

Neuropsychologisch onderzoek van bejaarden: een follow-up studie

door J. Lindeboom en C. Jonker

Inleiding

In een eerder artikel (Lindeboom, van de Wijngaard en Jonker, 1979) werd verslag uitgebracht van de opstelling van een eenvoudige neuropsychologische screeningsprocedure voor bejaarden. De procedure bestaat uit een serie gestandaardiseerde proeven, waarop cerebrale funktiestoornissen van uiteenlopende aard tot uiting kunnen komen. De onderdelen zijn:

1. oriëntatie (tijd en plaats),
2. sorteren,
3. meander (voortzetten van een afwisselend patroon),
4. fluency (opnamen van dieren en beroepen),
5. hand-eye-ear (uitvoeren van opdrachten met betrekking tot lichaamsschema),
6. natekenen van figuren,
7. visuele inprenting (inprenten en onthouden van plaatjes).

Om met dit materiaal afwijkingen op te kunnen sporen is het uiteraard noodzakelijk om te weten welk prestatieniveau als slecht kan worden beschouwd. Hiertoe werd een voorlopige ijking uitgevoerd bij een groep ingezetenen van een bejaardentehuis. Cutoff-scores werden gebaseerd op de resultaten die in deze groep werden behaald. Lage prestaties werden nader geanalyseerd op hun onderlinge samenhang en hun verband met opleiding, geslacht en leeftijd. De onderzoeksresultaten suggereren dat de procedure geschikt is voor de beoogde doeleinden, mits validatie met klinische gegevens plaatsvindt.

Vraagstelling van het follow-up onderzoek

Dit verslag betreft een herhalingsonderzoek dat 19 maanden na het eerste onderzoek plaatsvond. Met deze follow-up werd een antwoord gezocht op de volgende vragen:

Schrijvers zijn respectievelijk werkzaam als neuropsycholoog op de afdeling neurologie en als neuroloog op de afdeling neuropsychiatrie van de Valeriuskliniek, Valeriusplein 9, 1075 BG Amsterdam.

1. Zijn de testresultaten stabiel? In het bijzonder gaat het om de vraag of een eenmaal geregistreerde lage score bij herhaling van het onderzoek nog steeds laag is. Hoewel bij een tijdsinterval van deze lengte niet gesproken kan worden van een betrouwbaarheids-onderzoek, is het van belang na te gaan in hoeverre de scores voor-
 spellende waarde bezitten voor prestaties op een latere datum. De stabiliteit werd onderzocht door
 - a. vergelijking van de test- en hertestuitslagen per onderdeel, en
 - b. vergelijking van het aantal lage scores dat elke proefpersoon bij test en hertest behaalde.
2. Kunnen de testresultaten gegeneraliseerd worden naar gedrag in ruimere zin? Om dit te onderzoeken werden de testresultaten vergeleken met een gedragsbeoordeling volgens de B.O.P. (Beoordelingsschaal voor Oudere Patiënten, van der Kam, Mol, en Wimmers, 1971).
3. Zijn de resultaten van de Visuele Inprentingstest vergelijkbaar met de resultaten van een verbale inprentingstest? Daar de beoordeling van de geheugenfunctie meestal het grootste diagnostische probleem vormt, leek nadere aandacht hiervoor gewettigd. Tevens werd behoefte gevoeld aan een inprentingstest met een minder sterk plafond-effect. Er werd gebruik gemaakt van een tot 8 woorden verkorte versie van de 15-woordentest (Kalverboer en Deelman, 1964). De taak bestaat uit het memoriseren van een reeks woorden.

Proefpersonen

Van de oorspronkelijke groep konden in totaal 72 personen opnieuw getest worden. De redenen voor uitval zijn opgesomd in tabel 1

Tabel 1: Verloop van de proefgroep

Complete hertest	69
Incomplete hertest (1 à 2 onderdelen niet)	3
Overleden	14
Overgeplaatst naar ander bejaardentehuis	5
Overgeplaatst naar verpleegtehuis voor demente bejaarden	1
Opgenomen in ziekenhuis	3
Sterke achteruitgang in visus	1
Weigering	5
Niet te bereiken	3
Totaal	104

Ten aanzien van de ene proefpersoon die overgeplaatst werd naar een verpleegtehuis voor demente bejaarden, kan opgemerkt worden dat deze in principe correct geïdentificeerd werd door de test-procedure, daar hij slecht scoorde op 4 van de 7 onderdelen. Overigens lijken er geen systematische verschillen te bestaan tussen de afgevallen en de overblijvende proefpersonen; er zijn al-

thans geen statistisch aantoonbare verschillen in de mate waarin binnen beide groepen lage scores werden behaald tijdens het eerste onderzoek.

Resultaten

I. Stabiliteit per onderdeel

De resultaten per onderdeel zijn aangegeven in tabel 2. Voor de beoordeling van de scores als goed of slecht (+ of —) werden dezelfde cuttingscores gebruikt als bij het eerste onderzoek. De statistische significantie werd berekend met de Fisher Exact Probability test.

Tabel 2: Verband tussen uitslagen van test en hertest

	A	B	C	D	Tot.	p
uitslag 1e onderzoek	+	—	+	—		
uitslag follow-up	+	+	—	—		
Oriëntatie	61	7	1	3	72	.007
Inprenting	52	3	11	6	72	.004
Meander	54	2	8	5	69	.002
Sorteren	29	0	37	5	71	.065
Fluency	57	5	5	5	72	.003
Hand-eye-ear	58	9	2	2	71	.100
Natekenen	63	5	2	1	71	.218

Voor Oriëntatie, Inprenting, Meander en Fluency is de overeenkomst significant, hetgeen wijst op een zekere stabiliteit van de resultaten.

Kolom B toont de 'valse positieven', afwijkende scores die bij hertest verbeterd bleken. Op het kleine aantal lage scores vormen deze valse positieven soms een aanzienlijk percentage, met name bij Oriëntatie, Hand-eye-ear en Natekenen. Wat de laatste twee onderdelen betreft, ging het steeds om scores die net beneden de cuttingscore vielen, dat wil zeggen een score van 6 punten op Hand-eye-ear en 9 punten op Natekenen. Aan scores van deze hoogte kan daarom bij screening weinig waarde worden gehecht. De getallen in kolom C zijn moeilijk te interpreteren; door de grote tijdsduur tussen test en hertest kan het voor een deel om een inmiddels ontwikkelde achteruitgang gaan. Het aantal gedaalde scores op Sorteren is echter absurd hoog, en waarschijnlijk te wijten aan verschillen tussen proefleiders. We moeten daarom besluiten dat deze test niet objectief kan worden afgenomen, en om die reden onbruikbaar is.

II. Stabiliteit van het aantal lage scores

Tabel 3 toont de overeenkomst in het al of niet behalen van lage scores bij test en hertest.

Tabel 3: Verband tussen test en hertest in al of niet laag scoren

		uitslagen eerste onderzoek		
		ppn. zonder lage scores	ppn. met lage scores	
uitslagen tweede onderzoek	ppn. zonder lage scores	30	8	38
	ppn. met lage scores	8	23	31
		38	31	69

We zien in de eerste plaats dat in beide onderzoeken evenveel proefpersonen één of meer lage scores behaalden. Het verband tussen test en hertest is statistisch significant ($\chi^2 = 17.4$, $p < .001$), hetgeen wil zeggen dat degenen die bij het eerste onderzoek laag scoorden, dit bij herhaling ook neigden te doen.

In tabel 4 is het aantal lage scores per pp. nader gespecificeerd.

Tabel 4: Verband tussen test en hertest in aantal lage scores

		eerste onderzoek, aantal lage scores per pp.			
		0	1	>1	
tweede onderzoek, aantal lage scores per pp.	0	30	8	0	38
	1	7	9	5	21
	>1	1	2	7	10
		38	19	12	69

Alle gevallen waarin aanvankelijk meer dan één lage score werd behaald, vertonen bij hertest minstens één lage score. De gevallen waarin aanvankelijk slechts één lage score werd geregistreerd, vertonen echter aanzienlijk minder stabiliteit; in 8 gevallen, dit is 42 % waren er bij hertest geen afwijkingen. Van deze 8 gevallen betrof het viermaal een 'vals positieve' score op de Hand-eye-ear test (zie paragraaf I).

III. Generaliseerbaarheid van testresultaten

Subschaal 1 van de Beoordelingsschaal voor Oudere Patiënten (B.O.P.) geeft een maat voor de geobserveerde 'hulpbehoefte'. De scores op deze schaal zijn in zoverre geassocieerd met de mentale status, dat zij blijken te differentiëren tussen het arts-oordeel 'dement' en 'niet-dement' bij bejaarden in een Psychiatrische kliniek (Schouten, 1978). Een overeenkomst tussen de testresultaten en B.O.P.-scores zou betekenen dat de test afwijkingen registreert die verband houden met het gedrag van de proefpersoon in het dagelijks leven.

(a) Correlatie van BOP-scores en testprestaties. Er werd een correlatie berekend tussen de hoogte van de B.O.P. score en het gemiddelde aantal lage test scores dat behaald werd door proefpersonen met een score van die hoogte. Deze correlatie bedraagt .74 en is significant (Spearman rho, $p < .0005$). Er bestaat dus een sa-

menhang tussen testgedrag en 'hulpbehoevendheid' volgens de B.O.P.

(b) Verband tussen B.O.P. en specifieke onderdelen. Het verband tussen slecht scoren op schaal 1 van de B.O.P. en slecht scoren op aparte onderdelen werd onderzocht door toepassing van de Fisher Exact Probability test. Hierbij werden twee willekeurig gekozen cuttingscores voor de B.O.P. genomen: scores hoger dan 11 punten (16 % van de ppn) en scores hoger dan 12 punten (10 % van de ppn).

Tabel 5 toont de gevonden waarden van p.

Tabel 5: Verband tussen B.O.P. en testcores

onderdeel	cuttingscore	
	11	12
Oriëntatie	.012	.003
Vis.inprenting	.009	.004
Meander	.081	.111
Fluency	.007	.006
Hand-eye-ear	.275	.187
Natekenen	.579	.714

Het verband is duidelijk bij Oriëntatie, Inprenting en Fluency, gering bij Meander, en afwezig bij Hand-eye-ear en Natekenen. Ondanks minder stabiele resultaten lijkt Oriëntatie toch relevante informatie over de toestand van de pp. op te leveren. Hand-eye-ear en Natekenen leveren noch generaliseerbare, noch stabiele resultaten op. Hoewel dit voor een klinische populatie niet erg op hoeft te gaan, is voorzichtigheid geboden bij de interpretatie van deze of vergelijkbare proeven.

IV. Resultaten van de 8-woordentest

De betekenis van lage scores op de 8-woordentest werd eveneens onderzocht door vergelijking met de andere testonderdelen. Cuttingscores waren hier: lager dan 21 punten (17 % van de ppn.) en lager dan 20 punten (10 % van de ppn.).

Tabel 6: Verband tussen resultaten van 8-woordentest en overige onderdelen; waarden van p op Fisher's test

onderdeel	cuttingscore	
	21	20
Oriëntatie	.004	.0006
Vis. Inprenting	.018	.036
Meander	.025	.014
Fluency	.033	.035
Hand-eye-ear	.116	.300
Natekenen	.360	.057
(B.O.P.)	.10	.012

De uitslagen van de 8-woordentest vertonen een groot aantal significante verbanden. De Visuele Inprentingstest is daarentegen alleen significant geassocieerd met Oriëntatie. De visuele inprentingstest lijkt daarom een meer zuivere test voor de geheugenfunctie, terwijl de 8-woordentest eerder iets zegt over het globale niveau van functioneren.

Conclusies

In toepassing op deze niet-klinische groep bejaarden levert de onderzoeksprocedure redelijk stabiele resultaten op, die generaliseerbaar zijn naar gedrag in ruimere zin. Dit pleit voor de bruikbaarheid van de procedure bij klinisch onderzoek in een gerontologische of psychiatrische setting.

De stabiliteit en generaliseerbaarheid van de verschillende onderdelen is wisselend. De beste resultaten leverden Visuele Inprenting en Fluency. Oriëntatie is minder stabiel, maar een lage score daarop is niettemin relevant. Sorteren leverde zodanig inconsistente resultaten op, dat deze test uit de procedure moest worden verwijderd. Meander is stabiel, maar de relevantie is minder duidelijk. De resultaten van Hand-eye-ear en Natekenen zijn tot dusver onbevredigend; mogelijk kunnen zij echter in een klinische setting nuttiger blijken, waar meer en ernstiger afwijkingen kunnen voorkomen dan bij de onderhavige proefgroep. De 8-woordentest kan nonspecifieke indicaties van gestoordheid opleveren, maar kan niet beschouwd worden als een zuivere maat van de geheugenfunctie.

Literatuur

- Kalverboer, A. F., B. G. Deelman (1968), *Vijftienwoordentest, Testdocumentatieklapper deel I*, Afd. Klinische Psychologie, R.U. Groningen.
- Van de Kam, P., F. Mol en M. F. H. G. Wimmers (1971), *Beoordelingschaal voor oudere patiënten*. Van Loghum Slaterus, Deventer.
- Lindeboom, J., D. van de Wijngaard en C. Jonker (1979), Een neuropsychologische procedure voor dementie-onderzoek, *T. voor Psychiatrie*, 21, p. 241-252.
- Schouten, P. (1978), *Onderzoek naar de samenhang tussen het artsenoordeel 'dementie' en — veranderde — B.O.P.-scores*, Interne Publikatie Wilhelmina Gasthuis, Psychiatrische Kliniek, afdeling Psychologie, Amsterdam.

Met dank aan personeel en inwoners van huize Brentano te Amsterdam en aan drs. J. v. d. Broek voor het testwerk.