

Het gebruik van zelfrapportage vragenlijsten in de forensische psychiatrie

RUUD HORNSVELD, FLOOR KRAAIMAAT, PETER MURIS, HENK NIJMAN,
ALMAR ZWETS, SABINA ROZA & HJALMAR VAN MARLE

Samenvatting

De validiteit van zelfrapportage vragenlijsten is omstreden, vooral wanneer deze worden gebruikt voor het assessment van forensisch psychiatrische patiënten of de evaluatie van behandelprogramma's. Dit artikel geeft een overzicht van interne en externe factoren die van invloed kunnen zijn op het invullen van zelfrapportage vragenlijsten door verschillende forensische populaties. Interne factoren zijn factoren die betrekking hebben op eigenschappen van de individuele respondent, zoals de neiging tot sociaal wenselijk reageren, inzicht in eigen gedrag en moreel besef. Externe factoren hebben betrekking op de omstandigheden waaronder de vragenlijsten worden afgenomen, zoals de aan- of afwezigheid van een onderzoeker, het format en het onderwerp van de vragenlijsten en de setting waarin de respondent verblijft. Ook worden de psychometrische eigenschappen besproken van vragenlijsten die in de forensische psychiatrie veel gebruikt worden. Ten slotte worden er aanbevelingen gedaan voor het gebruik van zelfrapportage vragenlijsten in de forensische psychiatrie.

Trefwoorden: zelfrapportage vragenlijsten, psychometrische eigenschappen, forensische psychiatrie

INLEIDING

Zelfrapportage vragenlijsten worden veel gebruikt in de forensische psychiatrie, voor zowel het assessment van individuele patiënten als de evaluatie van behandelprogramma's. Het gebruik van dergelijke vragenlijsten staat

bij deze populatie echter ter discussie. In dit artikel worden de verschillende factoren besproken die van invloed kunnen zijn op het invullen van zelfrapportage vragenlijsten door forensisch psychiatrische patiënten, alsmede de psychometrische eigenschappen van enkele veelgebruikte, vertaalde vragenlijsten. Het artikel wordt afgesloten met conclusies en aanbevelingen.

Forensisch psychiatrische patiënten

Forensisch psychiatrische patiënten zijn delinquenten bij wie de rechter op grond van onderzoek door een psychiater en/of een psycholoog (rapportage pro Justitia) een verband heeft geconstateerd tussen 'een gebrekkige ontwikkeling of ziekelijke stoornis van de geestvermogens' en een gepleegd delict (van Marle, 2000). Zonder klinische of ambulante behandeling wordt de kans op herhaling reëel geacht. Belangrijkste doelstelling van de behandeling in of bij een forensisch psychiatrische instelling is daarom het bewerkstelligen van een zodanige gedragsverandering dat de kans op herhaling tot een aanvaardbaar minimum wordt teruggebracht. Terbeschikkinggestelden zijn forensisch psychiatrische patiënten die een delict hebben gepleegd met een gevangenisstraf van ten minste vier jaar. Behandeling vindt plaats in een gesloten, beveiligde instelling. Ambulante forensisch psychiatrische patiënten zijn door de rechter verplicht een behandeling op een forensisch psychiatrische polikliniek te volgen. Behandeling kan worden opgelegd als het gaat om: (1) een aanvullende voorwaardelijke straf bij delicten met een gevangenisstraf van maximaal drie jaar, (2) een alternatieve straf bij delicten met een gevangenisstraf van maximaal zes maanden, (3) een onderdeel van een penitentiair programma, en (4) een onderdeel van een ondertoezichtstelling door de jeugdzorg.

INTERNE FACTOREN

In deze paragraaf bespreken we literatuur over een aantal interne factoren die van invloed kunnen zijn op de beantwoording van zelfrapportage vragenlijsten, namelijk de neiging tot sociaal wenselijk reageren (*socially desirable responding*; SDR), inzicht in eigen gedrag en moreel besef.

Socially desirable responding (SDR)

SDR wordt meestal gedefinieerd als de neiging van respondenten om zichzelf in een gunstig daglicht te plaatsen (Nederhof, 1985; Paulhus, 2002). In het onderzoek naar de invloed van SDR wordt doorgaans gebruikgemaakt van twee instrumenten, namelijk de Marlow-Crowne Social Desirability Scale (MC-SDS; Crowne & Marlow, 1960) en de Balanced Inventory of Desirable Responding (BIDR; Paulhus, 1984, 1988).

De MC-SDS bestaat uit 33 items, waarvan er 15 negatief (bijvoorbeeld 'I am sometimes irritated by people who ask favors of me') en 18 positief (bijvoorbeeld 'I never hesitate to go out of my way to help someone in trouble') geformuleerd zijn. De antwoordmogelijkheden zijn 'waar' of 'niet waar'. Een sociaal wenselijk antwoord krijgt de score 1, in het andere geval is de score 0. De minimumscore is derhalve 0 en de maximumscore 33. Andrews en Meyer (2003) namen de oorspronkelijke versie van de MC-SDS met 33 items en de verkorte versie met 13 items (MC-C) af bij 1096 personen die werden verdacht van een delict. Delinquenten scoorden significant hoger op beide versies van de MC-SDS dan niet-delinquenten. De auteurs concludeerden derhalve dat beide versies gebruikt kunnen worden om zelfbedrog te meten als onderdeel van een forensisch assessment.

De BIDR kent twee schalen van elk 20 items, namelijk Zelfbedrog (de neiging om eerlijke, maar positief gekleurde antwoorden te rapporteren, bijvoorbeeld 'I rarely appreciate criticism') en Indruk Maken (doelbewust een positieve indruk maken op anderen, bijvoorbeeld 'I always declare everything at customs'). De Zelfbedrog-schaal werd later (Paulhus, 2002) onderverdeeld in twee subschalen, te weten Zelfbedrog Versterking (bijvoorbeeld 'My first impressions always turn out to be right') en Zelfbedrog Ontkenning (bijvoorbeeld 'I never regret my decisions'). De items worden gescoord op een zevenpuntsschaal, waarbij een score van 1 tot en met 5 telt als 0 en een score 6 of 7 als 1. Hoe hoger de score, des te sociaal wenselijker is er geantwoord. In een onderzoek van Paulhus (1988) bleek de totaalscore op de BIDR .71 te correleren met de totaalscore op de MC-SDS.

Onderzoek naar SDR bij universiteitsstudenten — Er is relatief veel onderzoek gedaan naar de invloed van SDR op de scores van zelfrapportage vragenlijsten voor persoonlijkheidstrekken bij universiteitsstudenten. Zo vond Holden (2007) dat de opdracht om zo gunstig mogelijk te antwoorden van grote invloed bleek op de criteriumvaliditeit, de mate waarin een test-score het criterium (in dit geval een persoonlijkheidstrek) voorspelt. Holden en Passey (2010) stelden echter vast dat dit niet het geval was als er voor de deelnemers weinig op het spel stond. Holden en Book (2012) concludeerden dat de relatie tussen SDR en zelfrapportage van persoonlijkheidskenmerken gerelateerd is aan een aantal factoren, te weten: (1) de manier waarop SDR is geoperationaliseerd, (2) de onderzoeksprocedure (SDR is impliciet of wordt experimenteel geïnduceerd), (3) het belang dat er door de proefpersonen aan de uitslag wordt toegekend, en (4) de psychometrische eigenschappen van de gebruikte persoonlijkheidsvragenlijst. Een vijfde factor die mogelijk een rol speelt is het onderwerp van een vragenlijst (Hendriks, 2011). Zo onderzochten Davis, Thake en Vilhena (2010) het alcoholgebruik van 568 universiteitsstudenten met gebruikmaking van onder andere de BIDR (Paulhus, 2002) voor het meten van SDR bij het antwoorden van vragenlijsten over alcoholmisbruik. Uit hun onderzoek kwam naar voren dat studenten

die hoog scoorden op de subschaal Indruk Maken 20% tot 33% minder alcoholgebruik en ongeveer 50% minder vaak dronkenschap rapporteerden dan studenten die laag op die subschaal scoorden. Tussen de subschaal Zelfbedrog en alcoholconsumptie werd geen significant verband gevonden. De auteurs concludeerden dat bij personen met hoge scores op Indruk Maken de validiteit van zelfgerapporteerd gebruik van alcohol met bijbehorende risico's beperkt is, omdat alcoholmisbruik door de respondenten als sociaal afwijkend gedrag wordt beschouwd. Het gebruik van zelfrapportage vragenlijsten bij dergelijke personen kan volgens de onderzoekers leiden tot een onderschatting van schadelijk en risicovol alcoholgebruik.

Onderzoek naar SDR bij delinquenten — Delinquenten worden vaak verondersteld zelfrapportage vragenlijsten sociaal wenselijk in te vullen, omdat ze zouden denken dat een gunstige score een positieve invloed heeft op hun juridische status (Kalmus & Beech, 2005). In de literatuur zijn verschillende studies te vinden waarin de manier van invullen door delinquenten is onderzocht. Mills, Loza en Kroner (2003) bijvoorbeeld, vonden in hun onderzoek bij 124 mannelijke delinquenten dat de Self-Appraisal Questionnaire (SAQ; Loza, 1996), een zelfrapportage vragenlijst over delictgerelateerd gedrag, significant negatief gerelateerd was aan beide subschalen van de BIDR, dat wil zeggen dat de scores op de SAQ niet beïnvloed werden door SDR. In een vervolgonderzoek (Mills & Kroner, 2006) bij 172 mannelijke delinquenten naar de validiteit van de Measure of Criminal Attitudes and Associates (MCAA; Mills & Kroner, 1999), een zelfrapportage vragenlijst over antisociale attitudes, bleken delinquenten met een hoge score op de subschaal Indruk Maken van de BIDR tegen de verwachting in minder antisociale attitudes te rapporteren. Deze opmerkelijke resultaten werden opnieuw gevonden bij een derde studie van Mills en Kroner (2005) in een groep van 97 mannelijke delinquenten: delinquenten die laag scoorden op SDR en die dus betrekkelijk eerlijk antwoordden, hadden een groter recidiverisico.

Onderzoek naar SDR bij seksueeldelictplegers — Bij deze populatie kunnen specifieke factoren bijdragen tot SDR, zoals het willen maken van een macho indruk of het willen vermijden van de diagnose 'seksueel deviant' (Kroner & Weekes, 1996). Bumby (1996) ontwikkelde de MOLEST-schaal en de RAPE-schaal voor het assessment van 'cognitieve vervormingen' (antisociale cognities die seksueeldelictgedrag entameren of rechtvaardigen) bij respectievelijk seksueel kindermisbruikers en verkrachters van volwassen vrouwen. Beide schalen bleken niet significant te correleren met de MCSDS. Arkowitz en Vess (2003) meenden echter dat de MOLEST-schaal en de RAPE-schaal te gevoelig zijn voor SDR om te gebruiken bij seksueeldelictplegers die een verplichte behandeling moeten ondergaan, maar ze onderzochten deze veronderstelling niet. Mathie en Wakeling (2011) onderzochten de invloed van SDR op zelfrapportage vragenlijsten bij 1730 mannelijke

seksueledelictplegers. De SDR werd gemeten met de BIDR (Paulhus, 1984), het sociaal functioneren met een negental vragenlijsten en de delictgerelateerde factoren met zeven vragenlijsten plus een risicotaxatie-instrument. Uit het onderzoek kwam naar voren dat de mate van SDR en de invloed ervan op zelfrapportage vragenlijsten minder groot was dan verwacht. Correlaties tussen SDR en delictspecifieke meetinstrumenten varieerden bij de subschaal Zelfbedrog van $-.34$ tot $.36$, en bij de subschaal Indruk Maken van $-.26$ tot $.20$. De onderzoekers concludeerden dat in de forensische psychiatrie gebruikte zelfrapportage vragenlijsten acceptabel accuraat en valide kunnen zijn. Alles bij elkaar lijkt voornamelijk een eerdere conclusie van Gannon, Ward en Collie (2007) van toepassing, namelijk dat 'the questionnaire literature presents a mixed bag of findings when examined as a whole' (p. 410).

Inzicht in eigen gedrag en moreel besef

Opvallend bij het onderzoek naar de invloed van SDR op de beantwoording van zelfrapportage vragenlijsten is dat daarbij de vraag steeds ontbrak in hoeverre respondenten inzicht hebben in hun eigen gedrag en in de morele standaarden van de maatschappij waarin zij leven. Een vragenlijst kan alleen eerlijk worden ingevuld als de respondent goed inzicht heeft in de wijze waarop hij zich in een bepaalde situatie gedraagt. Wil de respondent sociaal wenselijk antwoorden, dan moet hij wel besef hebben van het gedrag dat in een bepaalde situatie door de meeste mensen als normaal beschouwd wordt, de maatschappelijke norm.

Onderzoek naar inzicht in eigen gedrag is vooral gedaan bij psychiatrische patiënten zoals schizofrenen (bijvoorbeeld: Bell, Fiszdon, Richardson, Lysaker, & Bryson, 2007), manische patiënten (bijvoorbeeld: Adida et al., 2008) en adolescenten met psychiatrische stoornissen (bijvoorbeeld: Salbach-Andrae, Klinkowski, Lenz, & Lehmkuhl, 2009). Er is tot op heden echter nauwelijks onderzoek gedaan naar inzicht in eigen gedrag bij psychiatrische patiënten met een cluster B-persoonlijkheidsstoornis en bij niet-klinische volwassenen met een laag opleidingsniveau en/of een lage sociaal-economische status. Als het gaat om moreel besef hebben meerdere studies ondersteuning gevonden voor de bevinding dat veel delinquenten een beperkte kennis hebben van gangbare normen en waarden. Stams en collega's (2006) bijvoorbeeld, deden een meta-analyse van vijftig studies en concludeerden dat delinquenten gemiddeld een lager niveau van morele rijpheid hebben dan niet-delinquenten. Van Vugt en collega's (2011) voerden eveneens een meta-analyse uit, van negentien studies, en vonden dat een lagere fase van morele ontwikkeling gerelateerd was aan een hoger recidiverisico.

Samenvatting en bespreking

282

De MC-SDS (33- en 13-itemversie) en de BIDR lijken voornamelijk betrouwbare en valide instrumenten voor het meten van sociaal wenselijk reageren. Uit onderzoek bij verschillende populaties komt steeds naar voren dat sociaal wenselijk reageren lijkt samen te hangen met factoren als (1) het belang dat door de respondent aan de uitkomst wordt gehecht (bijvoorbeeld: bij delinquenten hun juridische status), (2) het thema van de vragenlijst (persoonkenmerken versus ongunstig gedrag), en (3) de onderzoeksprocedure (SDR is impliciet of wordt experimenteel geïnduceerd). De MC-SDS en de BIDR hebben echter als beperking dat deze vragenlijsten werden ontwikkeld met behulp van steekproeven universiteitsstudenten en dat de veronderstelde factorstructuur niet is onderzocht in populaties als delinquenten of forensisch psychiatrische patiënten. Helaas ontbreekt het ook aan onderzoek naar de relatie tussen SDR enerzijds en inzicht in eigen gedrag en moreel besef anderzijds.

EXTERNE FACTOREN

In deze paragraaf komen externe factoren aan de orde die het antwoorden op zelfrapportage vragenlijsten kunnen beïnvloeden, zoals de aan- of afwezigheid van een onderzoeker, het format van de vragenlijsten, het onderwerp ervan, en de setting waarin de respondent verblijft.

Aan- of afwezigheid van een onderzoeker

De afgelopen jaren is het digitaal afnemen van vragenlijsten sterk toegenomen. Campos, Zucoloto, Bonafé, Jordani en Maroco (2011) vergeleken bij een drietal zelfrapportage vragenlijsten over burn-out een elektronisch met een schriftelijk format en concludeerden dat elektronische vragenlijsten een gebruiksvriendelijke en redelijk betrouwbare manier zijn om data te verzamelen. Zij noemen als voordeel vooral het niet feitelijk aanwezig zijn van de onderzoeker. Hiermee zou eventuele ongewilde beïnvloeding door de onderzoeker worden verminderd.

In de forensische psychiatrie lijkt het echter raadzaam om de vragenlijsten te laten invullen in aanwezigheid van een onderzoeker. Veel patiënten zijn laagopgeleid en begrijpen soms bepaalde woorden niet (bijvoorbeeld: 'Wat is we-dijveren?' en 'Wat is een gedicht?'). Daarnaast kan een onderzoeker zien hoe de vragenlijsten worden ingevuld, variërend van overdreven secuur tot afgeraffeld. Vaak moeten patiënten gemotiveerd worden om door te gaan met invullen, omdat ze gauw verveeld of ontmoedigd raken. Als vragenlijsten worden afgenomen in het kader van wetenschappelijk onderzoek is het

van belang aan de respondenten te laten weten dat de uitkomsten niet worden opgenomen in het dossier (Drenth & Sijtsma, 2006, p. 150).

Vrije antwoorden of meerkeuzeantwoorden

Verscheidene auteurs hebben ervoor gepleit bij gedetineerden gebruik te maken van 'productie-instrumenten' in plaats van multiplechoice 'herkenningsinstrumenten' (Gavaghan, Arnold, & Gibbs, 1983; Stams et al., 2006; Van Vugt et al., 2011). Drenth en Sijtsma (2006) spreken van vrije-antwoordentests versus keuze-antwoordentests (p. 113). Bij vrije-antwoorden vragenlijsten wordt respondenten gevraagd hun reacties op een aantal plaatjes of uitspraken op te schrijven, waarna die reacties worden gescoord door een onafhankelijke beoordelaar. Deze geschreven reacties zouden directere informatie over het gedrag van respondenten geven, omdat ze niet hoeven na te denken over wat hun juiste score is op de likertschaal, zoals bij keuze-antwoorden vragenlijsten. Nadeel van de vrije-antwoorden vragenlijsten is onder andere dat het beoordelen en categoriseren van de antwoorden voor de onderzoeker een tijdrovende aangelegenheid is (Drenth & Sijtsma, 2006, p. 120).

Setting

In de literatuur wordt relatief weinig aandacht besteed aan de mogelijke invloed van de setting waarin de respondent verblijft (opgenomen of gedetineerd versus in ambulante behandeling versus niet-klinische vrijwilliger) op het beantwoorden van vragenlijsten. Zo vonden Olver, Kingston, Nicholaichuk en Wong (2014) bij 392 gedetineerde seksueel delictplegers dat 'they did not score in a particularly pathological manner on the battery of psychometric measures' (p. 549), dat wil zeggen dat hun score niet afweek van die van niet-klinische vrijwilligers. Kelsey, Rogers en Robinson (2015) namen de Psychopathic Personality Inventory-Revised (PPI-R; Lilienfeld & Widows, 2005), een zelfrapportage vragenlijst van 154 items voor het meten van psychopathie, af in een groep van 86 gedetineerden. De gedetineerden bleken lager te scoren op de vragenlijst dan studenten en andere niet-klinische vrijwilligers. Kanters en collega's (2016) vergeleken 28 klinische met 61 poliklinische kindermisbruikers, respectievelijk opgenomen in een forensisch psychiatrische instelling of in behandeling op een forensisch psychiatrische polikliniek. Gemeten met de Sexual Violence Risk-20 (SVR-20; Boer, Wilson, Gauthier, & Hart, 1997) bleek dat de poliklinische patiënten een significant lager recidiverisico hadden dan de klinische patiënten. Op de zelfrapportage vragenlijsten over agressie, vijandigheid en sociale angst scoorden de klinische kindermisbruikers echter significant lager dan de poliklinische. Bovendien deed de eerste groep significant vaker positieve zelfuitspraken dan de tweede groep. Een mogelijke verklaring voor deze opval-

lende verschillen kan zijn dat toezicht en structuur van een gevangenis of forensisch psychiatrische instelling een dempend effect hebben op negatief gedrag (Hornsveld, Kraaimaat, Bouwmeester, Polak, & Zwets, 2014) en dat klinische patiënten bijgevolg de vragenlijsten naar waarheid invulden.

Samenvatting en bespreking

Het lijkt raadzaam om in de forensische psychiatrie zo veel mogelijk gebruik te maken van vrije-antwoorden vragenlijsten, waarbij respondenten zelf een reactie formuleren in plaats van er een aan te kruisen. Echter, nader onderzoek is geboden naar de verschillen in validiteit tussen vrije-antwoorden en keuze-antwoorden vragenlijsten bij volwassen delinquenten of forensisch psychiatrische patiënten. De aanwezigheid van een intaker of onderzoeker bij de afname heeft als voordeel dat deze de patiënten kan motiveren de lijsten volledig in te vullen en vragen over moeilijke woorden kan beantwoorden. De invloed van de setting op de beantwoording van vragenlijsten lijkt aanzienlijk. De opvallende verschillen in scores op vragenlijsten tussen gedetineerden of klinische forensisch psychiatrische patiënten enerzijds en poliklinische patiënten of niet-klinische vrijwilligers anderzijds verdienen nader onderzoek. Kunnen die verschillen verklaard worden door de structuur en het toezicht van de forensisch psychiatrische instelling of spelen ook andere factoren een rol, zoals verschillen in psychopathologie? Of zijn gedetineerden in een gevangenis of forensisch psychiatrische instelling bedrever in het maskeren van hun probleemgedrag dan poliklinische patiënten of niet-klinische vrijwilligers?

VEELGEBRUIKTE, VERTAALDE VRAGENLIJSTEN

Vertaling

Bij een aantal veelgebruikte zelfrapportage vragenlijsten doet zich een probleem van heel andere aard voor, namelijk dat ze uit het Engels zijn vertaald en dat deze vertalingen niet bij Nederlandse forensisch psychiatrische patiënten op psychometrische eigenschappen zijn onderzocht. Zo bleek uit eigen onderzoek dat de oorspronkelijke factor- of subschalenstructuur van de Agressie Vragenlijst (AQ; Buss and Perry, 1992; Meesters, Muris, Bosma, Schouten, & Beuving, 1996), de Novaco Anger Scale-Provocation Inventory (NAS-PI; Novaco, 1994) en de Bermond-Vorst Alexithymia Questionnaire (BVAQ; Vorst & Bermond, 2001) niet kon worden geconfirmeerd in steekproeven van Nederlandse forensisch psychiatrische patiënten. Een uitzondering was de Aangepaste Versie van de Agressie Vragenlijst (AVL-AV; Hornsveld, Muris, Kraaimaat, & Meesters, 2009), een vertaling van de

Aggression Questionnaire-Short Form (AQ-SF; Bryant & Smith, 2001). Bovendien werd alleen de oorspronkelijke NAS-PI onderzocht in Amerikaanse, mannelijke delinquenten. De AQ, de AVL, de AQ-SF en de BVAQ werden ontwikkeld in gemengde groepen van mannelijke en vrouwelijke universiteitsstudenten in verschillende landen, veelal psychologiestudenten.

Normgroepen

.....

Een ander probleem bij het gebruik van de veelgebruikte zelfrapportage vragenlijsten in de forensische psychiatrie betreft het ontbreken van relevante normgroepen (Akkerman-Bouwsema, Biesbroek, Lancel, & Gerlisma, 2015). Als er voor een vragenlijst al normgroepen zijn, dan is vaak niet duidelijk wat het opleidingsniveau van de onderzochte steekproef is. Omdat veel forensisch psychiatrische patiënten een relatief laag opleidingsniveau hebben, werden om hieraan enigszins tegemoet te komen normen ontwikkeld door vragenlijsten te laten invullen door mannelijke (v)mbo-leerlingen (Hornsveld, Kraaimaat, Muris, Zwets, & Kanters, 2015).

CONCLUSIES EN AANBEVELINGEN

.....

In dit artikel werden drie soorten kanttekeningen geplaatst bij het gebruik van zelfrapportage in de forensische psychiatrie, namelijk (1) de invloed van interne factoren als SDR, inzicht en moreel besef, (2) de invloed van externe factoren als de aan- of afwezigheid van een onderzoeker, het format en het onderwerp van de vragenlijsten en de setting waarin de respondent verblijft, en (3) de psychometrische eigenschappen van enkele veelgebruikte, vertaalde zelfrapportage vragenlijsten. Op grond van de literatuur en van klinische ervaring trekken we de volgende conclusies.

Een belangrijk probleem bij het gebruik van zelfrapportage vragenlijsten voor forensisch psychiatrische patiënten is dat de meeste niet of onvoldoende op psychometrische eigenschappen bij deze doelgroep zijn onderzocht en dat relevante normen ontbreken. Het gebruik van subschaalscores is af te raden als de betreffende subschalen niet door psychometrisch onderzoek bij Nederlandse populaties worden ondersteund. Normen van niet-klinische vrijwilligers kunnen van belang zijn voor poliklinische patiënten, maar zijn dat minder voor klinische patiënten. Bij het gebruik van zelfrapportage vragenlijsten in de forensische psychiatrie zal er altijd rekening mee moeten worden gehouden dat respondenten de lijsten mogelijk sociaal wenselijk invullen en dat hun inzicht in eigen gedrag en hun moreel besef beperkt kunnen zijn. Dit is vooral van belang als de scores op de vragenlijsten gebruikt worden voor beslissingen op individueel niveau en er voor de respondent een groot belang op het spel staat.

Verschillende onderzoekers (bijvoorbeeld: Andrews & Bonta, 2010) pleiten voor de combinatie van verschillende meetinstrumenten, zoals gestructureerde interviews, observatieschalen en incidentenregistraties. Gebruik van meerdere soorten instrumenten is ook naar onze mening noodzakelijk, omdat elk instrument specifieke beperkingen heeft. In een gesloten instelling kan een vergelijking van de score op een vragenlijst met die op een observatieschaal relevante informatie opleveren over het inzicht van een patiënt in zijn eigen gedrag (Nijman, Bjørkly, Palmstierna, & Almvik, 2006). Per instrument moet echter steeds een afweging gemaakt worden tussen de betekenis van de verzamelde informatie enerzijds en de beperkingen van het betreffende instrument anderzijds (Drenth & Sijsma, 2006, p. 68).

In het kader van wetenschappelijk onderzoek lijkt het minder zinvol de scores van klinische patiënten op zelfrapportage vragenlijsten te vergelijken met die van poliklinische patiënten. Als zelfrapportage vragenlijsten worden gebruikt voor bijvoorbeeld de effectmeting van behandelprogramma's is het gebruik van een instrument voor het meten van sociaal wenselijk reageren aan te raden. De MC-SDS en BIDR kunnen voorlopig hiervoor gebruikt worden. Het is echter van belang dat de psychometrische eigenschappen van beide instrumenten alsnog bij Nederlandse populaties worden onderzocht.

NIEUWE ONTWIKKELINGEN

.....

Er zijn steeds meer aanwijzingen dat zelfrapportage vragenlijsten voor dynamische (dat wil zeggen veranderbare, door interventies te beïnvloeden) factoren een waardevolle bijdrage kunnen leveren aan de risicotaxatie van (seksueel)delictplegers (Douglas & Skeem, 2005). Beech, Friendship, Erikson en Hanson (2002) bijvoorbeeld, vonden in een groep van 140 kindermisbruikers dat het toevoegen van de vragenlijstgegevens over dynamische risicofactoren resulteerde in een significante verbetering van de risicotaxatie, die aanvankelijk gebaseerd was op statische (dat wil zeggen onveranderbare, vaak historische) factoren. De verschillende studies naar de meerwaarde van dynamische factoren voor risicotaxatie zijn echter uitgevoerd bij groepen delinquenten, zodat de resultaten niet zonder meer van toepassing zijn op individuele delinquenten. Volgens Hart, Michie en Cooke (2007) kan tot op heden risicotaxatie niet gebruikt worden om in individuele gevallen de kans op toekomstig geweld met een zekere mate van betrouwbaarheid te schatten en moet er daarom met de nodige voorzichtigheid maar liever helemaal geen gebruik van worden gemaakt (p. 60).

Voor een aantal dynamische factoren van recidiverisico zijn inmiddels specifieke vragenlijsten ontwikkeld, bijvoorbeeld voor het meten van vijandigheid (PSF-AV; Hornsveld, Nijman, Hollin, & Kraaimaat, 2007), morele rijpheid (SRM-AV; Hornsveld, Kraaimaat, & Zwets, 2012) en attitude jegens vrouwen (HVL; Hornsveld, Timonen, Kraaimaat, Zwets, & Kanters, 2014).

Alle drie de instrumenten zijn vrije-antwoorden vragenlijsten. Deze drie nieuwe vragenlijsten bevinden zich echter nog in een eerste stadium en verder psychometrisch onderzoek is dan ook geboden.

Een tweede interessante ontwikkeling is het gebruik van zelfrapportage vragenlijsten als risicotaxatie-instrument. Enkele onderzoekers vonden namelijk dat het voorspellend vermogen van sommige zelfrapportage vragenlijsten overeenkwam met dat van veelgebruikte actuariële instrumenten¹ (Loza & Green, 2003; Walters, 2006). Ook hier zal verder onderzoek moeten uitwijzen of bepaalde zelfrapportage vragenlijsten gebruikt kunnen worden voor het vaststellen van het recidiverisico op individueel niveau, vooral omdat er voor een respondent grote belangen op het spel kunnen staan.

Alles bij elkaar is er naar onze mening op dit moment geen reden om zelfrapportage vragenlijsten te verbannen uit de forensische psychiatrie, maar gebruikers dienen zich wel rekenschap te geven van de beperkingen van de in dit artikel besproken instrumenten.

Ruud H. J. Hornsveld is verbonden aan het Erasmus Universitair Medisch Centrum.

Floor W. Kraaiaam is verbonden aan de Radboud Universiteit, Nijmegen.

Peter Muris is verbonden aan Maastricht University, Maastricht.

Henk L. I. Nijman is verbonden aan APC Altrecht, Den Dolder en Radboud Universiteit, Nijmegen.

Almar J. Zwets is verbonden aan FPC de Kijvelanden, Poortugaal.

Sabina. J. Roza en **Hjalmar J. C. van Marle** zijn verbonden aan het Erasmus Universitair Medisch Centrum.

Correspondentieadres: Ruud H.J. Hornsveld, Lange Dreef 52, 2285 LA Rijswijk. E-mail: Ruud.Hornsveld@Ziggo.nl

Summary *The use of self-reporting questionnaires in forensic psychiatry*

The validity of self-reporting questionnaires is controversial, especially when used for the assessment of forensic psychiatric patients or the evaluation of treatment programs. This article provides an overview of the internal and external factors that may affect the responding to these scales in various forensic populations. Internal factors refer to factors that relate to individual respondent characteristics, such as the tendency to socially desirable responding, insight into one's own behavior, and moral awareness. External factors relate to the circumstances in which questionnaires are taken, such as the presence or absence of an assessor, the format and subject of the questionnaires, and the setting in which the respondent resides. Furthermore, the psychometric properties are discussed of a few questionnaires that are commonly used in Dutch forensic

1 Met behulp van een actuariële instrument kan delictgevaar worden geschat op basis van factoren waarvan in empirisch onderzoek is aangetoond dat zij samenhangen met gewelddadig gedrag.

psychiatry. Finally, conclusions and recommendations are presented for the use of self-report questionnaires in forensic psychiatry.

Keywords *self-report questionnaires, psychometric properties, forensic psychiatry*

Literatuur

- Adida, M., Clark, L., Pomietto, P., Kaladjian, A., Besnier, N., Azorina, J., ... Goodwin, G. M. (2008). Lack of insight may predict impaired decision making in manic patients. *Bipolar Disorders, 10*, 829-837.
- Akkerman-Bouwsema, G., Biesbroek, B., Lancel, M., & Gerlisma, C. (2015). De bruikbaarheid van agressievragenlijsten in de Nederlandse forensisch-psychiatrische praktijk. *Gedragstherapie, 48*, 282-300.
- Andrews, D. A., & Bonta, J. (2010). *The psychology of criminal conduct* (5th ed.). New Providence, NJ: Matthew Bender & Company.
- Andrews, P., & Meyer, R. G. (2003). Marlow-Crowne Social Desirability Scale and short form C: forensic norms. *Journal of Clinical Psychology, 59*, 483-492.
- Arkowitz, S., & Vess, J. (2003). An evaluation of the Bumby RAPE and MOLEST scales as measures of cognitive distortions with civilly committed sexual offenders. *Sexual Abuse: A Journal of Research and Treatment, 15*, 237-249.
- Beech, A., Friendship, C., Erikson, M., & Hanson, R. K. (2002). The relationship between static and dynamic risk factors and reconviction in a sample of U.K. child abusers. *Sexual Abuse: A Journal of Research and Treatment, 14*, 155-167.
- Bell, M., Fiszdon, J., Richardson, R., Lysaker, P., & Bryson, G. (2007). Are self-reports valid for schizophrenia patients with poor insight? Relationship of unawareness of illness to psychological self-report instruments. *Psychiatry Research, 151*, 37-46.
- Boer, D. P., Wilson, R. J., Gauthier, C. M., & Hart, S. D. (1997). Assessing risk for sexual violence: Guidelines for clinical practice. In C. D. Webster & M. A. Jackson (Eds.), *Impulsivity: Theory, assessment, and treatment* (pp. 326-342). New York: Guilford.
- Bryant, F. B., & Smith, B. D. (2001). Refining the architecture of aggression: A measurement model for the Buss-Perry Aggression Questionnaire. *Journal of Research in Personality, 35*, 138-167.
- Bumby, K. M. (1996). Assessing the cognitive distortions of child molesters and rapists: Development and validation of the MOLEST and RAPE Scales. *Sexual Abuse: A Journal of Research and Treatment, 8*, 37-54.
- Buss, A. H., & Perry, M. (1992). The aggression questionnaire. *Journal of Personality and Social Psychology, 63*, 452-459.
- Campos, J. A. D. B., Zucoloto, M. L., Bonafé, F. S. S., Jordani, P. C., & Maroco, J. (2011). Reliability and validity of self-reported burnout in college students: A cross randomized comparison of paper-and-pencil vs. online administration. *Computers in Human Behavior, 27*, 1875-1883.
- Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology, 24*, 349-354.
- Davis, C. G., Thake, J., & Vilhena, N. (2010). Social desirability biases in self-reported alcohol consumption and harms. *Addictive Behaviors, 35*, 302-311.
- Douglas, K. S., & Skeem, J. L. (2005). Violence risk assessment. Getting specific about being dynamic. *Psychology, Public Policy, and Law, 11*, 347-383.

- Drenth, P. J. D., & Sijtsma, K. (2006). *Testtheorie: Inleiding in de theorie van de psychologische test en zijn toepassingen*. Houten: Bohn Stafleu van Loghum.
- Gannon, T. A., Ward, T., & Collie, R. (2007). Cognitive distortions in child molesters: Theoretical and research developments over the past decades. *Aggression and Violent Behavior, 12*, 402-416.
- Gavaghan, M. P., Arnold, K. D., & Gibbs, J. C. (1983). Moral judgment in delinquents and nondelinquents: Recognition versus production measures. *Journal of Psychology, 114*, 267-274.
- Hart, S. D., Michie, C., & Cooke, D. J. (2007). Precision of actuarial risk assessment instruments: Evaluating the 'margins of error' of group v. individual predictions of violence. *British Journal of Psychiatry, 190* (suppl. 49), s60-s65.
- Hendriks, V. (2011). Meten en meetinstrumenten. In I. Franken & W. van den Brink (red.), *Handboek verslaving* (pp. 245-274). Utrecht: De Tijdstroom.
- Holden, R. R. (2007). Socially desirable responding does moderate personality scale validity both in experimental and non-experimental contexts. *Canadian Journal of Behavioural Science, 39*, 184-201.
- Holden, R. R., & Book, A. S. (2012). Faking does distort self-report personality assessment. In M. Ziegler, C. MacCann, & R. D. Roberts (Eds.), *New perspectives on faking in personality assessment* (pp. 71-84). New York: Oxford University Press.
- Holden, R. R., & Passey, J. (2010). Socially desirable responding in personality assessment: Not necessarily faking and not necessarily substance. *Personality and Individual Differences, 49*, 446-450.
- Hornsveld, R. H. J., Kraaimaat, F. W., Bouwmeester, S., Polak, M. A., & Zwets, A. J. (2014). Behavior on the ward of personality disordered inpatients and chronically psychotic inpatients during a three-year stay in a Dutch forensic psychiatric hospital. *Journal of Forensic Psychiatry and Psychology, 25*, 684-704.
- Hornsveld, R. H. J., Kraaimaat, F. W., Muris, P., Zwets, A. J., & Kanters, T. (2015). Aggression Replacement Training for violent young men in a forensic psychiatric outpatient clinic. *Journal of Interpersonal Violence, 30*, 3174-3191.
- Hornsveld, R. H. J., Kraaimaat, F. W., & Zwets, A. J. (2012). The adapted version of the Sociomoral Reflection Measure (SRM-AV) in Dutch forensic psychiatric patients. *International Journal of Forensic Mental Health, 11*, 1-9.
- Hornsveld, R. H. J., Muris, P., Kraaimaat, F. W., & Meesters, C. (2009). The Aggression Questionnaire in Dutch violent forensic psychiatric patients and secondary vocational students. *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 (PFS-AV) for the measurement of hostility in violent forensic psychiatric patients. *Criminal Behaviour and Mental Health, 17*, 45-56.
- Hornsveld, R. H. J., Timonen, B., Kraaimaat, F. W., Zwets, A. J., & Kanters, T. (2014). The Attitudes toward Women Inventory (AWI) in Dutch violent forensic psychiatric inpatients. *Journal of Forensic Psychology Practice, 14*, 383-397.
- Kalmus, E., & Beech, A. R. (2005). Forensic assessment of sexual interest: A review. *Aggression and Violent Behavior, 10*, 193-218.
- Kanters, T., Hornsveld, R. H. J., Nunes, K. L., Zwets, A. J., Buck, N. M. L., Muris, P., & van Marle, H. J. C. (2016). Aggression and social anxiety are associ-

- ated with sexual offending against children. *International Journal of Forensic Mental Health*, 15, 265-273.
- Kelsey, K. R., Rogers, R., & Robinson, E. V. (2015). Self-report measures of psychopathy: What is their role in forensic assessments? *Journal of Psychopathology and Behavioral Assessment*, 37, 380-391.
- Kroner, D. G., & Weekes, J. R. (1996). Balanced Inventory of Desirable Responding: Factor structure, reliability, and validity with an offender sample. *Personality and Individual Differences*, 21, 323-333.
- Lilienfeld, S. O., & Widows, M. (2005). *Professional manual for the psychopathic personality inventory-revised (PPI-R)*. Lutz, FL: Psychological Assessment Resources.
- Loza, W. (1996). *Self-Appraisal Questionnaire (SAQ): A tool for assessing violent and nonviolent recidivism*. Unpublished manuscript.
- Loza, W., & Green, K. (2003). The Self-Appraisal Questionnaire. A self-report measure for predicting recidivism versus clinician-administered measures: A 5-year follow-up study. *Journal of Interpersonal Violence*, 18, 781-797.
- Mathie, N. L., & Wakeling, H. C. (2011). Assessing social desirable responding and its impact on self-report measures among sexual offenders. *Psychology, Crime & Law*, 17, 215-237.
- 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.
- Mills, J. F., & Kroner, D. G. (1999). *Measures of Criminal Attitudes and Associates (MCAA)*. Unpublished instrument and user's guide.
- Mills, J. F. & Kroner, D. G. (2005). An investigation into the relationship between socially desirable responding and offender self-report. *Psychological Services*, 2, 70-80.
- Mills, J. F. & Kroner, D. G. (2006). Impression management and self-report among violent offenders. *Journal of Interpersonal Violence*, 21, 178-192.
- Mills, J. F., Loza, W., & Kroner, D. G. (2003). Predictive validity despite social desirability: Evidence for the robustness of self-report among offenders. *Criminal Behaviour and Mental Health*, 13, 140-150.
- Nederhof, A. (1985). Methods of coping with social desirability bias: a review. *European Journal of Social Psychology*, 15, 263-280.
- Nijman, H., Bjørkly, S., Palmstierna, T., & Almvik, R. (2006). Assessing aggression of psychiatric patients: Methods of measurement and its prevalence. In D. Richter & R. Whittington (Eds.), *Violence in mental health settings: Causes, consequences, management* (pp. 11-23). New York: Springer.
- 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* (pp. 21-59). Chicago: University of Chicago Press.
- Olver, M. E., Kingston, D. A., Nicholai-chuk, T. P., & Wong, S. C. P. (2014). A psychometric examination of treatment change in a multisite sample of treated Canadian federal sexual offenders. *Law and Human Behavior*, 38, 544-559.
- Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology*, 46, 598-609.
- Paulhus, D. L. (1988). *Manual for the Balanced Inventory of Desirable Responding (BIDR-7)*. Toronto/Buffalo: Multi-Health Systems.
- Paulhus, D. L. (2002). Socially desirable responding: The evolution of a construct. In H. I. Braun, D. N. Jackson, & D. E. Wiley (Eds.), *The role of con-*

- structs in psychological and educational measurement* (pp. 49-60). Mahwah, NJ: Erlbaum.
- Salbach-Andrae, H., Klinkowski, N., Lenz, K., & Lehmkuhl, U. (2009). Agreement between youth-reported and parent-reported psychopathology in a referred sample. *European Child & Adolescent Psychiatry, 18*, 136-143.
- Stams, G. J., Brugman, D., Dekovic, M., van Rosmalen, L., van der Laan, P., & Gibbs, J. C. (2006). The moral judgment of juvenile delinquents: A meta-analysis. *Journal of Abnormal Child Psychology, 24*, 697-713.
- van Marle, H. J. C. (2000). Forensic Services in the Netherlands. *International Journal of Law and Psychiatry, 23*, 515-531.
- Van Vugt, E., Gibbs, J., Stams, G. J., Bijleveld, C., Hendriks, J., & van der Laan, P. (2011). Moral development and recidivism: A meta-analysis. *International Journal of Offender Therapy and Comparative Criminology, 55*, 1234-1250.
- Vorst, H. C. M., & Bermond, B. (2001). Validity and reliability of the Bermond-Vorst Alexithymia Questionnaire. *Personality and Individual Differences, 30*, 413-434.
- Walters, G. D. (2006). Risk-appraisal versus self-report in the prediction of criminal justice outcomes: A meta-analysis. *Criminal Justice and Behavior, 33*, 279-304.