al., 1995; Zigmond & Snaith, 1983). Subsequently, audio-visual materials were prepared using an American English-speaking actor. Since the ACS tests that measure verbal learning and memory (Wordlist Learning, Delayed Recall, and Recognition) are language-based, these tests required additional attention during the translation process. To select appropriate English equivalents for the Dutch target and distractor words (see Appendix A), words were matched on: (1) word frequency, using the English SUBTLEX database (http://subtlexus.lexique.org/moteur2/index.php; (Brysbeart & New, 2009)) and the Dutch CELEX database (Baayen, Piepenbrock, & van Rijn, 1993); (2) word length; (3) semantic association; and (4) concreteness.

Study design

We studied concurrent validity by measuring consistency between scores on the ACS-EN and scores on well-established traditional neuropsychological tests (de Vet et al., 2011). As intended, the online ACS assessments were unsupervised, while the traditional assessments were conducted by a trained test leader. The study design was similar to that of the larger NL validation study (n=201; 56% female; mean age 53.5 (SD=12.3) years; 61% high education) (Feenstra et al., 2018b). Both assessments took place at the MSKCC Counseling Center within one day, using a counterbalanced design. Participants were randomized between two subgroups. Subgroup 1 first completed the traditional assessment and then the online ACS assessment (T-O), whereas subgroup 2 first completed the online ACS assessment and then the traditional assessment (O-T). The visit to the Counseling Center was scheduled to take 3 hours in total, including a 15 minute

break in between the two assessments.

Patient population

All participants were recruited from the MSKCC Survivorship Clinic (January to April 2015). Adult patients treated (with chemotherapy, immunotherapy, and/or hormonal therapy) for non-central nervous system cancer were screened by members of the protocol research team based on information from electronic patient files. Patients were excluded from the study in the case of evidence of active disease (metastasis or primary tumors), self-reported severe visual or auditory impairment, or diagnosis with a neurodegenerative disorder. Inclusion and exclusion criteria were similar to those from our NL validation study to enable comparison of results (see Appendix B for a full overview). After verification of eligibility, the research team mailed an invitation packet (consisting of an introductory letter describing the study, along with an informed consent and a self-addressed opt-out postcard) to prospective participants. Within approximately two weeks of mailing the invitation packet, potential participants were called by the research team to introduce the study (with a maximum of 5 follow-up calls, leaving no more than 3 messages). All eligible patients wanting to participate in the study were registered through the MSKCC Protocol Participant Registration (PPR) Office and signed an informed consent, meeting the requirements of the Code of Federal Regulations and the Institutional MSKCC Review Board/Privacy Board. Enrolled patients received an incentive of $20 for taking part.

Test administration

The assessments (online and traditional) took place in a quiet test room at the MSKCC Counseling Center. A laptop with a 17-inch screen using Microsoft Windows 7 and Google Chrome was used to run the ACS. Sound stimuli were played by an attached speaker set. A standard QWERTY keyboard and a two-button mouse were used to record response input. All online assessments were briefly introduced by the test leader and subsequently completed in an unmonitored setting. There was a telephone in the test room, which allowed participants to call the test leader if there were any problems or when they had finished the tests. All traditional assessments were conducted by the same test leader who introduced the online assessment using standardized instructions.

Materials

Online neuropsychological test battery - the Amsterdam Cognition Scan

Just as the ACS-NL, the ACS-EN consists of seven neuropsychological tests on the domains of attention, information processing speed, working memory, verbal learning and memory, visuo-spatial memory, executive functioning, and psycho-motor speed. All online tests were based on well-established traditional neuropsychological tests. Table 1 describes the test domains and main outcome measures of each ACS-EN test. Prior to the neuropsychological tests three tests on computer skills were presented, measuring typing skills and skills for clicking and dragging with the mouse. A detailed overview of all elements of the ACS can be found in Appendix A of Chapter 3.

Table 1. Tests of the ACS-EN and their equivalent traditional tests.

Online tests	Test domains	Main outcome measure	Traditional equivalent
Connect the Dots I; Connect the Dots II	Visuo-motor tracking, planning, cognitive flexibility, divided attention	Completion time (I & II)	Trail Making Test A Trail Making Test B
Wordlist Learning	Verbal learning	Total number of correct words (trial 1 to 5)	Rey Auditory Verbal Learning test, short form
Reaction Speed	Information processing speed & attention	Mean reaction time	Visual Reaction Time (subtest FePsy)
Place the Beads	Planning, response inhibition, visuo-spatial memory	Total number of extra moves	Tower of London, Drexel University (TOL-dx)
Box Tapping	Visuo-spatial short term memory	Total number of correctly repeated sequences	Corsi Block-tapping Test
Fill the Grid	Fine motor skills	Completion time	Grooved Pegboard
Wordlist Delayed Recall & Recognition	Retention of information: free recall and recognition	Total number of correct words; free recall and recognition	Rey Auditory Verbal Learning test, short form
Digit Sequences I; Digit Sequences II	I: Attention II: Working memory	Total number of correctly repeated sequences (I & II)	WAIS III Digit Span (forward & backward)

Traditional neuropsychological test battery

The traditional neuropsychological test battery consisted of conventional tests which require face-to-face assessment in English (see Table 1): (1) Trail Making Test (TMT) A & B (Reitan, 1958); (2) Rey Auditory Verbal Learning Test (RAVLT), short form (Mitrushina & Satz, 1991); (3) FePsy Visual Reaction Time (Alpherts & Aldenkamp, 1995); (4) Tower of London, Drexel University (TOL-dx) (Culbertson & Zillmer, 2001); (5) Corsi Block-tapping Test (Kessels et al., 2010); (6) Grooved Pegboard (Kløve, 1963); (7) WAIS-III Digit Span; forward & backward (Tulsky et al., 1997); and (8) Controlled Oral Word Association Test (COWAT) (Ruff, Light, Parker, & Levin, 1996). This test battery differed from the traditional battery of the NL validation study on the language dependent tests only: RAVLT and COWAT (note that results from the COWAT were not analyzed since the online version of this test is not completed to date). In addition, the Advanced Clinical Solutions Test of Premorbid Functioning (Chu, Y., Lai, Xu, & Zhou, 2012) was conducted to estimate level of IQ.

Questionnaires

Two questionnaires were presented both online—following the ACS—and via paper-and-pencil, following the traditional neuropsychological test battery: the Hospital Anxiety and Depression Scale (HADS) (Zigmond & Snaith, 1983), assessing symptoms of depression and anxiety, and the Multidimensional Fatigue Inventory (MFI-20) (Smets et al., 1995) assessing symptoms of fatigue. In addition, after completing the ACS, participants were online debriefed on test conditions during the online assessment (e.g., disruptions and technical issues) and on qualitative appreciation of several elements of the ACS.

Statistical analyses

Generally, statistical analyses were similar to those for the NL validation study (see Chapter 3,

Methods section for a description of statistical analyses in more detail). The NL validation study included a sample of 201 patients. Based on power calculations and feasibility, the target sample size for the US study was 35 to 40 patients. The final sample consisted of 35 patients.

Comparisons of demographics and clinical characteristics of the O-T/ T-O subgroups were performed using independent sample t-tests and chi-square tests. Outliers on neuropsychological test scores (traditional and online) were identified and excluded from data analyses similarly to the NL validation study, except that outliers were not identified per age group, but over the whole US sample because of the smaller sample size. This means that for reaction time outcomes (TMT/Connect the Dots, Visual Reaction Time/Reaction Speed, and Grooved Pegboard/Fill the Grid), we used the Median Absolute Distance (MAD) to detect outliers (Leys et al., 2013), while for tests that rely on number of correct responses and for which zero-scores are more likely to reflect usability issues than floor performance (RAVLT/Wordlist Learning, Corsi Block-tapping/ Box Tapping, and WAIS-III Digit Span/Digit Sequences), zero-scores were considered outliers. We applied an order-effect correction on all test scores from the second assessment, by subtracting the difference in mean test scores between subgroup O-T and T-O from all individual test scores. Concurrent validity of the ACS tests was assessed with Spearman and Pearson correlation coefficients (depending on the distributions of scores on the particular measurements). A criterion of ≥.40 was used to indicate acceptable validity (de Vet et al., 2011). Moreover, we took reliability of the traditional and online measures into account by calculating "ceiling consistencies": $\sqrt{}$*(reliability traditional test x reliability online ACS-NL test)*. Validity results were compared to results from the larger NL validation study (n=201) using Fischer-*z* tests.

All statistical analyses were performed using IBM SPSS Statistics for Windows, version 22.0 (Armonk, NY: IBM corp.). For correlational analyses, probabilities of $p<.01$ (two-tailed) were considered statistically significant to reduce the chance of type one error. For all other analyses, probabilities of $p<.05$ (two-tailed) were considered statistically significant in order to reduce the chance of type two error and to be more conservative.

RESULTS

Letters were sent to 147 MSKCC patients. Sixty-two patients (58%) were not interested or didn't have time to participate, 5 (5%) were excluded based on exclusion criteria, and 40 (37%) were unreachable. Of the 39 (26.5%) patients who agreed to participate, a total of 35 (90%) completed the study (54% female; mean age 57.1 (SD=10.6) years; 85.5% high education). Four patients dropped out before starting the study because of lack of time. Therefore, analyses are based on 35 US participants. Note that the Advanced Clinical Solutions Test of Premorbid Functioning (which is used to estimate IQ) could not be analyzed in four cases since two participants were non-native English speakers (even though their English language skills were sufficient to complete the assessments), one participant was dyslexic, and one participant stopped this test prematurely because it took too much effort. In addition, because of problems with the internet

connection, which required reloading the webpage, the results on Wordlist Learning, Delay, and Recognition were not valid for three participants and one participant was not able to complete the online questionnaires. Thirty-one percent of the participants experienced some type of technical problem, which was mainly caused by a weak internet connection at the hospital (10x internet connection disrupted; 1x automatic computer restart). In the NL study, 4 percent of the participants experienced technical problems (6x internet connection disrupted; 1x hardware problem).

Basic demographics and clinical characteristics for the US and the NL patient samples are provided in Table 2. Because of the observed differences in mean "estimated IQ" (US: 113.5; NL: 106.5) and "education level" (US: 85.5% high education; NL: 61% high education), we also selected a smaller sample from the NL validation study matched (1 to 1) on age and education of the US sample: n=35; 22 (63%) female; mean age 57.9 (SD=10.7) years; 85.5% high education. "High education" corresponds for the NL sample with Verhage 6 and 7 scores (Verhage, 1964), and in the US sample with academic degrees: Bachelor's, Master's, PhD/ Law school/ MD. Demographic

Table 2. Demographics and clinical characteristics of the US and the NL patient samples.

	US (n=35)	NL (n=201)	NL matched (n=35)
Gender (n)			
Female	54% (19)	56% (112)	63% (22)
Male	46% (16)	44% (89)	37% (13)
Age, M (SD) years	57.1 (10.6)	53.5 (12.3)	57.9 (10.7)
IQ estimate, M (SD)	113.5 (7.7) [a]	106.5 (9.6) [b]	111 (11.5) [b]
Education levelc, (n)			
Low	0	0	0
Medium	14.5% (5)	39% (78)	14.5% (5)
High	85.5% (30)	61% (123)	85.5% (30)
Tumor type, (n)			
Breast	40% (14)	41% (82)	48.5% (17)
Testis/ prostate	14.5% (5)	19% (39)	14.5% (5)
Other	45.5% (16)	40% (80)	37% (13)
Treatment type, (n)			
Surgery	94.5% (33)	68% (137)	69% (24)
Chemotherapy	51.5% (18)	77% (154)	74% (26)
Radiotherapy	40% (14)	75% (151)	69% (24)
Hormonal therapy	20% (7)	46% (92)	43% (15)
Immunotherapy	0	14% (28)	11.5% (4)

a = based on the Advanced Clinical Solutions Test of Premorbid Functioning. b = based on the Dutch reading tests for adults (Schmand, Bakker, Saan, & Louman, 1991). c = In the US sample, level of education is based on academic degree: Low = high school; Medium = associates; High = bachelors, masters, or PhD/ Law school/ MD. In the NL sample, level of education is based on Verhage education scores 1–7 (Verhage, 1964) (corresponding with the following US years of education; 1: 1–5 years; 2: 6 years; 3: 7–8 years; 4: 7–9 years; 5: 7–10 years; 6: 7–16 years; 7: 17–20 years): Low = Verhage 1 or 2; Medium = Verhage 3, 4 or 5; and High = Verhage 6 or 7.

differences that occurred between the US and the NL sample disappeared when comparing the US sample to the matched NL sample. There were no differences between US patient subgroups T-O (traditional tests first) and O-T (online tests first) on age, gender, IQ, and tumor type.

Online debriefing on usability, assessment type preference, and computer use

The main responses from the online debriefing are presented in Table 3. The instructions and practice sessions were generally rated as sufficiently clear. The majority of participants (71%) reported using the computer over 15 hours per week. Ninety-five percent of the participants reported preference for online (71% from home; 24% from the hospital/ medical center) over traditional neuropsychological assessment. Online debriefing from the NL validation study resulted in similar ratings and preferences, although the NL sample reported using computers less frequently (42% over 15 hours a week; see Table 3).

Table 3. Data from online debriefing after the ACS-EN and the ACS-NL (total and matched sample).

Question	ACS-EN n=31	ACS-NL n=199	ACS-NL matched n=34
	% (frequency)	% (frequency)	% (frequency)
Pace of instructions:			
Too slow	19 (6)	14 (28)	8.5 (3)
Too fast	0	0.5 (1)	0
Exactly right	81 (25)	85.5 (170)	88.5 (31)
Instructions *not* sufficiently clear	3 (1)	1.5 (3)	3 (1)
Tasks *not* clear after practice session	3 (1)	0.5 (1)	0
Number of hours computer use per week			
0-5	3 (1)	31 (61)	35.5 (12)
5-15	26 (8)	27 (54)	20.5 (7)
15-35	42 (13)	33.5 (67)	29.5 (10)
>35	29 (9)	8.5 (17)	14.5 (5)
Preference assessment type	n=21	n=177	n=28
Home online	71 (15)	68 (120)	60.7 (17)
Hospital paper and pencil	5 (1)	5 (9)	3.6 (1)
Hospital online	24 (5)	27 (48)	35.7 (10)

Concurrent validity

Table 4 shows concurrent validity results for all neuropsychological outcome measures of the ACS-EN (single outcome measures and total score). Questionnaire results are presented in Appendix C. Spearman/Pearson correlation coefficients were all significant (p<.01) and ranged from .51 to .81, indicating moderately-high to high validity. None of the tests correlated below .40 with their traditional equivalents. Consistencies between traditional and online performances ranged from 78 to 140 percent of the calculated ceiling consistencies (which indicates the maximum correlation

coefficient that can be expected, taking test reliabilities into account). Correlations found with the ACS-EN were generally stronger than those found with the ACS-NL (*r*/ρ: .36 to .70), although this difference was only significant for one test (Wordlist Delayed Recall: Fischer-*z*=2.29; *p*=.02).

Table 4. Concurrent validity for the US and NL patient samples - Spearman or Pearson correlation coefficients between the ACS tests and their traditional counterparts.

Tests	US n	Spearman's ρ/ Pearson's *r* (P) US	Spearman's ρ/ Pearson's *r* % of ceiling consistency[1]	NL n	Spearman's ρ/ Pearson's *r* (P) NL	Fisher-z	NL matched n	Spearman's ρ/ Pearson's *r* (P) NL matched	Fisher-z
Trail Making Test A/ Connect the Dots I	34	.51**	79 (.65)	195	.57***	-0.44	34	.59***	-0.45
Trail Making Test B/ Connect the Dots II	33	.70***	89 (.79)	196	.70***	0	35	.56***	0.92
RAVLT learning/ Wordlist Learning	33	.76*** (P)	112 (.68)	200	.64*** (P)	1.21	35	.78*** (P)	-0.19
RAVLT delayed/ Wordlist Delayed Recall	33	.81***	116 (.70)	200	.59***	2.29*	32	.68***	1.14
Visual Reaction Time/ Reaction Speed	33	.63***	140 (.45)	191	.49***	1.04	35	.58***	0.31
Tower of London/ Place the Beads	32	.59***	102 (.58)	189	.42***	1.15	33	-.07	-2.86**
Corsi Block-tapping/ Box Tapping	33	.54**	113 (.48)	194	.36***	1.16	33	.32*	1.06
Grooved Pegboard/ Fill the grid	34	.61*** (P)	109 (.56)	189	.45***	-0.13	32	.62***	-0.06
WAIS III Digit Span FW/ Digit Sequences I	34	.66***	89 (.74)	200	.43***	1.72	34	.43*	1.31
WAIS III Digit Span BW/ Digit Sequences II	34	.57***	78 (.73)	200	.52*** (P)	0.64	34	.55*** (P)	0.11
Total score	22	.58** (P)	N.A.	171	.78*** (P)	-1.62	28	.77*** (P)	-1.19

** Significant at p<.05; ** Significant at p<.01; *** Significant at p<.001. 1 = √(Pearson's r traditional test based on literature*

** Pearson's r/ Spearman's ρ online test based on NL validation study); N.A. = not applicable.*

Additional analyses on the matched NL sample yielded generally higher consistencies between traditional and online performances than those from the total NL sample; Spearman/Pearson correlations ranged from .32 to .78 (see Table 4). Based on this matched sample Fischer-*z* scores were generally lower and not significant, indicating similar validity results between the ACS-EN and the ACS-NL. The only difference was found for Place the Beads (Fischer-*z*=-2.86; *p*=.004).

DISCUSSION

This study assessed usability and validity of the English version of the newly developed online neuropsychological test battery – the Amsterdam Cognition Scan (ACS-EN), for North American populations.

Usability of the ACS-EN was good. Thirty-four out of 35 participants were able to complete the online assessment without help of a test leader and rated both the instructions and the practice sessions as sufficiently clear. The relatively large number of technical problems that occurred resulted from a weak internet connection at the test location. In a few cases, this caused extra presentation of memory stimuli, and test scores (Wordlist Learning, Delay, and Recognition) had to be excluded from the analyses. This stresses the importance of a solid internet connection during assessments with the ACS in order to collect reliable cognitive data.

Concurrent validity exceeded the .40 criterion for all ACS-EN measures (single tests and total score) and was relatively high when test-retest reliability of individual online and traditional tests was taken into account. This indicates that they provide valid measures of their cognitive measurement constructs in the current sample of US non-CNS cancer patients. When comparing the ACS-EN to the ACS-NL, performance between traditional and online assessments was found to be more consistent in the US sample. This is potentially due to the homogeneity of the US sample, which included almost exclusively highly educated patients. Additional analyses using a similarly highly educated matched NL sample provided validity results that were generally more comparable to those of the ACS-EN. As an exception, this was not the case for Place the Beads, possibly due to little variance in Place the Beads/ Tower of London performance in the matched sample. Other (smaller) remaining differences that tended to favor the ACS-EN might be related to the more frequent use of computers by the USA participants.

Previously, we have looked into the influences of tested computer skills as well as self-reported computer experience on ACS-NL performance and found associations with some of the ACS tests (see Chapter 3 (non-CNS cancer patients) and Chapter 4 (healthy adults)) as well as with several traditional face-to face tests (see Chapter 3). However, further research should clarify the influence of computer familiarity on ACS performance in more detail. In addition, future steps for implementing the ACS-EN include studies on reliability the collection of reference data.

The current study is limited by a small sample size (n=35). The inclusion rate was 26.5%,

whereas for a similar study previously performed at MSKCC a 62% inclusion rate was found. Possible reasons for this relatively low inclusion rate are severe weather conditions (prolonged periods of heavy snow and occasional winter storms) during the winter in which inclusion and the assessments took place. Furthermore, since the homogeneity of the US sample might limit generalizability to North American populations with lower education levels, further studies on the psychometric properties and reference data of the ACS-EN should include not only larger but also more diverse patient samples.

CONCLUSIONS

The ACS-EN provides results on cognitive functioning based on unsupervised testing that are comparable to those of traditional neuropsychological tests. To attain a viable research tool, additional norm data and additional studies on the influence of computer familiarity and reliability of the ACS-EN tests are required. A functional and validated ACS-EN will aid future international collaborations for cognitive research in the oncology field and beyond, and can foster larger-scale data collection in these areas.

APPENDIX A: ENGLISH EQUIVALENTS FOR THE DUTCH TARGET AND DISTRACTOR WORDS.

Target NL	Google English	Target USA	Length NL/ EN	Lexical frequency NL/EN	Distactor 1	Google English	Distractor USA 1	Length NL/ EN	Lexical frequency NL/EN	Distractor 2	Google English	Distractor USA 2	Length NL/ EN	Lexical frequency NL/EN
eland	moose	panther	5/ 7	47/ 46	ree	roe	leopard	3/ 7	96/ 84	hinde	doe	jaguar	5/ 6	46/ 63
kermis	Fair	parade	6/ 6	455/ 421	kraam	stall	fireworks	5/ 9	188/ 174	circus	circus	circus	6/ 6	292/ 450
polder	--none--	desert	6/ 6	203/730	vallei	valley	valley	6/ 6	400/ 600	heuvel	hill	hill	6/ 4	2237/ 920
viltstift	Felt-tip pen	marker	9/ 6	70/ 160	vulpen	Fountainpen	crayon	6/ 6	136/ 20	balpen	--none--	pencil	6/ 6	65/ 391
ballon	balloon	balloon	6/ 7	239/ 261	lampion	--none--	banner	7/ 6	63/ 135	slinger	pendulum	ribbon	7/ 6	222/ 156
afval	waste	garbage	5/ 7	541/ 852	puin	debris	rubbish	4/ 7	413/ 193	vuilnis	garbage	litter	7/ 5	55 / 146
loods	pilot	storage	5/ 7	236/ 326	kluis	safe	safe	5/ 4	191/ 511	kist	chest	box	4/ 3	1886/ 2353
rivier	river	river	6/ 5	2929/ 1211	sloot	Ditch	ditch	5/ 5	530/ 308	beek	brook	creek	4/ 6	629/ 241
gitaar	guitar	guitar	6/ 6	261/ 341	viool	Violin	violin	5/ 6	501/ 122	piano	Piano	piano	5/ 5	679/ 597
arts	doctor	doctor	4/ 6	3930/ 3703	dokter	Doctor	nurse	6/ 5	6140/ 1099	verpleger	nurse	physician	9/ 9	176/ 235
schuur	barn	attic	6/ 5	922/ 237	kelder	basement	cellar	6/ 6	1268/ 235	hok	hutch	loft	3/ 4	646/ 104
riem	belt	sweater	4/ 7	939/ 466	gordel	girdle	vest	6/ 4	352/ 203	ceintuur	belt	jacket	8/ 6	103/ 971
merg	pith	muscle	4/ 6	159/ 528	bot	bone	bone	3/ 4	68/ 855	been	leg	leg	4/ 3	7550/ 1588
zakdoek	handkerchief	plate	7/ 5	873/ 925	papier	paper	glass	6/ 5	4786/ 1841	handdoek	towel	dish	8/ 4	697/ 443
druppel	drop	rain	7/ 4	1112/ 1446	regen	rain	drop	5/ 4	2336/ 3548	plas	puddle	puddle	4/ 6	458/ 89

APPENDIX B: INCLUSION AND EXCLUSION CRITERIA.

Inclusion criteria:

- Non active cancer patients (currently no evidence of disease), all types without central nervous system involvement, who have undergone treatment (chemotherapy, immunotherapy, and/or hormonal therapy)
- Age 18-75 years
- In the judgment of the consenting professional, able to communicate well enough in English through verbal and written communication to complete the study assessments and provide informed consent

Exclusion criteria:

- As per medical record or self-report, diagnosis of neurodegenerative disorder that affects cognitive function (e.g., Alzheimer's, Parkinson's, Multiple Sclerosis, dementia, seizure disorders, etc.)
- As per medical record or as per self-report central nervous system involvement
- As per self-report or in the judgment of the consenting professional, visual or auditory impairment that would preclude ability to complete assessments

APPENDIX C: QUESTIONNAIRE RESULTS.

Questionnaire	US n	Spearman's ρ/ Pearson's *r* (P) US	NL n	Spearman's ρ/ Pearson's *r* (P) NL	Fisher-*z*	NL matched n	Spearman's ρ/ Pearson's *r* (P) NL matched	Fisher-*z*
HADS anxiety	34	.89* (P)	200	.94* (P)	-1.64	34	.93* (P)	-0.93
HADS deppression	34	.91* (P)	200	.94* (P)	-1.09	34	.94* (P)	-0.83
MFI	33	.87* (P)	200	.96* (P)	-3.13*	33	.97* (P)	-1.58

** Significant at p<.001.*