

Bedankt voor het downloaden van dit artikel. De artikelen uit de (online)tijdschriften van Uitgeverij Boom zijn auteursrechtelijk beschermd. U kunt er natuurlijk uit citeren (voorzien van een bronvermelding) maar voor reproductie in welke vorm dan ook moet toestemming aan de uitgever worden gevraagd.

# Boom

Behoudens de in of krachtens de Auteurswet van 1912 gestelde uitzonderingen mag niets uit deze uitgave worden verveelvoudigd, opgeslagen in een geautomatiseerd gegevensbestand, of openbaar gemaakt, in enige vorm of op enige wijze, hetzij elektronisch, mechanisch door fotokopieën, opnamen of enig andere manier, zonder voorafgaande schriftelijke toestemming van de uitgever.

Voor zover het maken van kopieën uit deze uitgave is toegestaan op grond van artikelen 16h t/m 16m Auteurswet 1912 jo. Besluit van 27 november 2002, Stb 575, dient men de daarvoor wettelijk verschuldigde vergoeding te voldoen aan de Stichting Reprorecht te Hoofddorp (postbus 3060, 2130 KB, [www.reprorecht.nl](http://www.reprorecht.nl)) of contact op te nemen met de uitgever voor het treffen van een rechtstreekse regeling in de zin van art. 16l, vijfde lid, Auteurswet 1912.

Voor het overnemen van gedeelte(n) uit deze uitgave in bloemlezingen, readers en andere compilatiewerken (artikel 16, Auteurswet 1912) kan men zich wenden tot de Stichting PRO (Stichting Publicatie- en Reproductierechten, postbus 3060, 2130 KB Hoofddorp, [www.cedar.nl/pro](http://www.cedar.nl/pro)).

*No part of this book may be reproduced in any way whatsoever without the written permission of the publisher.*

[info@boomamsterdam.nl](mailto:info@boomamsterdam.nl)  
[www.boomuitgeversamsterdam.nl](http://www.boomuitgeversamsterdam.nl)

# *De bruikbaarheid van agressievragenlijsten in de Nederlandse forensisch- psychiatrische praktijk*

GERJONNE AKKERMAN-BOUWSEMA, BRENDA BIESBROEK,  
MARIKE LANCEL & COBY GERLSMA

## *Samenvatting*

Agressie is een veelvoorkomend probleem in de forensische psychiatrie en daarom is het belangrijk dat agressie en gewelddadig gedrag van forensisch-psychiatrische patiënten kan worden beoordeeld en nauwkeurig worden voorspeld. In deze review worden de psychometrische eigenschappen van zelfrapportage agressievragenlijsten geëvalueerd. Bijna alle vragenlijsten beschikken over voldoende interne consistentie en test-hertestbetrouwbaarheid. Echter, onderzoek naar de predictieve validiteit van de vragenlijsten onder een volwassen (forensisch) klinische populatie bleek schaars. Daarnaast bleken er geen normen beschikbaar voor forensisch-psychiatrische patiënten en werd het risico van sociale wenselijkheid niet onderzocht. Empirisch onderzoek naar de predictieve waarde van deze vragenlijsten in de forensische psychiatrie en het opstellen van normen voor de forensisch-psychiatrische patiënten zijn wenselijk.

*Trefwoorden: agressie, zelfrapportagevragenlijsten, forensische psychiatrie, psychometrische eigenschappen, betrouwbaarheid en validiteit*

## INLEIDING

In de forensische psychiatrie worden personen behandeld die in het verleden, mede op basis van een psychische stoornis, een geweldsdelict hebben gepleegd. Een groot deel van deze patiënten kampt met agressieregulatieproblemen. De behandeling van deze patiënten, waarbij het voorkómen van recidive vooropstaat, richt zich voor een belangrijk deel op het (h)erkennen en leren omgaan met agressieve gedachten, gevoelens en gedrag (de Ruiter

(Hildebrand, 1999; Hornsveld, Soe-Agnie, Donker, & van der Wal, 2008). Voor het adequaat behandelen van agressieproblematiek is diagnostiek van groot belang: welke factoren dragen onder welke omstandigheden bij aan het agressieve gedrag van een patiënt? Theorieën over agressie kunnen richtinggevend zijn bij het opstellen van hypothesen over deze vraagstukken.

Agressief gedrag kan worden gedefinieerd als een handeling met het doel schade toe te brengen aan een persoon of object, anders dan uit zelfverdediging (Anderson & Bushman, 2002). Agressie kan zich daarbij voordoen in verschillende vormen. Een belangrijk onderscheid is reactieve versus proactieve agressie. Reactieve agressie is een emotionele, boze reactie op een waargenomen bedreiging, aanval of frustratie, bijvoorbeeld een woedeaanval na het verliezen van een spelletje. Bij proactieve agressie gaat het om het inzetten van agressie om een ander doel te bereiken dan het toebrengen van schade op zichzelf, bijvoorbeeld het dreigen met geweld om een telefoon te bemachtigen. Daarnaast wordt veelal onderscheid gemaakt tussen directe en indirecte agressie. Onder directe agressie vallen fysiek agressieve gedragingen, zoals slaan en duwen, alsmede verbaal agressief gedrag, zoals schelden en bedreigen. Indirecte agressie kan worden geconceptualiseerd als een alternatieve agressieve strategie die iemand gebruikt wanneer een directe confrontatie als te gevaarlijk of kostbaar wordt ingeschat (Archer & Coyne, 2005); men tracht het slachtoffer schade te berokkenen zonder zelf rechtstreeks betrokken te zijn, bijvoorbeeld door roddel en laster (indirect-verbaal), of door brandstichting en vandalisme (indirect-fysiek).

Gedragswetenschappers buigen zich al jaren over de vraag waarom mensen anderen pijn doen of anderszins schade toebrengen. Er bestaan diverse theorieën over het ontstaan van agressief gedrag. Het General Aggression Model (GAM; Anderson & Bushman, 2002) is een integratief model, waarin de belangrijkste theoretische benaderingen zijn samengevoegd. In dit model bepalen persoonlijke variabelen (zoals temperament en andere persoonlijkheidskenmerken) in combinatie met situationele variabelen op een bepaald moment de gemoedstoestand van de persoon. De gemoedstoestand bestaat uit cognities, affect en *arousal*, die met elkaar in verbinding staan en elkaar kunnen activeren. De gemoedstoestand bepaalt uiteindelijk of iemand wel of niet agressief wordt.

Met behulp van diagnostiek kan in kaart worden gebracht welke factoren vermoedelijk bijdragen aan agressief/gewelddadig gedrag en hoe behandeling kan bijdragen aan het voorkómen van recidive. Het 'meten' van agressieve gedachten, gevoelens en gedrag is daarbij een belangrijk middel om de stand van zaken en de voortgang van de behandeling te 'monitoren'.

Binnen de forensische psychiatrie worden grofweg twee verschillende methoden gebruikt om agressieve gevoelens, gedachten en gedrag bij individuele patiënten in kaart te brengen, namelijk risicotaxatie door een professional en zelfrapportage door de patiënt. Tegenwoordig is het vanuit het Ministerie van Justitie verplicht om jaarlijks voor alle forensisch-psychiatrische

patiënten een zogenaamd ‘gestructureerd klinische beoordeling’ van het gevaar voor gewelddadige recidive te geven. Hierbij scoort de behandelaar een risicotaxatie-instrument, dat bestaat uit statische en dynamische items die met het recidiverisico samenhangen. Op grond van een afweging van de scores voor de verschillende items beoordeelt de behandelaar vervolgens zelf in hoeverre de patiënt een lage, matige of hoge kans heeft op een gewelddadig recidive in het komende jaar. Voorbeelden van deze risicotaxatie-instrumenten zijn de Historical, Clinical and Risk Management-20 (HCR-20; Webster, Eaves, Douglas, & Wintrup, 1995) en de gereviseerde versie HCR-20 V3 (de Vogel, de Vries Robbé, Bouman, Chakhssi, & de Ruiters, 2013), alsook de Historische, Klinische en Toekomstige Risico-indicatoren-30 (HKT-30; Werkgroep Risicotaxatie Forensische Psychiatrie, 2003) en de gereviseerde versie HKT-R (Spreen, Brand, ter Horst, Willems, & Bogaerts, 2013).

Naast de verplichte risicotaxatie-instrumenten worden zelfrapportagevragenlijsten gebruikt om het agressieniveau te meten en toekomstig agressief gedrag te voorspellen. Bij deze vragenlijsten geeft de patiënt zelf aan in hoeverre hij agressieve gedachten en gevoelens ervaart en zich de afgelopen tijd agressief heeft gedragen. Veelgebruikte instrumenten in dit kader zijn de Buss Durkee Hostility Inventory (BDHI; Buss & Durkee, 1957; Nederlandse versie, de BDHI-D, door Lischewski & de Ruiters, 1989), de Aggression Questionnaire (AQ; Buss & Perry, 1992; Nederlandse versie, de Aggressie Vragenlijst (AVL) door Meesters, Muris, Bosma, Schouten, & Beuving, 1996), en de Novaco Anger Scale-Provocation Inventory (NAS-PI; Novaco, 1994; Nederlandse versie door Hornsveld, Muris, & Kraaimaat, 2011). Dit zijn multidimensionele maten van agressie, speciaal ontwikkeld om verschillende vormen van agressieve cognities (hostiliteit), affect (boosheid, woede) en/of gedrag te meten. Daarnaast worden ook wel eens relevante subschalen uit andere instrumenten gebruikt, zoals de Hostiliteitsschaal van de Symptom Checklist-90 (SCL-90; Derogatis, Lipman, & Cobi, 1973; Nederlandse versie door Arrindell & Ettema, 1986; 2003) en de schalen voor Zelfgenoegzaamheid en Verongelijkheid van de Nederlandse Persoonlijkheden-Vragenlijst (NPV-2; Barelds, Luteijn, van Dijk, & Starren, 2007).

Uit onderzoek van Blok, de Beurs, de Ranitz en Rinne (2010) blijkt dat de combinatie van het gestructureerd klinische oordeel op grond van risicotaxatie-instrumenten en de zelfrapportage van de patiënt de beste voorspelling geeft van gewelddadige recidive (Blok et al., 2010; Expertisecentrum Forensische Psychiatrie (EFP), 2013; Hornsveld, Muris, & Kraaimaat, 2009a).

Het meten van de mate van woede en agressie met behulp van zelfrapportagevragenlijsten heeft een aantal voordelen. Dergelijke vragenlijsten zijn over het algemeen praktische instrumenten, in die zin dat ze gemakkelijk, snel en op gestandaardiseerde wijze informatie opleveren. Door de standaardisatie is vergelijking over de tijd mogelijk. Standaardisatie maakt ook een vergelijking tussen personen mogelijk, bijvoorbeeld in de vergelijking van de scores van de patiënt met die van personen uit niet-klinische en

niet-criminele populaties, wat een beeld geeft van de huidige toestand van de patiënt ten opzichte van gewenst toekomstig functioneren. Daarom maken deze vragenlijsten veelal deel uit van het instrumentarium voor de ROM (*Routine Outcome Monitoring*), waarbij instrumenten jaarlijks worden afgenomen, om zo het verloop van de behandeling te evalueren. Hierbij wordt zowel op individueel niveau als op afdelingsniveau bekeken in hoeverre er sprake is van vooruitgang. Daarnaast kunnen zelfrapportagevragenlijsten worden benut om een uitspraak te doen over de mogelijke kans op agressieve incidenten binnen de muren van de kliniek (de Vogel, de Ruiters, Hildebrand, Bos, & van de Ven, 2004). Wanneer het risico op agressief gedrag hoog wordt ingeschat, zouden de mogelijkheden en vrijheden voor de patiënt moeten worden beperkt, en dient het betrokken personeel extra alert te zijn. Discrepancies tussen de inschatting door de behandelaar of diagnosticus enerzijds en die door de patiënt anderzijds kunnen ten slotte ook aanknopingspunten voor de behandeling geven, omdat ze kunnen wijzen op verschil van inzicht over de aan te pakken problematiek en, daarmee samenhangend, gebrek aan motivatie voor de behandeling bij de patiënt (EFP, 2013).

Een nadeel van de zelfrapportagevragenlijsten is dat men is aangewezen op het oordeel van de patiënt zelf en dus op diens inzicht in het eigen functioneren. Dit inzicht is bij de gemiddelde tbs-patiënt nog wel eens problematisch, wat kan leiden tot een vertekend, niet erg valide beeld (EFP, 2013; Hornsveld et al., 2009a). Evenzeer moet rekening worden gehouden met vertekening als gevolg van de neiging tot sociaal wenselijk antwoorden: patiënten kunnen antwoorden geven die, naar hun verwachting, 'goed' of 'passend' zijn, zodat ze zichzelf op een sociaal acceptabele manier neerzetten (EFP, 2013; Hornsveld et al., 2009a; King & Bruner, 2000) en daarmee bijvoorbeeld proberen meer vrijheden te krijgen, eerder op verlof te mogen en verlenging van tbs-maatregel te vermijden.

#### BETROUWBAARHEID EN VALIDITEIT

.....

Aangezien zelfrapportagevragenlijsten veelvuldig worden gebruikt in de Nederlandse forensische psychiatrie, is het belangrijk te weten in hoeverre de informatie die dergelijke instrumenten opleveren, betrouwbaar en valide is.

Betrouwbaarheid geeft aan in hoeverre de met een test gemeten waarde representatief is voor de werkelijke waarde die de test wil meten. Voor zelfrapportagevragenlijsten zijn vooral interne consistentie en test-hertestbetrouwbaarheid van belang. Interne consistentie geeft aan in hoeverre verschillende items van een meetinstrument die eenzelfde kenmerk beogen te meten ook daadwerkelijk met elkaar samenhangen. Dit wordt uitgedrukt met Cronbach's alfa. Hierbij zijn waarden tussen de 0,70 en 0,79 redelijk, tussen 0,80 en 0,89 goed en groter dan 0,90 uitmuntend (Cicchetti, 1994). Bij de

test-hertestbetrouwbaarheid wordt een test na een bepaald tijdsinterval bij dezelfde onderzoeksgroep herhaald. De correlatie – meestal Pearsons product-momentcorrelatiecoëfficiënt – geeft aan in hoeverre deze scores met elkaar overeenkomen.

Verschillende soorten validiteit kunnen worden onderscheiden. Begripsvaliditeit geeft aan of de test inderdaad de eigenschap meet die wordt verondersteld. Zij wordt afgeleid uit de convergente en divergente validiteit. Convergente validiteit heeft betrekking op de vraag in hoeverre verschillende instrumenten die hetzelfde begrip beogen te meten met elkaar samenhangen. Er is sprake van divergente validiteit wanneer twee instrumenten die verschillende begrippen beogen te meten, laag met elkaar correleren. Criteriumvaliditeit geeft aan in hoeverre de testscore een goede voorspeller is van het gedrag (het criterium) dat de test beoogt te meten. Wanneer de testscores en de criteriumgegevens op hetzelfde moment zijn verzameld, geeft de samenhang een indicatie van de concurrente validiteit. Wanneer de testscores en de criterium/uitkomstgegevens in een longitudinaal *design* zijn verzameld, waarbij de testscores voorafgaan aan de criteriumgegevens, kan de mate van predictieve validiteit, oftewel de voorspellende waarde van de test, vastgesteld worden.

#### DOELSTELLING VAN DIT ARTIKEL

.....

Dit artikel biedt een overzicht van de literatuur met betrekking tot de vragen (1) hoe betrouwbaar en (2) valide de meest gebruikte agressievragenlijsten in een forensisch-psychiatrische populatie zijn, en (3) of de predictieve validiteit voldoende is om met deze agressievragenlijsten uitspraken te doen over mogelijk toekomstig agressief gedrag bij forensisch-psychiatrische patiënten. We bespreken in dit overzicht alleen de bevindingen in Nederlandse onderzoeksgroepen.

De literatuur met betrekking tot de betrouwbaarheid en validiteit van de vragenlijsten is geselecteerd met de databases PsychINFO, Medline en Pubmed, waarbij de volgende zoektermen zijn gebruikt in zowel het Nederlands als het Engels: ‘agressie’ AND: ‘screening’, ‘beoordeling’, ‘meting’, ‘meten’, ‘zelfrapportage’, ‘psychometrie’, ‘forensische psychiatrie’, ‘tbs’, ‘forensisch psychiatrische kliniek’, ‘betrouwbaarheid’ en ‘validiteit’. Daarnaast werden de namen van de zelfrapportagevragenlijsten gebruikt als zoekterm, zowel in het Nederlands als Engels, in combinatie met bovenstaande zoektermen. Het gaat hierbij om de volgende vragenlijsten: de BDHI (Buss & Durkee, 1957), AVL (Buss & Perry, 1992), NAS-PI (Novaco, 1994), SCL-90 (Derogatis et al., 1973) en de NPV-2 (Barelds et al., 2007). In eerste instantie is gezocht naar psychometrisch onderzoek van de vragenlijsten onder forensisch-psychiatrische patiënten. Indien hier geen resultaten werden gevonden, werd gekeken naar onderzoek onder gedetineerden en psychiatrische patiënten.

## OVERZICHT

De psychometrische kwaliteiten van de BDHI-D, NPV-2 en SCL-90 werden door de COTAN (Evers, Braak, Frima, & van Vliet-Mulder, 2009-2014) beoordeeld als voldoende tot goed (zie tabel 1). Opgemerkt moet worden dat een goede COTAN-beoordeling niet betekent (1) dat het instrument daarmee ook geschikt is voor gebruik in een forensisch-psychiatrische setting, (2) dat de COTAN-beoordelingen voor de NPV-2 en SCL-90 betrekking hebben op de gehele vragenlijst en niet op de agressiegerelateerde subschalen, en (3) dat beoordelingen voor de AVL en NAS-PI vooralsnog ontbreken.

TABEL 1 *Samenvatting COTAN-beoordelingen van de geëvalueerde tests (Egberink, Vermeulen, & Frima, 2009-2014).*

Test	Theoretische achtergrond	Kwaliteit testmateriaal	Kwaliteit handleiding	Normen	Betrouwbaarheid	Begripsvaliditeit	Criteriumvaliditeit	Jaar laatste beoordeling
AVL*	-	-	-	-	-	-	-	-
BDHI-D	G	G	G	G <sup>1</sup>	G <sup>2</sup>	G	V	1997
NAS-PI*	-	-	-	-	-	-	-	-
NPV-2+	G	G	G	V <sup>3</sup>	G <sup>4</sup>	G	V	2009
SCL-90+	G	G	G	V	G <sup>5</sup>	G	G	2004

\* Deze tests zijn (nog) niet beoordeeld door de COTAN.

+ De COTAN-beoordelingen betreffen de gehele vragenlijst en niet de agressiegerelateerde subschalen ervan.

## Opmerkingen:

- 1 De normen zijn verouderd.
- 2 De betrouwbaarheid van de BDHI-subschaal Sociale Wenselijkheid is onvoldoende.
- 3 Voor bepaalde groepen 'Onvoldoende' wegens het ontbreken van relevante gegevens.
- 4 'Goed' voor minder belangrijke beslissingen, 'Voldoende' voor belangrijke beslissingen op individueel niveau.
- 5 De betrouwbaarheid van de subschaal voor hostiliteit is 'voldoende'.

In de volgende paragrafen worden per vragenlijst kenmerken zoals meetprentie, aantal items, *response format* en subschalen besproken. Een samen-

vatting hiervan is te vinden in tabel 2. Gegevens over de betrouwbaarheid en de validiteit van de instrumenten worden samengevat in tabel 3.

288

TABEL 2 *Kenmerken van zelfrapportagevragenlijsten voor het meten van agressie (naam, normen, responsformaat, aantal items en subschalen)*

Naam	Normen	Antwoordformaat	Items	Subschalen (aantal items) en Totaalscore
Agressie Vragenlijst (AVL)	Normen ontbreken	5-puntsschaal, van 'helemaal mee oneens' tot 'helemaal mee eens'	29	Fysieke agressie (9) Verbale agressie (5) Woede (7) Vijandigheid (8) Totaalscore = mate van agressie
Buss-Durkee Hostility Inventory-Dutch (BDHI-D)	Algemene bevolking; gedetineerden; psychiatrische patiënten	2-puntsschaal, met 'waar' en 'onwaar'	40	Directe agressie (19) Indirecte agressie (16) Sociale Wenselijkheid (5) Totaalscore = mate van hostiliteit en dispositie voor agressief gedrag
Novaco Anger Scale Provocation Inventory (NAS-PI)	Terbeschikking-gestelden; ambulante psychiatrische patiënten; langgestraften; vmbo-leerlingen	3-puntsschaal, met 'nooit waar', 'soms waar' en 'altijd waar'	48	Cognities (16) Arousal (16) Gedrag (16) NAS = Woededispositie PI = Woedereactiviteit
Symptom Checklist (SCL-90)	Algemene bevolking; ambulante psychiatrische patiënten	5-puntsschaal voor mate van woede van 'helemaal niet' tot 'heel veel'	90	Hostiliteit (6)
Nederlandse Persoonlijkheds Vragenlijst-2 (NPV-2)	Algemene bevolking; psychiatrische en somatische patiënten	3-puntsschaal met 'juist', '?' en 'onjuist'	140	Verongelijkheid (20) Zelfgenoegzaamheid (20)



TABEL 3 *Overzicht van de betrouwbaarheidsgegevens van zelfrapportagevragenlijsten voor het meten van agressie*

Test		Interne consistentie <i>Cronbach's alfa</i>	Test-heretest <i>Correlatie coëfficiënt</i>
AVL	Hornsveld et al., 2009b Meesters et al., 1996 Morren & Meesters, 2002	0,83 <sup>1</sup> (0,34-0,83) <sup>2</sup> 0,84 <sup>1</sup> (0,50-0,84) <sup>2</sup> 0,86 <sup>1</sup> (0,51-0,86) <sup>2</sup>	0,72 <sup>1</sup> (0,54-0,76) <sup>2</sup> 0,80 <sup>1</sup> (0,71-0,80) <sup>2</sup> -
BDHI-D	Lange et al., 1995c Lange et al., 1995b	(0,77 – 0,79) <sup>2</sup> (0,50 – 0,83) <sup>2</sup>	-(0,64 – 0,90) <sup>2</sup>
NAS-PI	Hornsveld et al., 2009a  Hornsveld et al., 2011	NAS: 0,93 <sup>1</sup> (0,77 – 0,93) <sup>2</sup> PI: 0,90 NAS-PI: 0,89 (0,70 – 0,95) <sup>2</sup>	NAS: 0,80 <sup>1</sup> (0,71 – 0,79) <sup>2</sup> PI: - NAS-PI: 0,84  (0,71 – 0,80) <sup>2</sup>
SCL-90 HOS	Arrindell & Ettema, 2003 Hafkenscheid, 1993	0,84 0,76	≥ 0,62 -
NPV: VE & ZE <sup>3</sup>	Luteijn et al., 2000 Kloens et al., 2002 Barelds et al., 2007	NPV NPV NPV-2	VE: 0,87 ZE: 0,76 - - VE: 0,91 ZE: 0,94

<sup>1</sup> Totaalscore<sup>2</sup> Hoogste en laagste waarde van de afzonderlijke subschalen<sup>3</sup> VE = Verongelijkheid; ZE = Zelfgenoegzaamheid

## KENMERKEN VAN DE VERSCHILLENDE VRAGENLIJSTEN

.....

*Buss Durkee Hostility Inventory (BDHI)*

.....

De BDHI (Buss & Durkee, 1957; Nederlandse versie, de BDHI-D, door Lischewski & de Ruiters, 1989) is gebaseerd op een multidimensionele theorie over vijandig-agressief gedrag; het instrument evalueert niet alleen de mate van agressiviteit maar ook verschillende niveaus van agressief gedrag. De BDHI meet gedachten en gevoelens die geassocieerd zijn met agressief gedrag, en bestaat uit 40 items die dichotoom worden beantwoord met 'waar'

of 'onwaar'. De items meten drie concepten, namelijk Directe Agressie, Indirecte Agressie en Sociale Wenselijkheid. De subschaal Directe Agressie bestaat uit negentien items met betrekking verbale en fysieke agressie, zoals schelden en slaan (voorbeeld: 'Ik heb mensen gekend die mij zo ver wisten te krijgen dat we op de vuist gingen'). Indirecte Agressie bestaat uit zestien items met betrekking tot indirecte, emotionele en cognitieve agressie, zoals vijandigheid, prikkelbaarheid, wantrouwen en woede (voorbeeld: 'Soms heb ik het gevoel dat mensen me uitlachen'). Sociale Wenselijkheid bestaat uit vijf items met betrekking tot de neiging tot sociaal wenselijke vertekening (voorbeeld: 'Het is voorgekomen dat ik iemand benadeelde'). De *a priori* schaalindeling zoals boven beschreven werd in factoranalytisch onderzoek overigens niet altijd bevestigd (Lange et al., 1995c; Lischewski & de Ruiter, 1989).

Er zijn Nederlandse normen beschikbaar voor de algemene bevolking (Kodde, Rullmann, Spaargaren, & Waltman, 1994), gedetineerden en psychiatrische patiënten (Lange, Hoogedoorn, Wiederspahn, & de Beurs, 1995b). Normen voor forensisch-psychiatrische patiënten ontbreken.

**Betrouwbaarheid** — Uitgaande van de oorspronkelijke schaalindeling toonde de BDHI-D in onderzoek onder de algemene bevolking ( $n = 472$ ), psychiatrische patiënten ( $n = 426$ ) en gedetineerden ( $n = 401$ ) een voldoende tot goede interne consistentie voor de subschalen Directe en Indirecte Agressie ( $\alpha \geq .79$ ); voor Sociale Wenselijkheidschaal was de interne consistentie minder sterk ( $\alpha \geq .50$ ) (Lange et al., 1995b). De test-hertestbetrouwbaarheid over een interval van zes weken was relatief hoog voor de subschalen Directe en Indirecte Agressie (respectievelijk  $r = .80$  en  $r = .90$ ); de schaal voor Sociale Wenselijkheid bleek minder stabiel ( $r = .64$ ) (Lange et al., 1995b). De matige interne consistentie en lagere test-hertestcoëfficiënt voor Sociale Wenselijkheid zou volgens de onderzoekers te wijten kunnen zijn aan het relatief kleine aantal items.

**Validiteit** — De BDHI-D-schalen voor Directe en Indirecte Agressie correleerden voldoende met enkele analoge instrumenten, wat steun voor de convergerende validiteit betekent. Zo bleek uit onderzoek van Lange et al. (1995b) onder de algemene bevolking, psychiatrische patiënten en gedetineerden dat Directe Agressie significant correleerde met de schaal Vijandigheid ( $r = .64$ ) uit de Vragenlijst Interpersoonlijke Relaties (Vertommen & Rochette, 1975), terwijl Indirecte Agressie significant samenhang met de subschaal Verbitterdheid ( $r = .54$ ). Lange, Dehghani, en de Beurs (1995a) vonden onder 67 jonge delinquenten sterke correlaties van zowel Directe Agressie ( $r = .56$ ) als Indirecte Agressie ( $r = .57$ ) met de Woededispositieschaal van de Zelf Analyse Vragenlijst (ZAV; van der Ploeg, Defares, & Spielberger, 1982). BDHI-D-scores lieten in diezelfde onderzoeksgroep ook een verband met gedragsevaluaties door onafhankelijke beoordelaars zien.

Lange et al. (1995a) vonden in hun onderzoek sterke correlaties van Directe en Indirecte Agressie met gedragsevaluaties door behandelaars en groepsbegeleiders (respectievelijk  $r = .61$  en  $r = .58$ ), waardoor zij van een goede concurrente validiteit spreken. We hebben echter geen onderzoek gevonden naar de predictieve validiteit van de BDHI-D in forensische populaties, zodat vooralsnog niet bekend is in hoeverre scores op dit instrument toekomstig agressief gedrag kunnen voorspellen.

### *Agressie Vragenlijst (AVL)*

De AVL (Meesters et al., 1996) is de Nederlandse versie van de Aggression Questionnaire (AQ; Buss & Perry, 1992). De AQ is ontwikkeld als verbeterde versie van de Buss-Durkee Hostility Inventory (BDHI; Buss & Durkee, 1957), aangezien de BDHI vrij lang is (75 items) en de factorstructuur niet helemaal stabiel lijkt.

De AVL meet zowel directe als indirecte agressie, waarbij de totaalscore een indicatie is voor de mate van agressie. Het instrument bestaat uit 29 uitspraken. De respondent geeft op een vijfpuntsschaal met de ankers 'helemaal mee oneens' (1) tot 'helemaal mee eens' (5) aan in hoeverre de uitspraken op hem van toepassing zijn. Factoranalyse van de AQ leverde vier factoren op (Buss & Perry, 1992): Fysieke Agressie (negen items; voorbeeld: 'Als iemand me slaat, sla ik terug'), Verbale Agressie (vijf items; voorbeeld: 'Als mensen me ergeren, zeg ik soms wel wat ik van ze vind'), Woede (zeven items; voorbeeld: 'Wanneer ik gefrustreerd ben, laat ik mijn irritaties merken') en Vijandigheid (acht items; voorbeeld: 'Wanneer personen erg aardig zijn, vraag ik me af wat ze van me willen'). Deze a priori structuur werd in factoranalytisch onderzoek niet altijd bevestigd (zie bijvoorbeeld Williams, Boyd, Cascardi, & Poythress, 1996) en ook de Nederlandse versie van deze lijst lijkt een niet geheel stabiele factorstructuur te hebben (Hornsveld, Muris, Kraaimaat, & Meesters, 2009b). Voor zover wij konden nagaan, ontbreken normen voor dit instrument.

**Betrouwbaarheid** — De interne consistentie is onderzocht onder studenten ( $n = 762$ ), klinische forensische patiënten ( $n = 138$ ), ambulante forensische patiënten ( $n = 206$ ), vmbo-leerlingen ( $n = 160$ ) en delinquente adolescenten ( $n = 73$ ), en bleek goed te zijn voor de totaalscore van de AVL ( $\alpha \geq .83$ ) (Hornsveld et al., 2009b; Meesters et al., 1996; Morren & Meesters, 2002). Voor de subschalen Fysieke Agressie, Woede en Vijandigheid werd een matige tot goede interne consistentie gevonden ( $.57 \leq \alpha \leq .81$ ); de subschaal Verbale Agressie bleek echter onvoldoende intern consistent ( $.34 \leq \alpha \leq .51$ ) (Hornsveld et al., 2009b; Meesters et al., 1996; Morren & Meesters, 2002). De AVL bleef redelijk stabiel onder de ambulante forensische patiënten over een interval van vier weken ( $r = .72$ ; Hornsveld et al., 2009b) en onder studenten over een interval van zes weken ( $r = .80$ ; Meesters et al., 1996).

**Validiteit** — De convergente validiteit van de AVL werd ondersteund onder 138 klinische en 206 ambulante forensische patiënten, en 160 vmbo-leerlingen, met positieve correlaties ( $r \geq .45$ ) van de AVL-totaalscore met schalen die analoge constructen beogen te meten, namelijk Woede als dispositie gemeten met de Zelf Analyse Vragenlijst (ZAV; van der Ploeg et al., 1982) en Woede als dispositie gemeten met de Novaco Anger Scale (NAS; Novaco, 1994) (Hornsveld et al., 2009a).

De begripsvaliditeit is onderzocht door Morren en Meesters (2002), die het gedrag van 73 delinquente adolescenten lieten beoordelen door leidinggevend personeel op de mate van fysiek geweld, verbale agressie, woede en negatieve attitude. Hieruit bleek dat de AVL-totaalscore de woedebeoordelingen redelijk voorspelde ( $r = .42$ ), maar niet de beoordelingen van fysieke en verbale agressie en negatieve attitude; dit duidt op een beperkte begripsvaliditeit. Er zijn geen studies gevonden naar de predictieve validiteit in een (forensisch-)klinische setting.

#### *Novaco Anger Scale-Provocation Inventory (NAS-PI)*

De NAS-PI (Novaco, 1994; Nederlandse versie door Hornsveld et al., 2011) is ontworpen als maat voor woededispositie met de nadruk op hoe een individu woede ervaart en hoe kwaad hij zich voelt in bepaalde situaties. Volgens Novaco (2003) is woede een normale emotie, die belangrijke adaptieve functies voor het individu heeft, zoals overlevings- en copingmechanismen, maar ook negatieve gevolgen kan hebben, zoals agressief gedrag en stressgerelateerde gezondheidsproblemen. De vragenlijst bestaat uit twee delen. Het eerste deel, de Novaco Anger Scale (Novaco AS), bestaat uit 48 items verdeeld over drie subschalen, namelijk Cognitie, Arousal en Gedrag. Bij Cognitie staat het interpreteren van sociale informatie, sociale kennis en attitudes centraal; personen die op deze schaal hoog scoren ervaren eerder dreiging of onrechtvaardigheid door anderen. De Cognitieschaal bestaat uit zestien items (voorbeeld: 'Iedere week ontmoet ik wel iemand die ik niet mag'). Arousal heeft betrekking op de mate van opwindning; personen die op deze schaal hoog scoren ervaren een toestand van hoge psychofysiologische opwindning, zoals versnelde hartslag. De Arousal-schaal bestaat uit zestien items (voorbeeld: 'Als ik boos word, duurt het uren voor ik eroverheen ben'). Gedrag betreft het aangeleerde reactiepatroon op een provocatie of dreiging; personen die op deze schaal hoog scoren beschikken over beperkte zelfcontrole tijdens een provocatie of dreiging. De Gedragschaal bestaat uit zestien items (voorbeeld: 'Ik ben gauw aangebrand en opvliegend van aard'). Respondenten geven op een driepuntsschaal aan in hoeverre de uitspraken van toepassing zijn, van 'nooit waar' (1) tot 'altijd waar' (3). De drie domeinen tezamen vormen de totaalscore die een indicatie geeft voor de mate waarin er sprake is van woededispositie.

Het tweede deel, de Novaco Provocation Inventory (Novaco PI), betreft de intensiteit van woede over verschillende situaties waarin men geprovoceerd kan worden. Het instrument bestaat uit 25 items waarin de persoon aangeeft welke specifieke situaties en provocaties hem woedend maken. Het gaat hierbij om vijf verschillende soorten van provocatieve situaties, namelijk respectloze behandeling, frustratie, oneerlijkheid, irritaties en irritante trekjes van anderen. Respondenten geven op een vierpuntsschaal aan in hoeverre de situatie hen boos maakt, van 'helemaal niet kwaad' (1) tot 'erg kwaad' (4). Dit deel meet de woedereactiviteit (Novaco & Taylor, 2004).

Er zijn normen voor de NAS-PI beschikbaar voor forensisch-psychiatrische patiënten, langgestraften en vmbo-leerlingen (Hornsveld et al., 2009a).

**Betrouwbaarheid** — Voor dit instrument werd in een forensische onderzoeksgroep, bestaande uit zowel klinische ( $n = 142$ ) als poliklinische patiënten ( $n = 194$ ), een goede tot zeer goede interne consistentie gevonden voor de Novaco AS-totaalscore ( $\alpha \geq .93$ ), de Novaco AS-subschalen ( $.77 \leq \alpha \leq .90$ ) en de Novaco PI-totaalscore ( $\alpha \geq .90$ ) (Hornsveld et al., 2009a; Hornsveld et al., 2011). Hornsveld et al. (2011) vonden bij 194 ambulante forensische patiënten redelijk stabiele scores over een interval van vier weken ( $.71 \leq r \leq .80$ ).

**Validiteit** — De convergente validiteit voor zowel de Novaco AS als de Novaco PI is onderzocht onder 142 klinische forensische patiënten, 194 poliklinische forensische patiënten en 320 vmbo-leerlingen. In alle onderzoeksgroepen correleerden de scores van de Novaco AS en de Novaco PI positief en significant ( $r > .24$ ) met andere maten voor agressiegerelateerde concepten, namelijk Woede als dispositie uit de ZAV (van der Ploeg et al., 1982), Vijandigheid uit de Aangepaste versie van de Picture Frustration Study (PFS-AV; Hornsveld, Nijman, Hollin, & Kraaimaat, 2007) en de subschalen Agressie en Woede uit de Aangepaste versie van de Agressie Vragenlijst (AVL-AV; Hornsveld et al., 2009b). Er werd geen onderzoek gevonden naar de predictieve validiteit van de Novaco AS en Novaco PI.

#### *Symptom Checklist-90 (SCL-90): Hostiliteit*

De SCL-90 (Derogatis et al., 1973; Nederlandse versie door Arrindell & Ettema, 1986, 2003) is een multidimensionele zelfbeoordelingsklachtenlijst waarvan de subschaal 'Hostiliteit' gedachten, gevoelens of gedragingen weergeeft die kenmerkend zijn voor de negatieve gemoedstoestand van woede en die verder wordt gekenmerkt door aspecten van agressie, irritatie en ressentiment (Arrindell & Ettema, 2003). De SCL-hostiliteitschaal bestaat uit zes items die woedegerelateerde klachten beschrijven waarvan de respondent op een vijfpuntsschaal van 'helemaal niet' (0) tot 'heel veel' (5) aangeeft in hoeverre de klacht de afgelopen week, inclusief vandaag, op hem van toepassing was (voorbeeld: 'Schreeuwen of met dingen smijten'). Door

de korte referentieperiode (namelijk alleen de afgelopen week) lijkt het instrument erg dynamisch en geschikt om ingezet te worden om op de korte termijn veranderingen in woede te evalueren. Er zijn normen beschikbaar voor volwassenen en ambulante psychiatrische patiënten (Arrindell & Ettema, 1986, 2003).

**Betrouwbaarheid** — De SCL-hostiliteitsschaal had een voldoende tot goede interne consistentie ( $.75 \leq \alpha \leq .84$ ) onder 437 klinisch psychiatrische patiënten (Hafkenscheid, 1993), 525 patiënten uit een eerstelijns psychologenpraktijk (Kloens, Barelds, Luteijn, & Schaap, 2002) en 2118 poliklinisch psychiatrische patiënten (Arrindell & Ettema, 2003). In deze laatste onderzoeksgroep bleek de schaal over een periode van twee maanden verrassend stabiel, met test-hertestcoëfficiënten van  $.73 \leq r \leq .97$ .

**Validiteit** — Divergente validiteit blijkt uit de verwaarloosbaar lage relaties van de SCL-hostiliteitsschaal met de subschalen Rigiditeit, Zelfgenoegzaamheid en Dominantie van de Nederlandse Persoonlijkheden-Vragenlijst (NPV-2; Barelds et al., 2007) onder 525 patiënten uit een eerstelijns psychologenpraktijk. Er is geen literatuur gevonden over de convergente validiteit, noch over de predictieve validiteit van de SCL-hostiliteitsschaal.

*Nederlandse Persoonlijkheden-Vragenlijst (NVP): Verongelijkheid en Zelfgenoegzaamheid*

.....

De NPV-2 (Barelds et al., 2007), een volledig herziene versie van de NPV (Luteijn, Starren, & van Dijk, 1975, 2000), meet zeven hoofddimensies van persoonlijkheid. In de forensische psychiatrie wordt voor het inschatten van de kans op agressief gedrag gebruikgemaakt van de subschalen Verongelijkheid en Zelfgenoegzaamheid. Bij Verongelijkheid staat kritiek op en wantrouwen tegenover andere mensen centraal; personen die op deze schaal hoog scoren zijn vaak intolerant, opvliegend en ongeduldig. De verongelijkheidschaal bestaat uit twintig items (voorbeeld: 'Mensen roddelen vaak over mij'). Zelfgenoegzaamheid meet de mate van tevredenheid met zichzelf en tegelijkertijd de desinteresse voor een ander en diens problemen; personen die op deze schaal hoog scoren zijn vaak egoïstisch en hebben weinig interesse in anderen, wat kan leiden tot agressief gedrag. De zelfgenoegzaamheidschaal bestaat eveneens uit twintig items (voorbeeld: 'Mijn vrije tijd breng ik het liefst alleen door'). Bij het invullen van de twee schalen geeft de respondent aan of de uitspraken 'juist', 'onbekend' of 'onjuist' zijn. Er zijn normen beschikbaar voor de algemene bevolking, somatische en psychiatrische patiënten en eerstelijns ggz-cliënten (Luteijn et al., 2000).

**Betrouwbaarheid** — De interne consistentie bleek bij verschillende onderzoeksgroepen, waaronder eerstelijns ggz-cliënten ( $n = 92$ ) en een groep

van klinisch en poliklinisch psychiatrische patiënten ( $n = 227$ ), goed voor de subschaal Verongelijkheid ( $\alpha = .87$ ) en voldoende voor de subschaal Zelfgenoegzaamheid ( $\alpha = 0.76$ ) (Luteijn et al., 2000). Als index van persoonlijkheidskenmerken worden de scores verondersteld relatief stabiel te zijn. Dit wordt bevestigd door de hoge test-hertestcoëfficiënten die zijn gevonden in een onderzoek onder een groep van 57 personen over een tijdsinterval van gemiddeld achttien dagen ( $r = .94$  voor Verongelijkheid en  $r = .91$  voor Zelfgenoegzaamheid) (Barelds et al., 2007). De stabiliteit van Verongelijkheid van de NPV bleek bij grotere tijdsintervallen tamelijk hoog onder 39 cliënten met psychische klachten bij een interval van ongeveer 24 maanden ( $r = .73$ ) en onder 96 reumapatiënten met een tussentijd van ongeveer 28 maanden ( $r = .77$ ) (Ouborg, 1988; Ploeg, 1980). Bij deze zelfde onderzoeksgroepen liggen de waarden voor Zelfgenoegzaamheid lager dan Verongelijkheid ( $r = .54$  bij de cliënten met psychische klachten en  $r = .61$  bij de reumapatiënten).

**Validiteit** — Verongelijkheid en Zelfgenoegzaamheid hadden een positieve en gemiddelde correlatie met de hostiliteitschaal van de SCL-90 in een groep van 106 personen uit de algemene bevolking ( $r = .38$  voor Verongelijkheid en  $r = .42$  voor Zelfgenoegzaamheid) (Barelds et al., 2007). Verongelijkheid hing samen ( $r = .66$ ) met Negativisme en Ernstige Psychopathologie van de Negativisme, ernstige Psychopathologie en Somatisering Test (NPST; Luteijn & Barelds, 2003) in een kleine groep personen uit de algemene bevolking ( $n = 10$ ) (Barelds et al., 2007). In datzelfde onderzoek bleek sprake te zijn van een redelijke divergente validiteit van Verongelijkheid en Zelfgenoegzaamheid door geringe samenhang met de (overige) NPV-schalen. We hebben geen onderzoek gevonden naar de concurrente en predictieve validiteit in forensische populaties.

#### CONCLUSIE

.....

De verzamelterm ‘agressievragenlijsten’ heeft betrekking op instrumenten voor zelfgerapporteerde agressieve gedachten (hostiliteit), gevoelens (boosheid, woede) en gedrag. Voor het gebruik van deze agressievragenlijsten in een forensisch-psychiatrische setting brengt dit overzicht enkele plus- en minpunten naar voren.

Het globale concept ‘agressiegerelateerde problematiek’ blijkt door de verschillende instrumenten voldoende betrouwbaar te kunnen worden gemeten, waarbij de totaalscores voor de multidimensionele instrumenten (AVL, Novaco AS en PI, BDHI-D) naar men mag aannemen een bredere dekking van het concept geven dan de subschaalscores uit de NPV en SCL-90. Immers, de multidimensionele instrumenten meten expliciet verschillende typen en aspecten van agressie, terwijl de subschalen uit de NPV en SCL-90 zich beperken tot één enkel aspect.

De interne consistentie voor de totaalscore bleek uitstekend, en de bevindingen met betrekking tot de test-hertestbetrouwbaarheid geven aan dat de instrumenten over het algemeen stabiel zijn over de tijd, maar voldoende dynamisch om veranderingen in de behandeling te meten. Verschillende subschalen van met name de AVL bleken echter niet altijd even betrouwbaar. Samenhangend met de niet helemaal stabiele en eenduidige factorstructuur die voor dit instrument werd gerapporteerd, toonden de subschalen een matige interne consistentie. Dit betekent dat het voordeel dat een multidimensioneel instrument kan bieden, namelijk het differentiëren tussen verschillende aspecten van agressie, niet ten volle kan worden benut: op subschaalniveau lijkt de AVL (en in mindere mate de BDHI-D) minder betrouwbaar. De bevindingen met betrekking tot de Novaco AS en PI waren wat dit betreft gunstiger.

De begripsvaliditeit van de multidimensionele vragenlijsten lijkt te worden gesteund door zinvolle samenhang met analoge instrumenten en met gedragsevaluaties. Vertekening van de vragenlijstcores door de neiging tot sociaal wenselijk antwoorden kan echter niet worden uitgesloten. Alleen de BDHI beschikt over een schaal waarmee de neiging tot sociaal wenselijke vertekening kan worden geschat; de betrouwbaarheid van deze subschaal leek echter matig. Om rekening te kunnen houden met tendensen tot sociaal wenselijke antwoorden zou dus een additioneel instrument moeten worden meegenomen, bijvoorbeeld de Marlowe-Crowne Social Desirability scale van Crowne en Marlowe (1960). Een alternatief is de inzet van een impliciete maat voor agressie (Gawronski, 2009), zoals de Weapons Entertainment Implicit Association Test van Polaschek, Bell, Calvert, en Takarangi (2010) of de Implicit Association Test of Aggression van Banse en Fischer (2002) (Richetin, South Richardson, & Mason, 2010). Dergelijke impliciete maten zouden volgens Gawronski (2009) bovendien superieur zijn ten opzichte van zelfrapportagematen wanneer er sprake is van verminderde cognitieve capaciteit of een lage capaciteit in het werkgeheugen. Bij de populatie forensisch-psychiatrisch patiënten, van wie volgens het Nationaal Kompas Volksgezondheid (2010) 20% zwakbegaafd is (IQ < 80), kunnen impliciete maten voor het meten van agressie ook daarom een zinvolle toevoeging betekenen.

De vraag of de agressievragenlijsten voldoende predictieve validiteit hebben om uitspraken te doen over mogelijk toekomstig agressief gedrag bij forensisch-psychiatrische patiënten kan op grond van dit literatuuroverzicht niet worden beantwoord. We hebben geen empirisch onderzoek kunnen vinden waarin scores op de agressievragenlijsten worden gerelateerd aan op later tijdstip gemeten agressief gedrag. Dat is een pijnlijk gemis, omdat juist deze functie van de instrumenten van groot belang is voor beslissingen rond behandeling en beveiliging. Gepaard aan het gebrek aan normgegevens voor de forensisch-psychiatrische populatie betekent dit een forse beperking van de bruikbaarheid van agressievragenlijsten in een forensisch-psychiatrische setting.



Samenvattend wijzen de gegevens in dit literatuuroverzicht erop dat de betrouwbaarheid en begripsvaliditeit van agressievragenlijsten zeker geen beletsel vormen voor het gebruik van deze instrumenten in een forensisch-psychiatrische setting; empirisch onderzoek naar de predictieve validiteit en het verzamelen van relevante normgegevens kan de bruikbaarheid nog belangrijk verbeteren.

**GERJONNE AKKERMAN-BOUWSEMA** is onderzoeker bij de Divisie Beveiligde Psychiatrie, GGZ Drenthe, Assen. *Adres:* Denneweg 9, 9404 LA Assen.

**BRENDA BIESBROEK** is psycholoog bij INTER-PSY, Winschoten. *Adres:* Mr. A.E.J. Moddermanstraat 3, 9671 AG Winschoten.

**MARIKE LANCEL** is hoofdonderzoeker bij de Divisie Beveiligde Psychiatrie, GGZ Drenthe, Assen. *Adres:* Denneweg 9, 9404 LA Assen.

**COBY GERLSMA** is universitair docent klinische psychologie, Rijksuniversiteit Groningen. *Adres:* Grote Kruisstraat 2/1, 9712 TS Groningen.

*Correspondentieadres:* Gerjonne (G. J.) Akkerman-Bouwsema, MSc, GGZ Drenthe, Forensisch Psychiatrische Kliniek, Afdeling Onderzoek, Denneweg 9, 9404 LA Assen. E-mail: Gerjonne.Akkerman-Bouwsema@ggzdrenthe.nl

**Summary** Aggression is a common problem in forensic psychiatric settings and it is therefore important that aggression and violent behaviour of forensic psychiatric patients can be assessed and predicted accurately. In this literature review, self-report instruments for measuring aggression that are employed in the Netherlands are evaluated. We found that nearly all questionnaires have sufficient internal consistency and test-retest reliability. However, information concerning the validity of these instruments in adult (forensic) clinical populations was sparse. Moreover, for most instruments there were no norms for forensic psychiatric patients, nor was the influence of social desirability bias examined. It would be worthwhile to further investigate the predictive power of the aggression self-report questionnaires and to delineate norms for forensic psychiatric populations.

**Keywords:** *aggression, self report questionnaires, forensic psychiatric practice, psychometric characteristics, reliability and validity*

## Literatuur

- Anderson, C. A. & Bushman, B. J. (2002). Human aggression. *Annual Review of Psychology*, 53, 27-51.
- Archer, J. & Coyne, S. M. (2005). An integrated review of indirect, relational, and social aggression. *Personality and Social Psychology Review*, 9, 212-230.
- Arrindell, W. A. & Ettema, H. (1986). *Handleiding bij een multidimensionele psychopathologie-indicator*. Lisse: Swets & Zeitlinger.
- Arrindell, W. A. & Ettema, H. (2003). *Handleiding bij een multidimensionele psychopathologie-indicator*. Lisse: Swets & Zeitlinger.
- Banse, R. & Fischer, I. (2002). *Implicit and explicit aggressiveness and the prediction of aggressive behavior*.

- Poster presentation at the 11<sup>th</sup> Conference on Personality of the European Association of Personality Psychology, Jena.
- Barelds, D. P. H., Luteijn, F., van Dijk, H., & Starren, H. (2007). *Nederlandse Persoonlijkheids Vragenlijst 2 (NPV-2): Handleiding*. Amsterdam: Harcourt.
- Blok, G. T., de Beurs, E., de Ranitz, A. G. S., & Rinne, T. (2010). Psychometrische stand van zaken van risicotaxatie-instrumenten voor volwassenen in Nederland. *Tijdschrift voor Psychiatrie*, 52, 331-340.
- Buss, A. H. & Durkee, A. (1957). An inventory for assessing different kinds of hostility. *Journal of Abnormal Social Psychology*, 21, 343-349.
- Buss, A. H. & Perry, M. (1992). The Aggression Questionnaire. *Journal of Personality and Social Psychology*, 63, 452-459.
- Cicchetti, D. V., (1994). Guidelines, criteria, and rules of the thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 2, 290.
- Crowne, D. P. & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, 24, 349-354.
- de Ruiter, C. & Hildebrand, M. (1999). *Behandelingsstrategieën bij forensisch psychiatrische patiënten*. Houten: Bohn Stafleu van Loghum.
- de Vogel, V., de Ruiter, C., Hildebrand, M., Bos, B., & van de Ven, P. (2004). Type of discharge and risk of recidivism measured by the HCR-20: A retrospective study in a Dutch sample of treated forensic psychiatric patients. *International Journal of Forensic Mental Health*, 3, 149-165.
- de Vogel, V., de Vries Robbé, M., Bouman, Y. H. A., Chakhssi, F., & de Ruiter, C. (2013). Innovatie in risicotaxatie van geweld: De HCR-20V3. *Tijdschrift voor Gedragstherapie*, 46, 107-118.
- Derogatis, L. R., Lipman, R. S. & Cobi. L. (1973). SCL-90: An outpatient psychiatric rating Scale – Preliminary report. *Psychopharmacology Bulletin*, 9, 13-28.
- Egberink, I. J. L., Vermeulen, C. S. M., & Frima, R. M. (2009-2014). *COTAN Documentatie* (www.cotandocumentatie.nl). Amsterdam: Boom test uitgevers.
- Evers, A., Braak, M. S. L., Frima, R. M., & van Vliet-Mulder, J. C. (2009-2011). *COTAN Documentatie*. Amsterdam: Boom test uitgevers.
- Expertisecentrum Forensische Psychiatrie (EFP) (2013). *Basis Zorgprogramma*. Utrecht: EFP.
- Gawronski, B. (2009). Ten frequently asked questions about implicit measures and their frequently supposed, but not entirely correct answers. *Canadian Psychology*, 50, 141-150.
- Hafkenscheid, A. (1993). Psychometric evaluation of the Symptom Checklist (SCL-90) in psychiatric inpatients. *Personality and Individual Differences*, 14, 751-756.
- Hornsveld, R. H. J., Muris, P., & Kraaimaat, F. W. (2009a). *Drie zelfrapportagevragenlijsten voor de forensische psychiatrie*. Rhooen: FPC de Kijvelanden.
- Hornsveld, R. H. J., Muris, P. & Kraaimaat, F. W. (2011). The Novaco Anger Scale – Provocation Inventory (1994 version) in Dutch forensic psychiatric patients. *Psychological Assessment*, 23, 937-944.
- Hornsveld, R. H. J., Muris, P., Kraaimaat, F. W. & Meesters, C. (2009b). Psychometric properties of the Aggression Questionnaire in Dutch violent forensic psychiatric patients and secondary vocational students. *Psychological Assessment*, 16, 181-192.
- Hornsveld, R. H. J., Nijman, H. L. I., Hollin, C. R., & Kraaimaat, F. W. (2007). An adapted version of the Rosenzweig Picture-Frustration Study for the measurement of hostility in violent forensic psychiatric patients.

- Criminal Behaviour and Mental Health*, 17, 45-56.
- Hornsveld, R. H. J., Soe-Agnie, S., Donker, J. & van der Wal, L. (2008). De ontwikkeling van een behandelprogramma voor gewelddadige forensisch psychiatrische patiënten met een cluster B persoonlijkheidsstoornis. *Tijdschrift voor Psychiatrie*, 50, 703-711.
- King, M. F. & Bruner, G. C. (2000). Social desirability bias: A neglected aspect of validity testing. *Psychology and Marketing*, 17, 79-103.
- Kloens, G. J., Barelds, D. P. H., Luteijn, F. & Schaap, C. P. D. R. (2002). De waarde van enige vragenlijsten in de eerste lijn. *Diagnostiek-wijzer*, 5, 520-522.
- Kodde, H., Rullmann, T., Spaargaren, M., & Waltman, M. (1994). *Normering van de Buss-Durkee Hostility Inventory*. Amsterdam: Vakgroep Klinische Psychologie, UvA.
- Lange, A., Dehghani, B., & de Beurs, E. (1995a). Validation of the Dutch adaptation of the Buss-Durkee Hostility Inventory. *Behavior Research and Therapy*, 33, 229-233.
- Lange, A., Hoogendoorn, M., Wieder-spahn, A., & de Beurs, E. (1995b). *Buss-Durkee Hostility Inventory-Dutch, BDHI-D. Handleiding, verantwoording en normering van de Nederlandse Buss-Durkee-agressievragenlijst*. Houten: Bohn Stafleu van Loghum.
- Lange, A., Pahlich, A., Sarucco, M., Smits, G., Dehghani, B., & Hanewald, G. (1995c). Psychometric characteristics and validity of the Dutch adaptation of the Buss-Durkee Hostility Inventory (The BDHI-D). *Behavior Research and Therapy*, 33, 223-227.
- Lischewski, J. J. & de Ruiter, C. (1989). *De Buss-Durkee Hostility Inventory: Handleiding, verantwoording en normering van de Nederlandse Buss-Durkee Agressievragenlijst*. Lisse: Swets & Zeitlinger.
- Luteijn, F., & Barelds, D. P. H. (2003). *Handleiding Negativisme, Ernstige Psychopathologie en Somatisering Test (NPST)*. Amsterdam: Harcourt Test Publishers.
- Luteijn, F., Starren, J., & van Dijk, H. (1975). *Nederlandse Persoonlijkheidsvragenlijst, Handleiding*. Lisse: Swets & Zeitlinger.
- Luteijn, F., Starren, J. & van Dijk, H. (2000). *Handleiding Nederlandse Persoonlijkheids-Vragenlijst* (herziene uitgave). Lisse: Swets & Zeitlinger.
- Meesters, C., Muris, P., Bosma, H., Schouten, E., & Beuving, S. (1996). Psychometric evaluation of the Dutch version of the Aggression Questionnaire. *Behaviour Research and Therapy*, 34, 839-843.
- Morren, M. & Meesters, C. (2002). Validation of the Dutch version of the Aggression Questionnaire in adolescent male offenders. *Aggressive Behavior*, 28, 87-96.
- Nationaal Kompas Volksgezondheid, versie 4.14, 12 december 2013; © RIVM, Bilthoven. Geraadpleegd op 21-01-2014, van [www.nationaalkompas.nl/object\\_binary/o8836.png](http://www.nationaalkompas.nl/object_binary/o8836.png)
- Novaco, R. W. (1994). Anger as a risk factor for violence among the mentally disordered. In J. Monahan & H. J. Steadman (Eds.), *Violence and mental disorder: Developments in risk assessment* (pp. 14-29). Chicago, USA: The University of Chicago Press.
- Novaco, R. W. (2003). *The Novaco Anger Scale and Provocation Inventory (NAS-PI)*. Los Angeles, USA: Western Psychological Services.
- Novaco, R. W. & Taylor, J. L. (2004). Assessment of anger and aggression in male offenders with developmental disabilities. *Psychological Assessment*, 16, 42-50.
- Ouborg, M. J. (1988). *Klinische psychodiagnostiek en psychotherapie*. Academisch proefschrift Rijksuniversiteit Groningen.
- Ploeg, F. A. E. (1980). De psychologische aspecten van reumatoïde artritis. *Ac-*

- demisch proefschrift Rijksuniversiteit Groningen.
- Polaschek, D. L. L., Bell, R. K., Calvert, S. W., & Takarangi, M. K. T. (2010). Cognitive-behavioural rehabilitation of high-risk violent offenders: Investigating treatment change with explicit and implicit measures of cognition. *Applied Cognitive Psychology, 24*, 437-449.
- Richetin, J. & South Richardson, D. & Mason, G. D. (2010). Predictive Validity of IAT Aggressiveness in the Context of Provocation. *Social Psychology, 41*, 27-34.
- Spren, M., Brand, E. F. J. M., ter Horst, P., Willems, M., & Bogaerts, S. (2013). *Handleiding HKT-R. Historische, Klinische en Toekomst – Revisie 2013*. HKT-R. Groningen.
- van der Ploeg, H. M., Defares, P. B., & Spielberger, C. D. (1982). *Handleiding bij de Zelf-Analyse Vragenlijst (ZAV)*. Lisse: Swets & Zeitlinger.
- Vertommen, H., & Rochette, F. (1975), *Handleiding Vragenlijst Interpersoonlijke Relaties (VIR)*. Lisse: Swets & Zeitlinger.
- Webster, C. D., Eaves, D., Douglas, K. S. & Wintrup, A. (1995). *The HCR-20 scheme: The assessment of dangerousness and risk*. Vancouver: Simon Fraser University and British Columbia Forensic Psychiatric Services Commission.
- Werkgroep Risicotaxatie Forensische Psychiatrie (2003). *Handleiding HKT-30 versie 2002. Risicotaxatie in de forensische psychiatrie*. Ministerie van Justitie. Dienst Justitiële Inrichtingen.
- Williams, T. Y., Boyd, J. C., Cascardi, M. A., & Poythress, N. (1996). Factor structure and convergent validity of the Aggression Questionnaire in an offender population. *Psychological Assessment, 8*, 398-403.