

Klinische bruikbaarheid van categoriegebonden woordproductie voor het onderscheiden van dementie en normale cognitieve veroudering

H.F.A. Diesfeldt, ^a W. van der Elst, ^b J. Jolles^b

Category fluency (animals, professions) in normal cognitive ageing and dementia

Regression-based normative data for commonly used semantic verbal fluency test (VFT) scores were derived from data of the Maastricht Aging Study ($N=1,825$; age range 24-81). Animal naming, profession naming, and the summed score (animal + profession naming) were shown to be profoundly affected by demographical variables, especially age and educational level. The clinical utility of the newly established regression-based VFT norms was evaluated in a large sample of psychogeriatric patients ($N=1,063$, of which $n=890$ met the criteria for Alzheimer's or vascular dementia). Results showed that the animal naming VFT score was the most sensitive measure to distinguish normal aging from dementia (when the specificities of the various VFT measures were equated). Fifty patients were found to commit stuck-in-set perseverations, i.e. they generated animal names during test administration of the profession naming VFT. The prevalence of stuck-in-set perseveration was 11.3% in Parkinson dementia and frontal variant frontotemporal dementia, versus 4.6% and 5.3% in dementia of the Alzheimer type and vascular dementia, respectively. Bivariate analyses revealed that people who committed stuck-in-set perseverations were more deficient on a measure of free recall and on several tests of executive control, such as the Expanded Mental Control Test, the Behavioral Dyscontrol Scale and the Alternating Graphical Sequences Test as a measure of response inhibition. Using multivariate logistic regression analysis, only the Alternating Graphical Sequences Test retained an independent effect in the association with stuck-in-set perseveration.

Key words: Semantic fluency, intrusion errors, response inhibition, executive function, cognition

Tijdschr Gerontol Geriatr 2009; 40: 00-00

Samenvatting

Met behulp van multiple regressieanalyse op gegevens van de Maastricht Aging Study (MAAS; N=1.825; leeftijdsbereik 24-81 jaar) zijn normen berekend voor een veelgebruikte semantische fluencytest (dieren en beroepen opnoemen). Het aantal genoemde dieren, beroepen en de somscore van dieren en beroepen samen bleek sterk afhankelijk van demografische variabelen, zoals leeftijd en opleiding. Het regressiemodel is toegepast op de gegevens van 1.063 psychogeriatrische patiënten, onder wie 890 met alzheimerdementie of vasculaire dementie. Bij gelijke specificiteit was de sensitiviteit voor het onderscheid tussen dementie en normale cognitieve veroudering het hoogst voor het aantal genoemde dieren. Vijftig patiënten persevereerden over taken (zij noemden tijdens afname van de beroepenfluencytaak spontaan weer een of meer 'dieren'). Dit verschijnsel kwam met 11,3% vaker voor bij parkinsondementie en de frontale variant van frontotemporale dementie dan bij alzheimerdementie (4,6%) en vasculaire dementie (5,3%). Persevereren over taken correleerde negatief met prestaties op een woordgeheugentaak, de EMCT (Expanded Mental Control Test), de BDS (Behavioral Dyscontrol Scale) en de meandertekening als een test voor verminderde responsinhibitie. In een multivariate logistische regressieanalyse correleerde alleen de meandertekening significant met persevereren over taken.

Trefwoorden: Maastricht Aging Study (MAAS); normen; specificiteit; sensitiviteit; voorspellende waarde

Inleiding

Hoeveel verschillende dieren kunt u opnoemen in één minuut? Voor onderzoekers van cognitieve functies is de prestatie op deze taak een indicatie van de toegankelijkheid van het woordstelsel. Bij dementie functioneert dat minder goed, omdat kennis verloren gaat, gedesorganiseerd raakt of minder gemakkelijk toegankelijk is.¹ Voorbeelden opnoemen uit een of meer ver-

trouwe categorieën is een verbale taak die een beroep doet op cognitieve controle, zoals actief organiseren en reguleren van het zoekproces om zoveel mogelijk verschillende voorbeelden te vinden binnen eenzelfde woordcategorie zonder in herhaling te vallen.² Vanwege de tijdlimiet is de taak ook gevoelig voor tempovertraging.³ Categoriegebonden woordproductie covarieert met indicatoren van alledaags functioneren.^{4,5} Dit alles verklaart waarom deze taak zich in het onderzoek van de uitvoerende controlefuncties een vaste plaats heeft verworven.^{6,7}

Om de resultaten van patiënten te interpreteren zijn vergelijkingsgegevens nodig uit de normale bevolking. Voor het Nederlandse taalgebied zijn deze op uitgebreide schaal verzameld door onderzoekers van de Maastricht Aging Study (MAAS).⁸ MAAS-deelnemers (N=1.825) werden geworven via het RegistratieNet Huisartspraktijken in de provincie Limburg. Na zorgvuldige screening op een eventuele cognitieve stoornis, werkten de deelnemers mee aan een serie psychologische tests, waaronder het opnoemen van zoveel mogelijk voorbeelden uit de categorieën dieren, en vervolgens beroepen, telkens in één minuut. In de klinische praktijk worden deze categorieën voor het onderzoek van de woordproductie veel gebruikt.⁹⁻¹¹

De MAAS-onderzoekers ontwikkelden via multiple regressie een normeringsmodel dat rekening houdt met de invloed van de demografische variabelen leeftijd, sekse en opleidingsniveau. Door het model toe te passen op gegevens van een patiënt kan worden bepaald in welke mate diens score afwijkt van het resultaat dat a priori kan worden verwacht van een cognitief intact persoon met dezelfde leeftijd, geslacht en vooropleiding als de patiënt. Deze methode van normeren, via een regressievergelijking, heeft als voordeel dat er een grotere nauwkeurigheid volgens een glijdende schaal mee wordt bereikt.⁸ Er zijn geen grote aantallen tabellen nodig, zoals bij de klassieke normen die zijn bepaald per relevante demografische categorie (i.c. voor vrouwen en mannen afzonderlijk, naar leeftijdscategorie en opleidingsniveau). Het is in de klassieke opzet niet altijd mogelijk om voor elke subgroep voldoende deelnemers te werven, waardoor gemiddelden en standaarddeviaties, en dus ook de ontwikkelde normgegevens, minder betrouwbaar zijn. Een ander nadeel van de klassieke methode is dat normen 'discontinu' veranderen. Bijvoorbeeld, wanneer twee onderzochte personen slechts een jaar in leeftijd van elkaar verschillen, kunnen beiden vanwege arbitraire grenzen in verschillende normgroepen vallen.¹²

De regressiemodellen uit het MAAS-onderzoek zijn beschikbaar voor toepassing in de klinische praktijk. In dit onderzoek worden gegevens gepresenteerd over de klinische bruikbaarheid van de regressiemodellen, zoals de sensitiviteit, de specificiteit en de voorspellende waarde voor het onderscheid tussen dementie en normale cognitieve veroudering.

^a PgD, De Stichtse Hof, Vivium zorggroep, Laren NH

^b Maastricht Brain and Behavior Institute, European Graduate School of Neuroscience (EURON), Department of Psychiatry and Neuropsychology, Maastricht University
Correspondentie: Dr. H.F.A. Diesfeldt, PgD, De Stichtse Hof, Naarderstraat 81, 1251 BG Laren. E:h.diesfeldt@vivium.nl

Een tweede uitbreiding op de eerdere publicatie van Van der Elst e.a. is dat hier ook het regressiemodel voor de somscore over dieren en beroepen wordt gepresenteerd. Dit sluit aan bij de gangbare klinische praktijk om uit te gaan van de somscore van dieren en beroepen samen.^{10,11} Een ander argument om de somscore te onderzoeken is dat deze meer items telt dan elke categorie afzonderlijk, en daarom een betrouwbaarder indicator van de woordproductie kan zijn dan de score voor één categorie.¹³

Hoewel beperking tot één categorie (meestal dieren) in de klinische praktijk niet ongebruikelijk is, biedt onderzoek met twee categorieën extra mogelijkheden om uitspraken te doen over uitvoerende mentale controlefuncties. Sommige patiënten geven, nadat zij voorbeelden hebben genoemd uit de tweede categorie (beroepen), onbedoeld weer voorbeelden van de eerste (dieren). Dit 'afdwalen' van de actuele taak is in de literatuur onder verschillende namen bekend, zoals intrusie, persevereren over taken, of stuck in set-perseveration. Het wijst op onvermogen om de meest recente taak of opdracht te onthouden, flexibel van taak te wisselen, of beide.¹⁴ Persevereren over taken komt vooral voor bij patiënten met een stoornis van de prefrontale hersensystemen.¹⁵ Volwassenen met een intacte cognitie vallen in wordfluencytaken niet terug op een voorgaande categorie.¹⁶ Hier wordt de prevalentie van dit probleem in een psychogeriatrische patiëntenpopulatie onderzocht. Tevens wordt nagegaan in hoeverre het persevereren over wordfluencytaken samenhangt met onafhankelijke indicatoren van episodisch geheugen en uitvoerende mentale controle.

Methode

Deelnemers

Maastricht Aging Study (MAAS). De werving van de deelnemers uit het MAAS-bestand is elders uitvoerig beschreven.⁸ Een groep van 10.396 potentiële deelnemers tussen 24 en 81 jaar oud werd verkregen door een aselechte trekking uit het *RegistratieNet Huisartspraktijken* in de provincie Limburg. Personen die volgens medische gegevens in het register leden aan een aandoening met een bekende nadelige invloed op het cognitief functioneren, zoals een cerebrovasculaire ziekte, hersentumor, multiple sclerose, epilepsie, parkinsonisme, dementie, verstandelijke beperking of psychose, werden bij voorbaat van deelname uitgesloten. Van de overblijvende potentiële deelnemers reageerden 4.490 door het insturen van een formulier positief op een verzoek van hun huisarts om aan het onderzoek mee te doen (3.531 personen lieten expliciet weten niet te willen meedoen; 2.375 personen stuurden geen formulier in). De 4.490 potentiële deelnemers werden geïnterviewd in verband met uitsluitingscriteria die niet in het huisartsregistratiesysteem aanwe-

zig waren, zoals een doorgemaakte TIA, hersenoperatie, nierdialyse, elektroconvulsieve therapie of langdurig gebruik van psychotrope stoffen. Dit leidde tot uitsluiting van 301 personen. Van de resterende 4.189 potentiële deelnemers werden er vervolgens 1.856 aselekt gekozen uit twaalf leeftijdscategorieën (25 +/- 1 jaar; 30 +/- 1 jaar; enz. tot 80 +/- 1 jaar). Bij een klein aantal deelnemers (31) rees alsnog twijfel aan intacte cognitie, bijvoorbeeld omdat hun MMSE-score lager was dan 24, of omdat zij bij een van de verbale *fluency*tests meer dan vijf fouten maakten, of vaker dan vijf keer eerder genoemde voorbeelden repeteerden. Gegevens van deze deelnemers werden niet in de regressiemodellen verwerkt, waardoor er uiteindelijk van 1825 deelnemers *fluency*gegevens zijn gebruikt. Elke leeftijdsgroep telde ongeveer 160 deelnemers (en per groep ongeveer evenveel vrouwen als mannen). Voor de oudste groep (79-81 jaar) werden minder deelnemers gevonden (n=67). Het opleidingsniveau van de deelnemers werd bepaald met behulp van de handleiding van De Bie.¹⁷ De oorspronkelijke indeling van acht opleidingsniveaus werd gehercodeerd naar drie opleidingsniveaus: lager algemeen en lager beroepsonderwijs ('laag'), middelbaar algemeen en middelbaar beroepsonderwijs ('midden'), en voortgezet algemeen of hoger onderwijs ('hoog'). Zie bijlage 1 voor een inventarisatie van de diverse scholingsvormen en hun indeling naar opleidingsniveau. In afwijking van de oorspronkelijke MAAS-indeling is 'voortgezet algemeen onderwijs' voor de berekening van de regressieformules hier ingedeeld bij 'hoog' (in plaats van bij 'midden'). Daarmee werd de indeling zo goed mogelijk in overeenstemming gebracht met het classificatiesysteem van Verhage, dat in de klinische praktijk gangbaar is.¹⁸ Testafname gebeurde tussen 1993 en 1996.

Psychogeriatrische dagbehandeling. De patiëntengroep in dit onderzoek werd gevormd door 1.116 achtereenvolgende deelnemers aan psychogeriatrische dagbehandeling (vanaf 21 juni 1994 tot 9 mei 2008). Zij werden voor dagbehandeling verwezen door geriaters of neurologen van de regionale geheugenpoli's. In het kader van de dagbehandeling werd elke patiënt uitgenodigd voor een psychologisch onderzoek, volgens een vast protocol. Het onderzoek was gespreid over twee sessies met een interval van meestal een week. Onderdeel van de tweede sessie was de Amsterdamse Dementie-Screeningstest (ADS), met daarin de subtest *Fluency*. Bij deze test probeert de onderzochte zoveel mogelijk dieren te noemen in één minuut, en daarna zoveel mogelijk beroepen, opnieuw in één minuut. Sommige deelnemers (n=53) waren voor de geplande tweede zitting, waarin de *fluency*test zou worden afgenomen, niet meer in dagbehandeling. Daardoor komt het totale aantal patiënten voor wie gegevens over het aantal genoemde dieren en beroepen beschikbaar zijn op 1.063. Enkele demografische gegevens van deze groep zijn weer-

gegeven in tabel 1. Mentale stoornissen werden gecodeerd volgens de systematiek van de DSM-IV, op basis van diagnostische gegevens van de verwijzende specialist.¹⁹ Bij 90,4% van de deelnemers werd een vorm van dementie vastgesteld, met een gemiddelde duur van 3,6 jaar ($SD=2,7$). Nosologische oorzaken waren ziekte van Alzheimer (71,3%), vasculaire dementie (12,4%) en overige dementieën (6,7%), bijvoorbeeld als gevolg van de ziekte van Parkinson of frontotemporale atrofie. Bij 9,6% van de patiënten was geen dementie gevonden, maar een affectieve of enkelvoudige cognitieve stoornis.

Tabel 1 Demografische kenmerken van deelnemers aan de Maas-tricht Aging Study (MAAS) en patiënten (deelnemers aan psychogeriatrische dagbehandeling).		
	MAAS	patiënten
N	1.825	1.063
periode	1993-1996	1994-2008
vrouwen (%)	49,86	59,36
mannen (%)	50,14	40,64
leeftijd		
M	51,62	78,66
SD	16,42	6,45
range	24-81	45-94
opleiding ^a		
laag (%)	36,71	57,01
midden (%)	34,58	20,04
hoog (%)	28,71	22,95

^a Voor indeling: zie Bijlage.

Procedure

Voor de scoring van dieren en beroepen werd de instructie in de handleiding van de ADS gevolgd. Deze komt overeen met de door de MAAS-onderzoekers gebruikte instructie voor Woordopnamen I (dieren) en II (beroepen) uit de Groninger Intelligentie Test (GIT).⁹ De *fluency* score is het totaal over de twee categorieën samen.²⁰ Voor elk van de drie afhankelijke variabelen (dieren, beroepen en de somscore) werd een multivariaat regressiemodel berekend. De onafhankelijke variabelen (predictoren) in deze stapsgewijze, hiërarchische multiple regressieanalyses

waren leeftijd, het kwadraat van de leeftijd, geslacht, opleiding en de tweeweginteracties tussen deze variabelen. Omdat voor opleiding drie niveaus werden onderscheiden, zijn twee dummyvariabelen geconstrueerd (dit zijn dichotome variabelen met uitkomst 0 of 1): lage, resp. hoge opleiding, met de middenniveau opleiding als referentiecategorie. Ter vermindering van multicollineariteit (of erg hoge correlaties tussen twee predictoren) werd de predictor leeftijd gecentreerd (gecentreerde leeftijd = kalenderleeftijd - 50). Vervolgens werd bij de introductie van gekwadraterde variabelen en interacties, leeftijd als gecentreerde variabele in de analyses opgenomen. Niet-significante predictoren ($p > 0,005$) bleven in het multivariate model buiten beschouwing. In de hier gebruikte analyses is een groot aantal onafhankelijke variabelen getoetst. Daarom is de gekozen p-waarde voor significantie een factor 10 kleiner dan de gebruikelijke 5%.

Persevereren over taken werd vastgesteld zodra de onderzochte tijdens afname van de beroepen-*fluency* taak overging op 'dieren'. In univariate analyses werd eerst nagegaan welke indicatoren van episodisch geheugen en uitvoerende mentale controle onderscheid lieten zien tussen patiënten die perseveratiefouten maakten en wie dat niet deden. Vervolgens werden significant correlerende indicatoren in een multivariate logistische regressieanalyse op hun onafhankelijkheid als predictoren van perseveratiefouten getoetst. Als indicatoren van episodisch geheugen gebruikten we het aantal gereproduceerde voorwerpsnamen uit de Vijf Afbeeldingen Test (somscore over drie trials), de uitgestelde herkenning van vijf eenmalig getoonde afbeeldingen (incidenteel geheugen) en de subtest Visueel Geheugen uit de ADS.²¹ Indicatoren van uitvoerende mentale controle waren de resultaten op de *Expanded Mental Control Test* (EMCT), de *Behavioral Dyscontrol Scale* (BDS), de meandertekening uit de ADS, en de voor leeftijd en opleiding gecorrigeerde woord*fluency* score.^{20,22-24}

Resultaten

Regressiemodellen

Het regressiemodel voor het door MAAS-deelnemers genoemde aantal dieren bevat als onafhankelijke en statistisch significante voorspellers leeftijd en opleidingsniveau. De regressiegewichten (B) in tabel 2 laten zien dat de voorspelde dieren*fluency* score gelijk is aan 24,6 voor een 50-jarige met een 'midden' opleidingsniveau. De voorspelde dieren*fluency* score daalt lineair als functie van leeftijd met ongeveer 0,1 diernamen per jaar leeftijdsverschil. Onafhankelijk van de leeftijd heeft schoolopleiding een duidelijke invloed op de testprestatie. Deelnemers met een relatief lage opleiding noemden gemiddeld twee tot drie dieren minder

(-2,653) dan deelnemers met een middenniveau opleiding, en ruim vier dieren minder dan deelnemers met een ‘hogere’ opleiding ($2,653+1,572=4,225$). Daarmee is de invloed van opleiding equivalent aan 27,6, respectievelijk 44 jaren leeftijdsverschil ($2,653/0,096$, respectievelijk $4,225/0,096$). De negatieve invloed van een

relatief lage opleiding op het aantal genoemde dieren is verhoudingsgewijs groter dan de positieve invloed van een relatief hoge opleiding. Zie de gestandaardiseerde regressiegewichten in de kolom ‘bèta’. Leeftijd en opleiding verklaren 17,7% van de variantie in het aantal genoemde dieren (zie kolom R²).

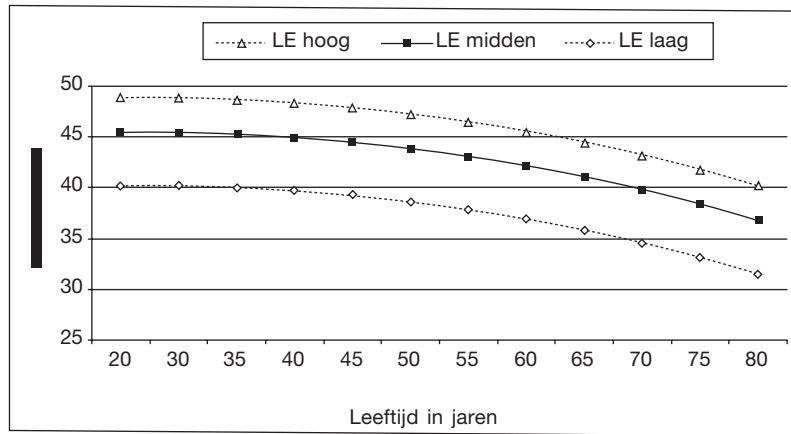
Tabel 2 Multipeleregressiemodellen voor categoriegebonden woordproductie (dieren, beroepen en de somscore van dieren en beroepen).								
categorie	variabele	B	standaardfout B	T	P	bèta	R ²	SD residu
dieren ^a	(constante)	24,630	0,231	106,547	<0,001			
	leeftijd	-0,096	0,009	-10,917	<0,001	-0,246		
	LE (laag)	-2,653	0,332	-7,991	<0,001	-0,200		
	LE (hoog)	1,572	0,345	4,562	<0,001	0,111	0,177	5,800
beroepen ^b	(constante)	18,459	0,251	73,459	<0,001			
	leeftijd	-0,053	0,007	-7,046	<0,001	-0,162		
	leeftijd ²	-0,002	<0,001	-4,153	<0,001	-0,090		
	sekses	0,754	0,228	3,305	0,001	0,070		
	LE (laag)	-2,610	0,276	-9,445	<0,001	-0,235		
	LE (hoog)	1,624	0,289	5,627	<0,001	0,137	0,189	4,823
somscore ^c	(constante)	43,829	0,435	100,673	<0,001			
	leeftijd	-0,142	0,014	-9,828	<0,001	-0,221		
	leeftijd ²	-0,003	0,001	-3,783	<0,001	-0,080		
	LE (laag)	-5,300	0,534	-9,917	<0,001	-0,241		
	LE (hoog)	3,346	0,557	6,010	<0,001	0,143	0,223	9,337

LE = Opleidingsniveau (voor indeling: zie bijlage). Leeftijd = Kalenderleeftijd - 50. Leeftijd² = (Kalenderleeftijd - 50)². Sekse: codeer man als 1, vrouw als 0. LE (laag): codeer een laag opleidingsniveau als 1, gemiddeld of hoog als 0. LE (hoog): codeer een hoog opleidingsniveau als 1, gemiddeld of laag als 0. ^a N = 1823; ^b N = 1824; ^c N = 1822.

Het regressiemodel voor beroepen laat zien dat de basiswaarde (constante) voor beroepen lager is dan voor dieren (18,5 versus 24,6). Het model is iets ingewikkelder omdat het als significante onafhankelijke voorspellers niet alleen leeftijd en opleiding, maar ook het kwadraat van de leeftijd, en sekses bevat. Het significante kwadratische leeftijdseffect betekent dat het effect van leeftijd op beroepen *fluency* groter wordt bij stijgende leeftijd. Mannen noemden gemiddeld ongeveer één beroep meer dan vrouwen. De combinatie van voorspellers verklaart 18,9% van de variantie in het aantal genoemde beroepen.

Het regressiemodel voor de somscore verklaarde met 22,3% de meeste variantie in verge-

lijking met de twee andere modellen. De somscore daalde non-lineair (kwadratisch) als functie van leeftijd, en was hoger of lager voor mensen met een hoger, resp. lager opleidingsniveau. Het non-lineaire (kwadratische) leeftijdseffect wordt zichtbaar in figuur 1. Leeftijd beïnvloedt de *fluency* slechts in geringe mate bij mensen jonger dan 45 jaar. Daarna daalt de *fluency* sneller als functie van leeftijd.



Figuur 1

:Voorspelde somscore (dieren en beroepen samen) als functie van leeftijd en opleidingsniveau (LE).

Voor elk regressiemodel zijn de residuen per leeftijdsgroep geïnspecteerd. Het residu in een regressieanalyse bevat het gedeelte van de voorspelde score dat niet verklaard kan worden vanuit de predictoren in het model. De nauwkeurigheid van de voorspelde score kan afnemen naarmate de waarde van een predictor, in dit geval leeftijd, meer afwijkt van het gemiddelde. De variabiliteit in de residuen van de diverse *fluency*-scores bleek echter niet beïnvloed te worden door leeftijd, wat betekent dat de voorspelde *fluency*-scores voor elke leeftijdsgroep even nauwkeurig zijn, ook voor de 'extreme' groepen van oudste (en jongste) deelnemers.

Toepassing van de MAAS-normen op patiëntresultaten

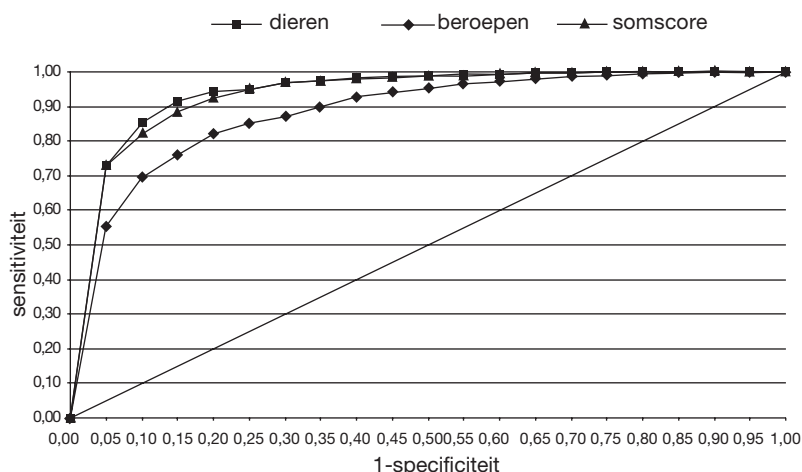
Resultaten naar diagnose. Met behulp van de MAAS-regressiemodellen zijn gestandaardiseerde residuscores (z-scores) berekend voor elk van beide categorieën dieren en beroepen, en voor de somscore. Tabel 3 geeft de gemiddelden en standaarddeviaties voor vier diagnosegroepen (voor cognitief gezonde personen, niet weergegeven in deze tabel, zijn gemiddelde en stan-

daarddeviatie van de z-score theoretisch gelijk aan 0, respectievelijk 1). De gemiddelden verschilden niet voor patiënten met een alzheimerdementie en patiënten met een vasculaire dementie. Dit komt overeen met bevindingen van andere onderzoekers.^{4,25-27} De gestandaardiseerde residuscores voor patiënten met de diagnose 'overige dementie' waren significant lager dan die voor patiënten met een alzheimerdementie. Verschillen tussen patiënten met een vasculaire dementie en patiënten met de diagnose 'overige dementie' waren echter statistisch niet significant. Voor patiënten met een affectieve of cognitieve stoornis (maar geen dementie) was de woordproductie op de drie indicatoren gemiddeld significant hoger dan voor elk van de drie andere diagnosegroepen.

Specificiteit en sensitiviteit voor het onderscheid tussen dementie en intacte cognitie. De specificiteit van een gestandaardiseerde residuscore is bepaald als het percentage MAAS-deelnemers met een hogere gestandaardiseerde residuscore. Bijvoorbeeld: voor de categorie dieren behaalde 95% van de MAAS-deelnemers een gestandaardiseerde residuscore > -1,539. De score -1,539 heeft daarmee een specificiteit van 0,95. Op deze wijze werd ook voor andere z-waarden de specificiteit

diagnose	n	dieren		beroepen		somscore	
		M	SD	M	SD	M	SD
alzheimerdementie	758	-2,03 _a	0,79	-1,65 _d	0,92	-2,09 _g	0,91
vasculaire dementie	132	-2,06 _b	0,89	-1,77 _e	0,97	-2,16 _h	1,01
overige dementie	71	-2,35 _{a, c}	0,97	-2,07 _{d, f}	0,99	-2,49 _{g, i}	1,06
geen dementie	102	-1,13 _{a, b, c}	0,76	-0,90 _{d, e, f}	0,84	-1,14 _{g, h, i}	0,82

F-toetsen voor verschillen in gemiddelde residuscores tussen diagnosegroepen. Gemiddelden met hetzelfde subscript verschillen significant volgens Tukey's test ($p \leq 0,05$).



Figuur 2 Receiver Operating Characteristic (ROC)-curves voor het aantal genoemde dieren, beroepen en de somscore (dieren en beroepen).

bepaald, in dalende stappen van 0,05. Bij elke specificiteit werd de sensitiviteit berekend, als het percentage patiënten dat een score op of onder de z-waarde behaalde. Voor bepaling van de sensitiviteit werden de scores van patiënten met een alzheimerdementie en patiënten met een vasculaire dementie gecombineerd, omdat het gemiddeld aantal genoemde woorden tussen deze diagnosegroepen niet significant verschilde (zie tabel 3). Bijvoorbeeld: bij een gestandaardiseerde residuscore (z-waarde) van -1,539 voor dieren (specificiteit: 0,95) hoort een sensitiviteit van 0,73 omdat 73% van de patiënten met een dementie (alzheimer of vasculair) een gelijke of lagere z-waarde behaalde. Op vergelijkbare wijze werd de sensitiviteit bij opeenvolgende niveaus van specificiteit bepaald. De verhouding specificiteit-sensitiviteit wordt grafisch weergegeven door de sensitiviteit op de verticale as af te zetten tegen 1-specificiteit op de horizontale as. Het resultaat is een *Receiver Operating Characteristic* (ROC)-curve voor de drie parameters: dieren, beroepen en de somscore (zie figuur 2).

Naarmate de ROC-curve verder verwijderd ligt van de diagonaal, is het onderscheidend vermogen van de parameter beter. In een oogopslag is te zien dat de diagnostische nauwkeurigheid van het aantal genoemde dieren groter is dan die van het aantal genoemde beroepen. Bij gelijke specificiteit is het aantal genoemde beroepen een minder gevoelige indicator voor dementie (alzheimer of vasculair) dan het aantal genoemde dieren. De diagnostische nauwkeurigheid van de somscore benadert die van het aantal genoemde dieren. Voor het gebied tussen de ROC-curve en de diagonaal is een kwantitatieve maat berekend, de zogenoemde *area under the curve* (AUC). Voor een perfect onderscheidende test is de AUC gelijk aan 1, en 0,50 voor een niet-onderscheidende test.²⁸ In tabel 4 zijn de AUC-waarden weergegeven voor het onderscheid tus-

sen normale cognitieve veroudering en dementie. Daarbij wordt bovendien onderscheid gemaakt naar ernst van dementie, geoperationaliseerd volgens de 14-itemversie van de Cognitieve Screening Test (CST).²⁹ De items van de CST betreffen oriëntatie in tijd (datum, dag van de week, tijd van de dag), persoonlijke gegevens (woonadres, geboortedatum, leeftijd) en elementaire feitenkennis (regerend vorstin en haar ambtsvoorganger). Voor een differentiatie naar ernst van dementie is de groep patiënten op basis van de mediaanscore van de CST verdeeld in een groep met een score op of onder de mediaan (matig sterke dementie) en een groep met een CST-score boven de mediaan van 10,5 (lichte dementie).

Tabel 4 laat zien dat de AUC voor dierenfluency hoger is dan voor beroepenfluency. Het verschil is statistisch significant, zoals valt af te leiden uit de niet-overlappende 95%-betrouwbaarheidsintervallen. Het onderscheidend vermogen van de fluencytests is groter voor patiënten met een matig sterke dementie dan voor patiënten met een lichte dementie. Dit geldt overigens minder voor dierenfluency dan voor beroepenfluency. Het aantal genoemde dieren heeft in de vergelijking van normale cognitieve veroudering en dementie zowel voor lichte als voor matig sterke dementie een AUC-waarde van ruim boven de 0,90.

Tabel 4 geeft ook de optimale grensscores voor de verschillende fluencytaken en voor twee ernstniveaus van dementie. De optimale grensscore is de gestandaardiseerde residuscore (z-score) met de beste combinatie van sensitiviteit en specificiteit. Zonder te differentiëren naar ernst van dementie (zie het bovenste deel van tabel 4), heeft een gestandaardiseerde residuscore voor dierenfluency van -1,139 de optimale combinatie van hoge sensitiviteit (0,88) en hoge specificiteit

(0,89). Anders gezegd: 88% van patiënten met een dementie (alzheimer of vasculair) behaalde een residuscore minder dan 1,139 standaarddeviatie onder het gemiddelde van een normale vergelijkingsgroep. Voor de categorie beroepen is de optimale grenswaarde -0,850 (met een sensitiviteit van 0,82 en een specificiteit van 0,80), en voor de somscore -1,058 (met een sensitiviteit van 0,88 en een specificiteit van 0,86). In een addendum bij dit artikel wordt de klinische toepassing van deze gegevens toegelicht.

Persevereren over taken

Vijftig van de ruim duizend patiënten in dit onderzoek noemden bij het tweede onderdeel van de fluencytaak (beroepen) opnieuw weer een of meer dieren. Illustratief is het voorbeeld van een 61-jarige man. Gevraagd naar dieren noemde hij er 13 (hond, kat, vogel, mus, meeuw, leeuw, tijger, olifant, bacterie, hermelijn, hagedis, krokodil, piranha). Gevraagd naar beroepen schakelde hij soepel over naar 'arts, psycholoog, vuilnisman, tuinman', om vervolgens zonder aarzeling

Tabel 4 Onderscheidend vermogen van drie fluencycores (dieren, beroepen en de som van het aantal genoemde dieren en beroepen) voor het onderscheid tussen dementie en normale cognitieve veroudering.				
CST ₁₄ (0-14)	AUC (95%-BI)	optimale grensscore	sensitiviteit	specificiteit
dieren	0,95 (0,94-0,96)	-1,139	0,88	0,89
beroepen	0,89 (0,88-0,90)	-0,850	0,82	0,80
somscore	0,95 (0,94-0,95)	-1,058	0,88	0,86

CST ₁₄ ≤10,5	AUC (95%-BI)	optimale grensscore	sensitiviteit	specificiteit
dieren	0,97 (0,96-0,97)	-1,139	0,92	0,89
beroepen	0,93 (0,91-0,94)	-1,080	0,84	0,86
somscore	0,97 (0,96-0,97)	-1,363	0,86	0,93

CST ₁₄ >10,5	AUC (95%-BI)	optimale grensscore	sensitiviteit	specificiteit
dieren	0,93 (0,92-0,94)	-0,955	0,89	0,84
beroepen	0,85 (0,83-0,87)	-0,817	0,77	0,79
somscore	0,92 (0,91-0,94)	-0,866	0,89	0,80

In het bovenste deel van de tabel zijn de gegevens van 890 patiënten gebruikt, ongeacht de ernst van de dementie volgens de Cognitieve Screening Test (CST₁₄). Het middelste deel van de tabel berust op gegevens van 471 patiënten met een matig sterke dementie (CST₁₄ ≤10,5). Het onderste deel van de tabel berust op gegevens van 419 patiënten met een lichte dementie (CST₁₄ >10,5).

AUC = area under the curve, met het 95%-betrouwbaarheidsinterval. De optimale grensscore is de gestandaardiseerde residuscore (z-score) met de beste combinatie van sensitiviteit en specificiteit. Sensitiviteit is het percentage patiënten met een score op of onder de gestandaardiseerde residuscore. Specificiteit is het percentage deelnemers met intacte cognitie dat beter presteerde dan de grensscore.

verder te gaan met ‘haas, konijn, vos, vlieg, bij, dar’. Opmerkelijk is dat hij geen moeite had met de overgang van de ene taak (dieren) naar de andere (beroepen). Hij verschilde hierin niet van de meeste patiënten (96,3% noemde na dieren een of meer beroepen). Ook opmerkelijk is dat de dieren die hij na beroepen noemde, nog niet in de voorafgaande taak waren genoemd. Ook daarmee vormde hij geen uitzondering, want slechts negen (18%) van de vijftig patiënten die afdwaalden naar de voorafgaande categorie herhaalden dezelfde dieren; 82% noemde na beroepen nog spontaan een of meer nieuwe dierexemplaren, alsof het zoeken naar dieren autonoom was doorgestaan. Persevereren kwam bij mannen niet vaker voor dan bij vrouwen, en hield geen verband met leeftijd.

De prevalentie van persevereren was 4,6% bij alzheimerpatiënten, 5,3% bij patiënten met een vasculaire dementie, en 11,3% bij patiënten met een andere vorm van dementie. Het betrof hier patiënten met dementie bij de ziekte van Parkinson (inclusief *Lewy-body*-dementie) en patiënten met een frontotemporale dementie (frontale variant). Bij hen was de kans op perseveratiefouten significant groter dan bij patiënten met een alzheimerdementie (odds ratio en 95%-betrouwbaarheidsinterval: 2,62 (1,17-5,90)). De odds ratio voor de vergelijking van patiënten met een vasculaire dementie en ‘overige dementie’ was 2,27 (0,79-6,54) en verschilde, gelet op het 95%-betrouwbaarheidsinterval, niet significant van 1,00. Ook de vergelijking van patiënten met een alzheimerdementie en een vasculaire dementie leverde geen significant van 1 afwijkende odds ratio op (1,16 (0,50-2,66)). Persevereren over taken kwam niet voor bij patiënten met een affectieve of cognitieve stoornis zonder dementie.

Het is de vraag of persevereren over taken samenhangt met een slechte episodische geheugenfunctie (de actuele taak vergeten), een stoornis van de aandachtscontrole (afdwalen van de actuele taak) of beide. Tabel 5A geeft de vergelijking tussen patiënten die persevereerden en wie dat niet deden. Van de drie episodische geheugentaken liet de test die vraagt om woordelijke weergave van de getoonde items (de Vijf Afbeeldingen Test; 5AT) een significant verschil zien tussen patiënten die terugvielen op de eerste *fluency*taak en degenen die dat niet deden. Een goede of minder goede herkenning van eerder aangeboden items hing echter niet samen met de kans op persevereren over *fluency*taken. Herkenningstests doen, anders dan woordreproductietaken, minder een beroep op executieve functies. Taken die sterker een beroep doen op uitvoerende mentale controle, zoals de meandertekening, de BDS en de EMCT, brengen duidelijke verschillen aan het licht. Deficiëntie op deze taken verhoogt de kans dat een deelnemer persevereert. De odds ratio's wijken significant af van 1,00.

De vijf tests met een significante odds ratio zijn vervolgens in een multivariate analyse opgenomen. Tabel 5B laat zien dat alleen de meandertekening een onafhankelijke en significante bijdrage levert aan de kans op persevereren wanneer de invloed van de andere onafhankelijke variabelen in de logistische regressieanalyse wordt verdisconteerd. Patiënten die in de meandertekening perseveratiefouten maakten (de score < 3 duidt daarop), door niet afwisselend een driehoek en een vierkant te tekenen, maar meerdere driehoeken of vierkanten achter elkaar, hadden een verhoogd risico om ook in de woord*fluency*taak te persevereren.

Tabel 5A	Kenmerken (%) van patiënten die persevereerden over taken (n = 50) en patiënten die dat niet deden (n=1.013).			
	persevereren over taken		OR	95%-BI
	ja	Nee		
uitgestelde herkenning (< 5)	52	52,5 ^b	-	-
uitgestelde herkenning (5)	48	47,5	1,02	0,58-1,80
ADS-Visueel geheugen (> 0 fouten)	60	49,2 ^c	-	
ADS-Visueel geheugen (0 fouten)	40	50,8	0,65	0,36-1,15
5AT (M; SD)	8,14 (3,10)	9,39 (3,13) ^d	0,89*	0,82-0,97
ADS-Meander (< 3)	80	44,8 ^e	-	-

ADS-Meander (≥ 3)	20	55,2	0,20*	0,10-0,41
BDS1-8 (M; SD)	7,96 (4,11) ^a	10,44 (4,10) ^f	0,88*	0,82-0,94
EMCT (M; SD)	12,92 (5,78)	16,04 (5,86) ^g	0,92*	0,88-0,96
Fluency (M; SD)	-2,40 (0,90)	-2,02 (0,97)	0,66	0,49-0,89

5AT = Vijf Afbeeldingen Test (somscore van gereproduceerde items over drie trials); BDS1-8 = Behavioral Dyscontrol Scale (somscore over de eerste acht items). EMCT = Expanded Mental Control Test. Fluency = gestandaardiseerde residuscore van de somscore (dieren plus beroepen). OR = Odds ratio, met het 95% betrouwbaarheidsinterval (BI). [00e2][0080][00a8]* $p \leq 0,05$. ^a $n = 45$; ^b $n = 998$; ^c $n = 1008$; ^d $n = 1007$; ^e $n = 1005$; ^f $n = 978$; ^g $n = 1010$.

Tabel 5B Aanwezigheid van perseveratiefouten. Multivariate analyse (logistische regressie) van voorspellers die volgens univariate analyses een significant van 1 afwijkende odds ratio hadden. $N = 1011$.

predictor	B	SE (B)	OR	95%-BI
(constante)	-1,85	0,91	-	-
ADS-Meander (≥ 3)	-1,29	0,41	0,28*	0,12-0,61
5AT	0,00	0,06	1,00	0,89-1,12
BDS1-8	-0,05	0,05	0,95	0,86-1,04
EMCT	-0,03	0,04	0,97	0,90-1,05
Fluency	-0,05	0,21	0,95	0,63-1,43

* $p \leq 0,05$. De verklaarde variantie (R^2) van het model is 3,01%.

Discussie

Voor de klinische praktijk zijn de normaalwaarden uit de *Maastricht Aging Study* (MAAS) van grote betekenis (zie het addendum bij dit artikel). Zij helpen de clinicus beter dan algemene aanbevelingen in de literatuur, zoals 'For the category animals, normal subjects usually produce 20 exemplars'. Met behulp van de gegevens in tabel 2 kan worden bepaald hoe de score van een individuele patiënt zich verhoudt tot de verdeling van scores bij 'normalen', met een correctie voor variabelen zoals geslacht, leeftijd en opleiding. Deze toepassing is niet beperkt tot onderzoek van oudere patiënten, of onderzoek van dementie, maar strekt zich ook uit tot neuropsychologisch onderzoek van jongere volwassenen, bijvoorbeeld bij hersentrauma of beroerte. Nieuw in dit onderzoek is dat de fluencytaken zijn onderzocht op hun vermogen onderscheid te maken tussen dementie en normale cognitieve veroudering. Het onderscheidend vermogen van het aantal opgenoemde dieren is beter dan van beroepen. Bij gelijke specificiteit, en ongeacht de ernst van de

dementie, is het aantal genoemde beroepen een minder gevoelige indicator voor dementie dan het aantal genoemde dieren. Anders gezegd: verschillen tussen patiënten en leeftijdgenoten met een intacte cognitie worden bij het opnoemen van dieren beter zichtbaar dan bij beroepen.

Over de mogelijke oorzaak van dit verschil kan zonder nader onderzoek slechts worden gespeculeerd. Recent onderzoek levert steun op voor de hypothese dat woorden waarmee dieren worden aangeduid neurale gerepresenteerd zijn in de onderste winding van de linker frontotemporale cortex, anders dan woorden uit andere categorieën, die een minder specifieke, meer gespreide neurale representatie kennen. Lexicale categorieën die worden ondersteund door een gespreid neurale netwerk zouden de beginfase van een degeneratieve hersenziekte beter 'overleven' dan categorieën die gerepresenteerd zijn in hersengebieden die het eerst getroffen worden, zoals de temporaalkwab bij de ziekte van Alzheimer.

Conceptueel is de definitie van wat mag worden aangemerkt als een geldig beroep enigszins

diffuus. Volgens de handleiding bij de ADS en de GIT worden mannelijke en vrouwelijke vormen van hetzelfde beroep (zoals steward en stewardess) beide meegeteld. Ook wanneer een beroep een synoniem of specificatie is van een ander woord, worden beide goed gerekend (zoals verpleegster en zuster, of fietsenmaker en rijwielhersteller). Zelfs niet geheel juiste formuleringen, zoals aanduidingen van bedrijven ('kledingwinkel', 'benzinepomp') worden als juist geaccepteerd. Misschien noemen patiënten wat vaker specifieke termen voor beroepen (zoals mannelijke en vrouwelijke benamingen, bedrijfsaanduidingen, of omschrijvingen) dan mensen met een intacte cognitie, die zichzelf hierin misschien wat strengere criteria opleggen. Het effect kan zijn dat kwantitatieve verschillen tussen patiënten en anderen met een intacte cognitie kleiner worden. Als de scoring van 'beroepen' minder eenvoudig en eenduidig is dan die van 'dieren', zou dat een argument kunnen zijn om in de klinische praktijk minder op 'beroepen' te vertrouwen dan op 'dieren'. De geringere sensitiviteit van het aantal genoemde beroepen roept dan ook twijfels op over de klinische bruikbaarheid van deze categorie.

Dieren, of dieren en beroepen?

Voor het onderscheid tussen dementie en intacte cognitie is de sensitiviteit van beroepen bij gelijke specificiteit minder dan de sensitiviteit van dieren. Wat is dan de toegevoegde waarde van een extra categorie? De klinisch onderzoeker kan op basis van dit onderzoek besluiten om de woordproductie met de meest sensitieve categorie te testen. Er zijn in de internationale literatuur diverse rapporten over de toepasbaarheid van een enkele categorie. Ook voor het Nederlandse taalgebied zijn er voorbeelden van tests die woordproductie enkel via 'dieren' testen. Daartegenover staat dat het gebruik van twee categorieën (dieren en beroepen) een integraal onderdeel is van bestaande tests, zoals de GIT, de ADS en de Woord-Fluency Test, waarbij in de twee laatste alleen de somscore over twee categorieën betekenis heeft. Dit onderzoek toont aan dat het onderscheidend vermogen van de somscore nauwelijks verschilt van dat van de best onderscheidende subtest (dieren opnoemen), waarmee de klinische toepassing van de somscore wordt ondersteund.

Door toepassing van twee categorieën kon in dit onderzoek bovendien het klinisch interessante verschijnsel van perseveratie over taken worden onderzocht. Dit type gedrag kwam in de hier onderzochte psychogeriatrische populatie echter weinig voor. Andere onderzoekers hebben dat ook gemerkt, maar omdat hun steekproeven kleiner waren, zagen zij ervan af dit symptoom nader te analyseren. In dit onderzoek waren er echter voldoende patiënten ($n=50$) van wie via een logistische regressieanalyse kon worden vastgesteld dat het optreden van perseveraties

samenhangt met afwijkende resultaten op tests voor cognitieve controle. Daarmee wordt bevestigd dat terugvallen op een voorafgaande taak een klinische aanwijzing is voor een stoornis van de uitvoerende mentale controlefuncties, met name van responsinhibitie. Perseveren over taken heeft dus een klinische betekenis, maar is door de zeer lage prevalentie in de klinische praktijk niet zo bruikbaar. Voor het onderzoek van executieve functies blijven specifiek daarop gerichte methoden dan ook onmisbaar.

Beperkingen van het onderzoek

Dit onderzoek was niet primair opgezet om methoden te vinden voor een betrouwbaar onderscheid tussen dementie en normale veroudering. De klinische bruikbaarheid van de MAAS-normen is onderzocht in een gelegenheidssteekproef, aan de hand van een grote hoeveelheid beschikbare patiëntgegevens. De diagnosegegevens waren afkomstig van verschillende klinieken, wat variatie in de betrouwbaarheid van de diagnose met zich mee kan brengen. Daar staat tegenover dat de betrouwbaarheid van de diagnostiek in het algemeen is toegenomen doordat klinici gebruikmaken van richtlijnen voor de diagnostiek van dementie. Het valt niet uit te sluiten dat *fluency*gegevens gebruikt zijn bij het stellen van een diagnose, omdat *fluency*-tests in veel ziekenhuizen en geheugenpoli's een vast onderdeel zijn van het neuropsychologisch onderzoek. Hiertegen kan worden ingebracht dat de diagnose dementie gesteld wordt op basis van diverse criteria, zoals meervoudige cognitieve stoornissen (van episodisch geheugen, taal, waarneming, handelen en oordeelsvermogen), gedragsveranderingen en beperkingen in het uitvoeren van instrumentele taken in het dagelijks leven.

De ernst van de dementie varieerde in de onderzochte patiëntengroep van licht tot matig sterk. Het onderscheidend vermogen van cognitieve tests is voor lichte dementie meestal minder groot dan voor matig sterke dementie. Dat blijkt ook in dit onderzoek (zie tabel 4), wanneer onderscheid wordt gemaakt tussen patiënten met relatief gunstige resultaten op de Cognitieve Screening Test (CST) en patiënten met een minder goed resultaat. Echter voor de categorie dieren liggen de AUC-waarden voor het onderscheid tussen lichte dementie, respectievelijk matig sterke dementie en normale cognitieve veroudering niet ver uit elkaar. Het lijkt dan ook weinig zinvol om in de klinische praktijk andere grensscores te gebruiken bij een 'lichte' dementie dan bij een 'matig sterke' dementie. Een bijkomend probleem is dat zonder nader onderzoek vooraf niet is vast te stellen of iemand een 'lichte' of 'matig sterke' dementie heeft.

In de patiëntengroep was 35% ouder dan 81 jaar, dat wil zeggen ouder dan de oudste deelnemer aan het onderzoek waarop de MAAS-normen zijn gebaseerd. Het aantal patiënten

met een leeftijd hoger dan 81 nam echter snel af, via 17% van 85 jaar of ouder tot 2% van 90+. Analyse van de residuen per leeftijdsgroep gaf geen aanwijzingen voor een grotere onnauwkeurigheid van voorspelde *fluency* scores bij mensen in de oudste leeftijdsgroep. Dit rechtvaardigt vertrouwen in de voorspellingen van de regressie-modellen voor mensen in de oudste leeftijdsgroepen. Zonder empirische gegevens voor mensen ouder dan 81 blijft echter onzeker in hoeverre extrapolatie van de MAAS-normen verantwoord is. Dit probleem is overigens niet inherent aan de hier gebruikte methode van regressie-analyse, maar doet zich ook voor bij de klassieke

opzet, wanneer een onderzochte een leeftijd heeft die uitstijgt boven die van de oudste deelnemer in de normtabel.

We hechten eraan te benadrukken dat resultaten van een *fluency* test niet toereikend zijn voor een eventuele diagnose van dementie. Er kunnen tal van redenen zijn voor een lage *fluency* score, zoals bijvoorbeeld afasie, tempovertraging bij ziekte van Parkinson, of concentratieverlies bij depressie. Dit neemt niet weg dat een zorgvuldige interpretatie van resultaten bij categoriegebonden woordproductie gebaat is bij de beschikbaarheid van de hier beschreven normgegevens.

Literatuur

- Rogers S, Friedman RB. The underlying mechanisms of semantic memory loss in Alzheimer's disease and semantic dementia. *Neuropsychologia* 2008;46:12-21.
- Rosen VM, Engle RW. The role of working memory capacity in retrieval. *Journal of Experimental Psychology: General* 1997;126:211-227.
- Henry JD, Crawford JR, Phillips LH. Verbal fluency performance in dementia of the Alzheimer's type: a meta-analysis. *Neuropsychologia* 2004;42:1212-1222.
- Giovannetti T, Schmidt KS, Gallo JL, Sestito N. Everyday action in dementia: evidence for differential deficits in Alzheimer's disease versus subcortical vascular dementia. *Journal of the International Neuropsychological Society* 2006;12:45-53.
- Ramsden CM, Kinsella GJ, Ong B, Storey E. Performance of everyday actions in mild Alzheimer's disease. *Neuropsychology* 2008;22:17-26.
- Haaxma R. Neurologie van cognitie en gedrag in hoofdlijnen. Tweede, herziene druk. Maarsse: Elsevier Gezondheidszorg, 2008.
- Hodges JR. Cognitive assessment for clinicians. Second edition. Oxford: Oxford University Press, 2007.
- Van der Elst W, Van Boxtel MPJ, Van Breukelen GJP, Jolles J. Normative data for the Animal, Profession and Letter M Naming verbal fluency tests for Dutch speaking participants and the effects of age, education, and sex. *Journal of the International Neuropsychological Society* 2006;12:80-89.
- Luteijn F, Barelds DPH. GIT2 Groninger Intelligentie Test 2. Amsterdam: Harcourt Test Publishers, 2004.
- Lindeboom J, Jonker C. Amsterdamse Dementie-Screeningstest (ADS6). In: Bouma A, Mulder J, Lindeboom J, editors. *Neuropsychologische diagnostiek: handboek*. Lisse: Swets & Zeitlinger, 1998: I.1-I.9.
- Mulder JL, Dekker PH, Dekker R. Woord-Fluency Test & Figuur-Fluency Test. Handleiding. Leiden: PITS, 2006.
- Crawford JR, Garthwaite PH. Using regression equations built from summary data in the neuropsychological assessment of the individual case. *Neuropsychology* 2007;21:611-620.
- Troyer AK. Normative data for clustering and switching on verbal fluency tasks. *Journal of Clinical and Experimental Neuropsychology* 2000;22:370-378.
- Faust ME, Balota DA. Inhibition, facilitation, and attentional control in dementia of the Alzheimer's type. The role of unifying principles in cognitive theory development. In: Gorfein DS, MacLeod CM, editors. *Inhibition in cognition*. Washington, DC: American Psychological Association, 2007: 213-238.
- Sandson J, Albert ML. Perseveration in behavioral neurology. *Neurology* 1987;37:1736-1741.
- Pekkala S, Albert ML, Spiro A, Erkinjuntti T. Perseveration in Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders* 2008;25:109-114.
- De Bie SE. Standaardvragen 1987. Voorstellen voor uniformering van vraagstellingen naar achtergrondkenmerken in interviews. Leiden: Vereniging van Onderzoek Instituten, 1987.
- Verhage F. Intelligentie en leeftijd bij volwassenen en bejaarden. Assen: Van Gorcum, 1964.
- American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*. Washington, DC: American Psychiatric Association, 1994.
- Lindeboom J, Jonker C. Amsterdamse Dementie-Screeningstest. Lisse: Swets and Zeitlinger, 1989.
- Diesfeldt HFA. Constructvaliditeit van enkele tests voor episodisch geheugen in de psychogeriatric. *Tijdschrift voor Gerontologie en Geriatrie* 2006;37:59-66.
- Lindeboom J, Koene T, Matto D. De diagnostische waarde van tests voor mentale controle. *Tijdschrift voor Gerontologie en Geriatrie* 1993;24:105-109.

- 23 Grigsby J, Kaye K, Robbins LJ. Reliabilities, norms and factor structure of the Behavioral Dyscontrol Scale. *Perceptual and Motor Skills* 1992;74:883-892.
- 24 Diesfeldt HFA. Executive functioning in psychogeriatric patients: scalability and construct validity of the Behavioral Dyscontrol Scale (BDS). *International Journal of Geriatric Psychiatry* 2004;19:1065-1073.
- 25 Duff Canning SJ, Leach L, Stuss D, Ngo L, Black SE. Diagnostic utility of abbreviated fluency measures in Alzheimer's disease and vascular dementia. *Neurology* 2004;62:556-562.
- 26 Graham NL, Emery T, Hodges JR. Distinctive cognitive profiles in Alzheimer's disease and subcortical vascular dementia. *Journal of Neurology, Neurosurgery and Psychiatry* 2004;75:61-71.
- 27 Jonsson Laukka E, Jones S, Small BJ, Fratiglioni L, Bäckman L. Similar patterns of cognitive deficits in the preclinical phases of vascular dementia and Alzheimer's disease. *Journal of the International Neuropsychological Society* 2004;10:382-391.
- 28 De Koning I, Luteijn F. Psychometrie. In: Hendriks M, Kessels R, Gorissen M, Schmand B, editors. *Neuropsychologische diagnostiek*. Amsterdam: Boom, 2006: 103-118.
- 29 De Graaf A, Deelman BG. *Cognitieve Screening Test*. Lisse: Swets en Zeitlinger, 1991.
- 30 Venneri A, McGeown WJ, Hietanen HM, Guerrini C, Ellis AW, Shanks MF. The anatomical basis of semantic retrieval deficits in early Alzheimer's disease. *Neuropsychologia* 2008;46:497-510.
- 31 Gomez RG, White DA. Using verbal fluency to detect very mild dementia of the Alzheimer type. *Archives of Clinical Neuropsychology* 2006;21:771-775.
- 32 Grober E, Hall C, McGinn M, Nicholls T, Stanford S, Ehrlich A, et al. Neuropsychological strategies for detecting early dementia. *Journal of the International Neuropsychological Society* 2008;14:130-142.
- 33 Skjerve A, Nordhus IH, Engedal K, Pallesen S, Braekhus A, Nygaard HA. Seven minute screen performance in a normal elderly sample. *International Journal of Geriatric Psychiatry* 2007;22:764-769.
- 34 Cummings JL. The one-minute mental status examination. *Neurology* 2004;62:534-535.
- 35 Deelman BG, Koning-Haanstra M, Liebrand WBG, Van den Burg W. *SAN Test. Een afasietest voor auditief taalbegrip en mondeling taalgebruik*. Lisse: Swets & Zeitlinger, 1981.
- 36 Peters MJV, Ponds RWHM, Van den Ham P, Scheltens P, Verheij FRJ. Detectie van dementie van het Alzheimer type (DAT) met de '7 Minuten Test' (7MT). *Tijdschrift voor Gerontologie en Geriatrie* 2004;35:114-120.
- 37 Baldo JV, Shimamura AP. Letter and category fluency in patients with frontal lobe lesions. *Neuropsychology* 1998;12:259-267.
- 38 Possin KL, Filoteo JV, Roesch SC, Zizak V, Rilling LM, Davis JD. Is a perseveration a perseveration? An evaluation of cognitive error types in patients with subcortical pathology. *Journal of Clinical and Experimental Neuropsychology* 2005;27:953-966.
- 39 Kwaliteitsinstituut voor de Gezondheidszorg CBO. *Diagnostiek en medicamenteuze behandeling van dementie*. Utrecht: Kwaliteitsinstituut voor de Gezondheidszorg CBO, 2005.
- 40 Crawford JR. <http://www.abdn.ac.uk/~psy086/dept/regdisc.htm>.
- 41 Crawford JR, Garthwaite PH. Comparing patients' predicted test scores from a regression equation with their obtained scores: a significance test and point estimate of abnormality with accompanying confidence limits. *Neuropsychology* 2006;20:259-271.
- 42 Bland M. *An introduction to medical statistics*. Third edition. Oxford: Oxford University Press, 2000.
- 43 Menard S. *Applied logistic regression analysis*. Second edition. Thousand Oaks: Sage, 2002.