



CAT-NL

ERNSTBEPALING
KRITIEKE WAARDEN

WHITE PAPER 3

White paper

CAT-NL

Ernstbepaling Kritieke Waarden

White paper 3

M.G. van der Staij-Mulder

Logopedist, Klinisch linguïst, Treant Zorggroep, locatie Weidesteyn, Hoogeveen, lid
Werkgroep Logopedische richtlijn 'Diagnostiek en behandeling van Afasie'

E. G. Visch-Brink

Klinisch linguïst, afd. neurologie en neurochirurgie, Erasmus MC, Rotterdam

P. Mariën

Klinisch neurolinguïst, Klinische en experimentele neurolinguïstiek, Vrije Universiteit Brussel,
& Departement Neurologie en Geheugenkliniek, ZNA Middelheim, Antwerpen

D. Vandenborre

Logopedist, Linguïst, Klinische en experimentele neurolinguïstiek, Vrije Universiteit Brussel,
Thomas More Opleiding Logopedie en Audiologie, Antwerpen & CEPOS, Duffel

M. Kerkmeer

Psychometrist, Pearson Benelux B.V.

G.I. Leemans

Product Developer, Pearson Benelux B.V.

Augustus 2016

PEARSON

Inhoud

1	Inleiding	4
2	Methodisch-logopedisch handelen en de CAT-NL	5
3	Ernstbepaling	7
4	Kritieke waarden	10
5	Conclusie	12
	Literatuur	13

1 Inleiding

De CAT-NL is een gestandaardiseerd instrument, dat door de logopedist / (klinisch) linguïst gebruikt kan worden bij de diagnostiek van afasie. De CAT-NL voorziet in een kwantificering van verschillende taaltaken, zoals benoemen, naspreken, taalbegrip, lezen en schrijven. De test is gebaseerd op een breed gehanteerd taalverwerkingsmodel, dat onder andere in Patterson & Shewell (1987) en Ellis & Young (1988) beschreven staat.

Voor de normering van de CAT-NL zijn gegevens verzameld van 573 gezonde en 257 afatische respondenten uit Nederland en Vlaanderen. Per subtest, subdomein en domein zijn afkappwaarden en standaardscores berekend. De CAT-NL heeft een goede betrouwbaarheid. De validiteit van de test wordt ondersteund door onder meer een hoge correlatie tussen de verschillende subtests van de CAT-NL en respectievelijk de Token Test (versie AAT, Graetz, de Bleser & Willmes, 1992) en de ScreeLing (Visch –Brink, van de Sandt-Koenderman, & El Hachoui, 2010) (zie verder Visch-Brink, Vandenborre, de Smet, & Mariën, 2014; Vandenborre, Kerkmeer, Visch-Brink & Mariën, 2016).

In deze whitepaper worden aanvullende gegevens verstrekt voor:

1. het indelen van de ernst van de stoornis in drie categorieën (Ernstige, Matig-ernstige en Lichte afasie).
2. het beoordelen van significante verschillen tussen twee testafnames.

Hiermee wordt direct aangesloten bij de *Logopedische Richtlijn Diagnostiek en behandeling van afasie* (Berns et al., 2015) en het *NVAT Afasie Interventie Schema (NAIS)* (NVAT, 2015). Deze nieuwe analyses zijn gebaseerd op de bestaande klinische steekproef van de CAT-NL.

De tabellen worden opgenomen in een herziene uitgave van de CAT-NL Handleiding.

2 Methodisch-logopedisch handelen en de CAT-NL

De logopedist werkt op methodische wijze volgens gegeven stappen in een cyclisch proces. Dit 'methodisch handelen' kenmerkt zich door bewustheid, doelgerichtheid, systematiek en procesmatigheid (Meulenberg-Brouwer et al., 2008). Binnen alle stappen van het methodisch handelen gaat het om klinisch redeneren en evidence-based handelen (Houweling, Nijenhuis & Lipke-Steenbeek, 2008). Elke logopedische behandeling wordt voorafgegaan door gedegen diagnostiek (Commissie CVA, 2001; Kuipers, 2007; Cox, van Hemert & van der Vloed, 2012). Bij deze diagnostiek worden de stoornissen van de persoon met afasie klinisch beschreven en via meetinstrumenten in kaart gebracht (NVAT 2015).

Meetinstrumenten kunnen voor verschillende doeleinden gebruikt worden. Het doel waarvoor het meetinstrument is gebruikt, bepaalt hoe er met de testresultaten wordt omgegaan (Berns et al., 2015):

1. *Diagnostisch:*
 - stellen van een diagnose: in de context van een (vermoede) hersenaandoening wordt de aanwezigheid, de ernst en de aard van een afasie vastgesteld;
 - toetsen van een hypothese: achterhalen van de onderliggende stoornis van een vastgesteld probleem;
 - vaststellen van een baseline: voordat er gestart wordt met een specifieke therapie wordt de beginsituatie objectief gemeten.
2. *Prognostisch:*
 - bijdragen aan het formuleren van een prognose. De score op de fonologische schaal van de ScreeLing (Visch –Brink, van de Sandt-Koenderman, & El Hachoui, 2010) kan bijvoorbeeld worden gebruikt bij het bepalen van het te verwachten herstel van de afasie (El Hachoui et al., 2013; zie ook Berns et al (2015), Uitgangsvraag 2, p. 42).
3. *Evaluatief:*
 - meten van het effect van een behandeling: na een periode van gerichte therapie kan het effect hiervan worden vastgesteld.
 - meten van voor- en/of achteruitgang tijdens het verloop van de ziekte.

De CAT-NL is een veelomvattende, maar toch beknopte diagnostische testbatterij waarmee in relatief korte tijd de essentiële kenmerken van een afasie onderzocht kunnen worden. Gezien de uitgebreide normering bij gezonde en afatische proefpersonen voldoet de test aan de eisen die gesteld worden aan een gedegen diagnostisch instrument voor het meten van taalstoornissen ten gevolge van een neurologische aandoening.

De CAT-NL kan o.a. ingezet worden:

- voor de diagnose van de stoornis en de planning van de stoornisgerichte behandeling;
- voor het meten van de ernst van de taalstoornis (o.a. via een vergelijking met andere afasiepatiënten);
- voor het voorspellen en volgen in de tijd van veranderingen in de ernst van de afasie;

In deze white paper worden per (sub) domein ernstgradaties en kritieke verschillen beschreven. De inschaling in ernstgradaties geeft inzicht in het klinische profiel van individuele patiënten. De kritieke waarden maken het mogelijk om het verloop van de stoornis te volgen al of niet in relatie tot de therapie; bij hertesten kan worden bepaald of er sprake is van een significante vooruitgang.

3 Ernstbepaling

De ernst van de afasie is een bepalende factor voor de prognose (Lazar, Minzer, Antoniella et al., 2010 in NVAT, 2015). De ernstgraad beïnvloedt in sterke mate de inhoud van de interventie (Lazar et. Al, 2010 in NVAT, 2015).

In de NAIS (NVAT, 2015) en de Logopedische richtlijn ‘Diagnostiek en behandeling van afasie’ (Berns et al, 2015) zijn op basis van gestandaardiseerde meetinstrumenten indicaties van ernst opgesteld op stoornis- en activiteitsniveau. Aanbevolen wordt dat een onderzoeker bij elk meetmoment de ernst van de afasie bepaalt en dat hij of zij een basisset meetinstrumenten kiest ten behoeve van de diagnostiek en voor de evaluatie van de therapie. Het klinisch oordeel op basis van het totaalbeeld is hierbij doorslaggevend.

Voor de CAT-NL zijn ernstwaarden vastgesteld voor elk domein en elk subdomein van de Taaltest. De bepaling van de ernstwaarden is gebaseerd op de groep afatische respondenten (de patiëntengroep). De procedure per (sub)domein was als volgt:

1. *Afbakening van de groep*: het gaat om het deel van de patiëntengroep met een ruwe score op of beneden de afkapwaarde: het stoornisgebied. Een score boven de afkapwaarde geeft aan dat er geen sprake is van een stoornis. Voor elk onderdeel geldt immers dat respondenten punten krijgen voor taken die zij goed uitvoeren. Dus hoe hoger de score, des te minder waarschijnlijk is het uitvoeren van deze taak gestoord.
2. *Frequentieverdeling van de scores*: dit is een lijst met elke ruwe score binnen het stoornisgebied (dus van ruwe score 0 tot en met de afkapwaarde) en het aantal patiënten dat die score behaald heeft.
3. *Ernstbepaling*: de verdeling van de patiënten in drie categorieën: Ernstige, Matig-ernstige en Lichte afasie. Uitgangspunt was een evenredige verdeling van het aantal patiënten met een ruwe score in het stoornisgebied over de ernstcategorieën. De verdeling is niet overal precies gelijk geworden: het aantal patiënten met een score op of onder de afkapwaarde verschilde van domein tot domein. Zo scoorden bij Taalproductie 171 patiënten onder de afkapwaarde, en bij Taalbegrip 139. Ook moesten de aantallen patiënten met een bepaalde score volgens de frequentieverdeling gehandhaafd blijven.

De gehanteerde procedure is alleen zinvol als er binnen de patiëntengroep van de CAT-NL-steekproef een redelijke ernstverdeling is. Dit is inderdaad het geval: de frequentie-verdeling laat zien dat bij alle subdomeinen alle scores goed vertegenwoordigd zijn.

De resultaten staan in Tabel 1.

Tabel 1

Ernst van de stoornis (Ernstige, Matig-ernstige, Lichte afasie) per domein en subdomein van de CAT-NL Taaltest, gebaseerd op de ruwe scores¹

Domein	Bereik		N
Taalbegrip			
Ernstig	6 of lager		48
Matig-ernstig	7 of 8		41
Licht	9 of 10		50
Afkapwaarde**		10	139*
Maximale Ruwe Score**		20	

* aantal patiënten dat op of beneden de afkapwaarde scoorde

** zie noot 1

Subdomein	Bereik		N
Begrip gesproken taal			
Ernstig	40 of lager		45
Matig-ernstig	41-49		45
Licht	50-55		43
Afkapwaarde		55	133*
Maximale Ruwe Score		66	

* aantal patiënten dat op of beneden de afkapwaarde scoorde

Subdomein	Bereik		N
Begrip geschreven taal			
Ernstig	33 of lager		50
Matig-ernstig	34-45		48
Licht	46-52		49
Afkapwaarde		52	147*
Maximale Ruwe Score		62	

* aantal patiënten dat op of beneden de afkapwaarde scoorde

Domein	Bereik		N
Taalproductie			
Ernstig	13 of lager		54
Matig-ernstig	14-19		59
Licht	20-24		58
Afkapwaarde**		24	171*
Maximale Ruwe Score**		40	

* aantal patiënten dat op of beneden de afkapwaarde scoorde

** zie noot 1

¹ Let op: De Domeinscores (Taalbegrip en Taalproductie) zijn samengestelde scores: hun waarde in deze tabel is een optelling van de C-scores van de subdomeinen waaruit ze zijn samengesteld. Strikt genomen zijn het dus geen "echte" ruwe scores, want ze zijn niet gebaseerd op items. Het zijn echter ook geen standaardscores. Ze worden, net zoals "echte" ruwe scores, gebruikt voor de berekening van de Domein-standaardscores (ook C-scores). Daarom worden ze in deze tabel aangeduid als ruwe scores.

Tabel 1 (vervolg)

Subdomein	Bereik		N
Nazeggen			
Ernstig	20 of lager		46
Matig-ernstig	21-40		46
Licht	41-48		43
Afkapwaarde		48	135*
Maximale Ruwe Score		61	

* aantal patiënten dat op of beneden de afkapwaarde scoorde

Subdomein	Bereik		N
Benoemen			
Ernstig	18 of lager		60
Matig-ernstig	19-44		59
Licht	45-58		59
Afkapwaarde		58	178*
Maximale Ruwe Score		87+	

* aantal patiënten dat op of beneden de afkapwaarde scoorde

Subdomein	Bereik		N
Lezen			
Ernstig	21 of lager		56
Matig-ernstig	22-48		57
Licht	49-62		53
Afkapwaarde		62	166*
Maximale Ruwe Score		70	

* aantal patiënten dat op of beneden de afkapwaarde scoorde

Subdomein	Bereik		N
Schrijven			
Ernstig	23 of lager		50
Matig-ernstig	24-49		50
Licht	50-70		50
Afkapwaarde		70	150*
Maximale Ruwe Score		82	

* aantal patiënten dat op of beneden de afkapwaarde scoorde

De CAT-NL geeft geen maat voor de ernst van de afasie in het algemeen. De ernst wordt alleen gegeven per linguïstisch (sub)domein, juist om te differentiëren tussen de resultaten op deze testonderdelen. Dit is van belang voor een gerichte interventie. Bovendien geven de ernstgradaties van de stoornissen in de verschillende taalmodaliteiten directe aanknopingspunten voor eventueel nader onderzoek.

4 Kritieke waarden

Bij de interpretatie van scoreverschillen tussen twee testafnames moet een onderscheid gemaakt worden tussen de statistische significantie van het gevonden verschil en de klinische relevantie ervan. Als een scoreverschil statistisch significant is, dan wil dat zeggen dat dit verschil met een grote waarschijnlijkheid (bijvoorbeeld 95% of 90%) niet aan toevallige factoren te wijten is, maar duidt op een werkelijk verschil. Wanneer een verschilscore statistisch significant is, betekent dat echter nog niet dat dit verschil ook klinisch relevant is. Het eigen professionele oordeel en factoren als de achtergrond en de medische en/of lichamelijke conditie van de patiënt moeten worden meegewogen in de beoordeling of de waargenomen verschilcores een klinische betekenis hebben, dat wil zeggen of er sprake is van een daadwerkelijke verbetering of verslechtering.

Een betrouwbaar criterium om vast te kunnen stellen of het verschil in score tussen twee testafnames bij dezelfde patiënt als een daadwerkelijke verbetering (of verslechtering) geïnterpreteerd kan worden, is de 'kritieke waarde'. De kritieke waarde is afhankelijk van de betrouwbaarheid van de score-eigenschappen (interne consistentie) van de test.

Om te kunnen spreken van een klinisch relevant significant verschil moet de afwijking tussen de twee testafnames gelijk zijn of hoger dan de kritieke waarde.

Tabel 2

Kritieke waarden voor verschilcores tussen twee testafnames, berekend op ruwe scores en standaardcores, per domein en subdomein van de CAT-NL

2a. Kritieke waarden berekend op ruwe scores²

(Sub)domein	SD	r	kritieke waarde (95%-niveau)	kritieke waarde (90%-niveau)
Begrip van gesproken taal	12,53	0,91	10,7	8,9
Begrip van geschreven taal	14,26	0,94	9,7	8,1
Taalbegrip	3,73	0,96	2,1	1,7
Nazeggen	18,17	0,96	9,9	8,3
Benoemen	23,79	0,93	17,6	14,7
Lezen	24,27	0,98	10,3	8,6
Schrijven	26,81	0,91	22,3	18,6
Taalproductie	6,70	0,98	2,6	2,2

² Zie noot 1

2b. Kritieke waarden berekend op standaardscores (C-scores)

(Sub)domein	SD	r	kritieke waarde (95%-niveau)	kritieke waarde (90%-niveau)
Begrip van gesproken taal	2,00	0,91	1,7	1,4
Begrip van geschreven taal	2,00	0,94	1,4	1,1
Taalbegrip	2,00	0,96	1,1	0,9
Nazeggen	2,00	0,96	1,1	0,9
Benoemen	2,00	0,93	1,5	1,2
Lezen	2,00	0,98	0,8	0,7
Schrijven	2,00	0,91	1,7	1,4
Taalproductie	2,00	0,98	0,8	0,7

De gegevens in Tabel 2 zijn gebaseerd op de gehele klinische groep (N=257) van de CAT-NL. De gebruikte formule voor de berekening van de kritieke verschillen (Allen & Yen, 1979) is:

$$Z \times sd \times \sqrt{2 \times (1 - \text{betrouwbaarheid})}$$

Z is hierbij de waarde uit de normaalverdeling waarmee een bepaald significantieniveau samenhangt. We hebben hier gekozen voor twee niveaus: 95% en 90%. Dit komt overeen met een Z van respectievelijk 1,96 en 1,64. SD is de standaardafwijking. Bij ruwe scores verschilt die van testonderdeel tot testonderdeel, bij standaardscores is die overal gelijk. De in de CAT-NL gebruikte standaardscore is de C score, met een schaal die loopt van 0-10, een gemiddelde van 5 en een SD van 2, en daarom is in Tabel 2 bij standaardscores de kritieke waarde overal gelijk aan 2. De betrouwbaarheid die we hier gebruiken is Guttman's lambda-2.

Een voorbeeld van de beoordeling van verschillen tussen twee testafnames aan de hand van kritieke waarden:

Stel: bij de eerste testafname heeft een afatische patiënt op het subdomein 'Begrip van geschreven taal' een ruwe score van 43/62. Om te kunnen spreken van een daadwerkelijke verbetering moet hij/zij bij een hertest 10 punten meer halen dan de ruwe score op het eerste testmoment, dat wil zeggen 53/62 correct gescoord (95% betrouwbaarheid), of 8 punten meer, namelijk 51/62 correct gescoord (90% betrouwbaarheid).

De onderzoeker kan kiezen of de beoordeling op ruwe scores of standaardscores wordt gebaseerd. Bij de Domeinen Taalbegrip en Taalproductie is de ruwe score (Tabel 2a) de optelling van de C-scores van de subdomeinen waaruit ze zijn samengesteld. De standaardscores (C-scores, Tabel 2b) van de Domeinen zijn berekend op deze "tussenscores" (Zie ook noot 1).

Bij de hertestmomenten moet rekening worden gehouden met de plafondscores van eerder afgenomen tests. In het algemeen worden bij patiënten met een CVA alleen die (sub)testen opnieuw afgenomen waarop bij een eerder onderzoek op of lager dan de afkapwaarde werd gescoord. Op deze manier wordt onnodig hertesten voorkomen. Subjectieve klachten bij een persoon met een lichte of restafasie kunnen een reden zijn om een test waarop een plafondscores is behaald toch te herhalen, om te bepalen of er een werkelijk verschil is tussen de afnames. Uiteraard is de koppeling met een neurologische etiologie doorslaggevend.

5 Conclusie

De bepaling van de ernstgraden van een afasie voor de verschillende testonderdelen en de berekening van klinisch relevante verschillcores vergroot de waarde van de CAT-NL in de klinische praktijk. Deze gegevens verduidelijken de interpretatie van de testgegevens en kunnen meer helderheid brengen in de interactie tussen de onderzoeker en de patiënt en zijn/haar naasten.

En paar afsluitende opmerkingen bij het hanteren van deze gegevens:

Een verandering van ernstcategorie, bijvoorbeeld van Matig-ernstig naar Licht afatisch, zonder een kritiek verschil tussen de twee testafnames, is alleen van betekenis voor een subjectief oordeel. Een verschillscore van 1 of 2 punten kan namelijk al een verschuiving naar een andere ernstcategorie impliceren, terwijl het kritieke, dus betekenisvolle, verschil veel groter is. Het gemeten verschil van 1 of 2 punten kan dan louter op toeval berusten. Uiteraard is het wel zo, dat een verschuiving naar een 'normale waarde' klinisch van belang is voor de patiënt, zeker als dit bij meerdere testonderdelen het geval is. Het is aan de logopedist/ (klinisch) linguïst om dit goed te beoordelen.

Het is aan de onderzoeker om de keuze te maken tussen het significantieniveau van 90% of 95%. Het significantieniveau van 95% is wat strenger dan het 90% niveau.

De kritieke waarden zijn in decimalen aangegeven. Het is aan de onderzoeker om deze waarden af te ronden: 7,5 wordt 8; 7,4 wordt 7.

De ernstcategorieën en de kritieke waarden zijn alleen berekend op de totaalscores van de verschillende (sub)domeinen. Een vergelijking van het aantal en van de soort gemaakte fouten kan ook een patroon van verbetering aangeven, zoals bij de productietaken een verschuiving van het aantal neologismen naar fonematische parafasieën. Dit is onder meer van belang bij de subtests met een zeer beperkt aantal items, bijvoorbeeld het nazeggen en hardop lezen van complexe en nonwoorden.

De ernstcategorieën gelden niet voor de mondelinge en schriftelijke beschrijving van de situatieplaat. Hiervoor zijn aparte categorieën bepaald, verwerkt in de scoring.

Met dank aan Ineke van der Meulen en Mieke van de Sandt-Koenderman, Rijndam Revalidatie Centrum, voor een zinvolle discussie.

Literatuur

- Visch-Brink, E.G., van de Sandt-Koenderman, M. & El Hachioui, H. (2010) *ScreeLing*. Houten: Bohn Stafleu van Loghum.
- Blomert, L., Koster, C.H. & Kean, M.L.L. (1995). *Amsterdam-Nijmegen test voor alledaagse taalvaardigheid*. Lisse: Swets & Zeitlinger.
- Allen, M.J. & Yen, W.M. (1979). *Introduction to measurement theory*. Long Grove, IL: Waveland Press.
- Berns, P.E.G., Jünger, N., Boxum, E., Nouwens, F., van der Staaij, M. G., van Wessel, S., van Dun, W., van Lonkhuijzen, J.G. & CBO. (2015). *Logopedische richtlijn 'Diagnostiek en behandeling van afasie bij volwassenen'*. Woerden: Nederlandse Vereniging voor Logopedie en Foniatrie.
- Cox, E., van Hemert, A. & van de Vloed, J. (2011). *Standaarden Revalidatie*. Woerden: NVLF).
- Cox, E., van Hemert, A. & Koolhaas, C. (2010). *Standaarden Verpleeghuiszorg*. Woerden: NVLF. www.NVLF.nl
- Graetz, P., de Blezer, R & Willmes, K (1992). *De Akense Afasie test (AAT)*. Amsterdam: Hogrefe.
- Meulenberg-Brouwer, A., van der Pol-Top, H. & Kuiper, C. (2008). *Dialog. Een theoretisch en praktisch perspectief op de beroepsrollen van de logopedist*. Utrecht: Lemma Uitgeverij.
- NVAT (2015). *Afasie Interventie Schema van de Nederlandse Vereniging van Afasietherapeuten*. (NAIS). www.afasietherapie.info/NAIS
- Visch-Brink, E., Vandenborre, D., de Smet, H.J. & Mariën, P., (2014). *Comprehensive Aphasia Test - Nederlandse bewerking - Handleiding*. Pearson, Amsterdam
- Ellis, A. W. & Young, A. W. (1988). *Human Cognitive Psychology*. London: Erlbaum.
- Patterson, K.E. & Shewell, C. (1987). Speak and spell: Dissociations and word-class effects. In M.Coltheart, R. Job & G. Sartori (eds.), *The cognitive neuropsychology of language* (p. 273-296). Hillsdale, NJ: Lawrence Erlbaum Associates Inc.
- Vandenborre, D., Kerkmeer, M., Visch-Brink, E.G. & Mariën, P. (2016). Betrouwbaarheid en validiteit van de CAT-NL. *Logopedie*, 1, 19-35.



Pearson Benelux B.V.

Gatwickstraat 1, 1043 GK Amsterdam

Postbus 78, 1000 AB Amsterdam

T: +31 (0)20 581 5500

E: info-nl@pearson.com

www.pearsonclinical.nl

www.pearsonclinical.be

Twitter: [@PearsonNL](https://twitter.com/PearsonNL)