

De ontwikkeling van de aangepaste versie van het Sociomorele Reflectie Meetinstrument (SRM-AV)

Ruud H.J. Hornsveld¹, Lotte Vermeulen² & Geert van Veldhuizen²

SAMENVATTING

We ontwikkelden een nieuw instrument voor het meten van moreel besef bij forensisch psychiatrische patiënten, omdat tot voor kort een betrouwbaar en valide instrument ontbrak dat onderzocht was bij een Nederlandse populatie. Daarbij fungeerde de *Sociomoral Reflection Measure-Short Form* als uitgangspunt. Dit instrument omvat 11 stellingen waarvan respondenten moeten aangeven hoe belangrijk ze de in de stelling genoemde kwestie vinden en vervolgens welke argumenten ze voor hun keuze hebben. Aan deze 11 stellingen werden 20 nieuwe stellingen toegevoegd. Het aldus ontstane instrument werd onderzocht op psychometrische eigenschappen bij 80 terbeschikkinggestelden en 40 voetbalsupporters. Om tot een handzaam instrument te komen, verwijderden we van de in totaal 31 items 11 items met een lage interbeoordelaarbetrouwbaarheid. De interbeoordelaar betrouwbaarheid en interne consistentie van het nieuwe instrument met de overgebleven 20 items (SRM-AV) waren goed. De test-hertest betrouwbaarheid over een periode van twee maanden was significant maar relatief laag. Validiteit kon worden aangetoond door betekenisvolle relaties met persoonlijkheidstrekken als altruïsme en consciëntieusheid. De terbeschikkinggestelden scoorden significant lager dan de voetbalsupporters, maar het onderscheid was zodanig klein dat in de klinische praktijk met behulp van de SRM-AV vooralsnog geen conclusies kunnen worden getrokken verschillen in moreel besef tussen terbeschikkinggestelden en normalen. Een omgekeerde relatie tussen psychopathie en moreel besef werd niet gevonden. De SRM-AV lijkt vooralsnog moreel besef bij forensische patiënten betrouwbaar en valide te meten. Verder onderzoek bij grotere groepen patiënten en andere populaties is echter geboden.

1. Klinisch psycholoog/wetenschappelijk onderzoeker, verbonden aan FPC De Kijvelanden te Poortugaal. Correspondentieadres: Dr. R.H.J. Hornsveld, Lange Dreef 52, 2285 LA RIJSWIJK. E-mail: r.hornsveld@tiscali.nl

2. Psycholoog, verbonden aan FPC De Kijvelanden te Poortugaal.

INLEIDING

In Nederlandse instellingen voor forensische psychiatrie zijn de laatste jaren steeds meer cognitief-gedragstherapeutische programma's ontwikkeld en geïmplementeerd. Een voorbeeld is de Agressiehanteringstherapie, die een drietal modules omvat van elk vijfwekelijkse bijeenkomsten, te weten 'Woedebeheersing', 'Sociale vaardigheden' en 'Moreel redeneren' (Hornsveld, Nijman, Hollin & Kraaimaat, 2008). Om de behandelprogramma's te kunnen evalueren ontstond behoefte aan specifiek op probleemgedragingen afgestemde meetinstrumenten, die bovendien in een Nederlands forensisch psychiatrische populatie op psychometrische eigenschappen zijn onderzocht. Voorbeelden van zulke instrumenten, zijn de *Aangepaste Versie van de Picture-Frustration Study* (PFS-AV; Hornsveld, Nijman, Hollin & Kraaimaat, 2007a), de *Observatie Schaal voor Agressief Gedrag* (OSAB; Hornsveld, Nijman, Hollin & Kraaimaat, 2007b), de *Forensic Inpatient Observation Scale* (FIOS; Timmerman, Vastenburg & Emmelkamp, 2001), de FP40 (Brand en Van Emmerik (2001) en de Agressie Vragenlijst (AVL; Meesters, Muris, Bosma, Schouten & Beuving, 1996). De module Moreel Redeneren van de Agressiehanteringstherapie kon echter tot op heden niet geëvalueerd worden omdat een betrouwbaar en valide instrument ontbrak dat in Nederlandse populaties was onderzocht. In dit artikel beschrijven we de ontwikkeling van een dergelijk instrument. We definiëren moraliteit als principes betreffende het verschil tussen goed en kwaad. Met moreel redeneren wordt de rechtvaardiging bedoeld van waarom iets als goed of kwaad moet worden bestempeld. Moraliteit is dus een breder begrip dan altruïsme, het terzijde kunnen zetten van eigenbelang.

Het eerste instrument om het niveau van moreel redeneren te bepalen was het *Moral Judgement Interview* (MJI; Colby & Kohlberg, 1987; Colby, Kohlberg, Spechier, Hewer, Candee, Gibbs & Power, 1987). Dit instrument was gebaseerd op Kohlbergs drie niveaus van morele ontwikkeling ('preconventioneel', 'conventioneel' en 'postconventioneel') met elk twee stadia, en omvatte een aantal hypothetische dilemma's waarop respondenten moesten reageren door een keuze te maken om vervolgens die keuze te beargumenteren. De argumenten werden ten slotte gescoord op de verschillende fasen voor moreel ontwikkeling. De MJI heeft echter als nadeel dat het scoren van de antwoorden ingewikkeld en tijdrovend is. Bovendien kunnen bij een interview respondenten ontvankelijk zijn voor onbedoelde suggestieve beïnvloeding door de interviewer (Carlo, Eisenberg & Knight, 1992).

Gibbs, Basinger en Fuller (1992) meenden dat het presenteren van morele dilemma's niet de enige manier is om morele oordelen uit te lokken. Zij ontwikkelden op basis van de MJI de *Sociomoral Reflection Measure-Short Form* (SRM-SF), een vragenlijst die in plaats van morele dilemma's stellingen omvat (bijvoorbeeld: 'Stel: een vriend van je is in levensgevaar en je bent de enige die hem kan redden'), gevolgd door vragen naar morele oordelen ('Hoe belangrijk

is het voor iemand om het leven van een vriend te redden, zonder daarbij het eigen leven te verliezen') en argumenten voor de gemaakte keuze. Voor de scoring van de 11 items, gebruikten Basinger et al. (1995) alleen de eerste vier stadia van Kohlberg, te weten 1. Unilateraal gezag, 2. Pragmatische oplossing, 3. Basis voor interpersoonlijke relaties en 4. Maatschappelijke betrokkenheid (Vermeulen, 2007). Basinger et al. (1995) onderzochten de psychometrische eigenschappen van de SRM-SF in subgroepen van 165 basisschool leerlingen, 89 highschool leerlingen, 72 universiteitsstudenten, 58 volwassenen en 89 delinquente jongens in een medium-security gevangenis. Vijf beoordelaars vonden bij 25 respondenten (drie tot vijf per subgroep) een interbeoordelaar betrouwbaarheid van .94 tot .99 en test-hertest betrouwbaarheid van .61 tot .78. Cronbach's α bedroeg voor de totale groep van niet-delinquenten .93 ($N = 384$) en voor de groep delinquenten .70 ($N = 89$). Concurrente validiteit kon worden aangetoond door significante correlaties met onder andere het Moral Judgement Interview. Daarnaast scoorden de delinquente jongens naar verwachting significant lager dan mannelijke leeftijdgenoten op de SRM-SF.

Verscheidene auteurs hebben er voor gepleit bij gedetineerden gebruik te maken van 'production instruments' in plaats van 'recognition instruments' (Cavaghan, Arnold & Gibbs, 1983, Hornsveld, Nijman, Hollin & Kraaimaat, 2007a, Stams, Brugman, Dekovic, Rosmalen, Laan & Gibbs, 2006). Bij productie instrumenten worden 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 meer directe informatie over het gedrag van respondenten geven, omdat ze niet hoeven na te denken over wat hun juiste score is op de Likert-schaal, zoals bij 'recognition instruments' het geval is. We besloten daarom uit te gaan van Basinger, Gibbs en Fuller's SRM-SF, maar items toe te voegen die naar onze mening voor forensisch psychiatrische patiënten een meer volledige weergave geven van morele problemen die ze tegen kunnen komen in het dagelijks leven. Door gebruik te maken van een productie instrument verwachten we een significant verschil te vinden in moreel besef tussen een groep terbeschikkinggestelden en een groep min of meer in opleiding vergelijkbare, niet-gedetineerde, mannen (Stams, Brugman, Dekovic, Van Rosmalen, Van der Laan & Gibbs, 2006).

We onderzochten de SRM-SF met toegevoegde items eerst in een groep van 25 terbeschikkinggestelden (onderzoek 1) op interbeoordelaar betrouwbaarheid. Na verwijdering van de onbetrouwbare items, berekenden we de interne consistentie en de test-hertest betrouwbaarheid van het instrument met de overgebleven items. In een vervolgstudie (onderzoek 2) onderzochten we ten behoeve van de validiteit de relatie van scores op het nieuwe instrument met de scores op instrumenten voor het meten van persoonlijkheidstrekken en probleemgedrag.

ONDERZOEK 1: DE CONSTRUCTIE VAN HET NIEUWE INSTRUMENT

De 11 items van de *Sociomoral Reflection Measure-Short Form* werden vertaald en aangevuld met 20 nieuwe items, die werden geformuleerd op grond van de literatuur over moraliteit en op basis van klinische ervaring met terbeschikkinggestelden. De nieuwe items refereerden aan de volgende aspecten van moreel besef: anderen respecteren (bijvoorbeeld: 'Stel: twee joodse jongens komen uit de synagoge en worden uitgescholden. Hoe belangrijk is het dat mensen elkaar niet discrimineren?'), hulpbehoevende personen helpen (bijvoorbeeld: 'Stel: iemand met een handicap durft de straat niet over te steken. Hoe belangrijk is het die persoon met oversteken te helpen?'), trouw blijven (bijvoorbeeld: 'Stel: je merkt dat je vriendin/vrouw vreemd is gegaan. Hoe belangrijk is het trouw te blijven in een relatie?'), anderen aanspreken op hun wangedrag (bijvoorbeeld: 'Stel: je merkt dat een vriend van je in harddrugs dealt. Hoe belangrijk is het dat dealen verboden is?') en nemen van verantwoordelijkheid door ouders (bijvoorbeeld: 'Stel: twee ouders drinken teveel en vragen hun kind naar de slijter te gaan om een krat bier te halen. Hoe belangrijk is het dat ouders niet teveel drinken?'). Net als bij de SRM-SF werd respondenten bij de nieuwe items eerst gevraagd hoe belangrijk zij de stelling vinden die in een item wordt geponeerd, gevolgd door de vraag waarom ze de kwestie (on)belangrijk vinden die in de stelling aan de orde komt.

Voor de scoring van de antwoorden op de belangrijkheid van een stelling werd gebruik gemaakt van een vijfpuntsschaal, lopende van 1 = erg onbelangrijk tot 5 = erg belangrijk. De antwoorden op de vraag naar de argumentatie werden gescoord met behulp van een zevenpuntsschaal. De punten op deze schaal waren omschreven conform de vier fasen en drie overgangsfasen van Gibbs et al. (1992). Bij de fasen 3 en 4 werd echter afgezien van het onderscheid die Gibbs et al. maken in de twee typen moreel besef (Gibbs et al., 1992; Gibbs, 2003). Bizarre of niet ter zake doende antwoorden werden gescoord met een '1', omdat ze blijk geven van een onbegrip voor de (morele) inhoud van de stelling (zie Bijlage 1).

In de pilotstudie werden de interne consistentie, de interbeoordelaar betrouwbaarheid en de test-hertest betrouwbaarheid van de SRM-AV onderzocht. Vervolgens werd voor een eerste indruk van de validiteit ook gekeken naar correlaties met andere instrumenten. De verwachting daarbij was dat een hoge score op de SRM-AV negatief correleert met psychopathie en positief met altruïsme.

METHODE

Deelnemers

Het pilootonderzoek werd uitgevoerd bij 25 mannelijke terbeschikkinggestelden die veroordeeld waren voor een ernstig geweldsdelict. Hun gemiddelde leeftijd was 37.9 jaar ($SD = 11.8$, bereik 19-56 jaar). De patiënten hadden

als hoofddiagnose een cluster B persoonlijkheidsstoornis op as II of een (chronisch) psychotische stoornis op as I in combinatie met een cluster B persoonlijkheidsstoornis op as II (DSM-IV: American Psychiatric Association, 1994). Het psychiatrisch toestandsbeeld van de psychotische patiënten was zodanig gestabiliseerd dat de aanwezige persoonlijkheidsstoornis op de voorgrond was komen te staan. De diagnose was vastgesteld in een multidisciplinair team, gebaseerd op dossiergegevens, klinische rapportage en recent psychiatrisch/psychologisch onderzoek. Ten tijde van het onderzoek werd bij het classificeren van de patiënten volgens de DSM-IV nog geen gebruik gemaakt van specifieke meetinstrumenten voor as I en as II stoornissen.

Meetinstrumenten

Psychopathy Checklist-Revised. (PCL-R; Hare, 1991; Nederlandse versie: Vertommen, Verheul, De Ruiter en Hildebrand, 2002) wordt gebruikt voor het meten van psychopathie. De checklist bestaat uit 20 items die gescoord moeten worden met behulp van een driepuntsschaal met 0 = 'niet van toepassing', 1 = 'enigszins van toepassing' en 2 = 'van toepassing'. Vertommen et al. (2002) vonden ondersteuning voor de betrouwbaarheid en validiteit van de Nederlandse versie van de PCL-R en zij bevestigden Hare's twee-factoren structuur: (1) egoïstisch, ongevoelig en zonder wroeging gebruik maken van anderen (bijvoorbeeld: 'gladde prater/oppervlakkige charme'), en (2) chronisch instabiel en antisociaal gedrag (bijvoorbeeld: 'gedragsproblemen op jonge leeftijd'). In het onderhavige onderzoek gebruiken we de vier-factoren structuur (Hare & Neumann, 2006) met de volgende facetten: Interpersoonlijk (bijvoorbeeld: 'Sterk opgeblazen gevoel van eigenwaarde'), Affectief (bijvoorbeeld: 'Kil/ gebrek aan empathie'), Leefstijl (bijvoorbeeld: 'Impulsiviteit') en Antisociaal (bijvoorbeeld: 'Jeugdcriminaliteit').

NEO Five Factor Inventory. (NEO-FFI; Costa & McCrae, 1992; Nederlandse versie: Hoekstra, Ormel & De Fruyt, 1996) heeft 60 items en meet vijf persoonlijkheidsdomeinen, te weten: Neuroticisme (bijvoorbeeld: 'Ik voel me vaak de mindere van anderen'), Extraversie (bijvoorbeeld: 'Ik houd er van veel mensen om me heen te hebben'), Openheid (bijvoorbeeld: 'Ik houd er niet van mijn tijd met dagdromen te verdoen'), Altruïsme (bijvoorbeeld: 'Sommige mensen vinden mij zelfzuchtig en egoïstisch') en Consciëntieusheid (bijvoorbeeld: 'Ik houd mijn spullen netjes en schoon'). Respondenten geven op een vijfpuntschaal aan in hoeverre ze de uitspraak onderschrijven, lopend van: 1 = helemaal oneens tot 5 = helemaal eens. Een Nederlands onderzoek bij een grote groep normale volwassenen en ouderen wees uit dat de betrouwbaarheid van de NEO-FFI voldoende was en de validiteit goed (Hoekstra; Ormel & De Fruyt, 1996).

Zelf-Analyse Vragenlijst. (ZAV; Van der Ploeg, Defares & Spielberger, 1982) is een Nederlandse vertaling van de *Spielberger State-Trait Anger Scale* (Spielberger, 1980). Van deze Vragenlijst worden de tien 'trait' items gebruikt voor de bepaling van boosheid als dispositie (bijvoorbeeld: 'Ik zou wel iets kapot willen smijten'). Respondenten moeten aangeven hoe ze zich in het algemeen voelen en daarbij gebruik te maken van een vierpuntsschaal: 1 = geheel niet, 2 = een beetje, 3 = tamelijk veel en 4 = zeer veel. In een onderzoek bij een grote groep normale Nederlandse volwassenen (Van der Ploeg, Defares & Spielberger, 1982) bleken interne consistentie, test-hertest betrouwbaarheid en validiteit goed.

Aangepaste versie van de Rosenzweig Picture-Frustration Study. (PFS-AV; Hornsveld, Nijman, Hollin & Kraaimaat, 2007a) wordt gebruikt voor het meten van vijandigheid. De test vraagt respondenten hun reacties op te schrijven op 12 cartoonachtige plaatjes waarop steeds twee personen staan afgebeeld. In elk plaatje maakt de linker persoon een opmerking die door de ander als frustrerend kan worden opgevat. De rechter persoon heeft een lege tekstballon. Proefpersonen worden verzocht de situaties te bekijken (bijvoorbeeld tegen winkelier: 'Dit is al de derde keer dat dit nieuwe horloge stilstaat'), en vervolgens in de lege tekstballon hun eerste reactie op te schrijven. Antwoorden werden gescoord door een onderzoeksassistent die bij zijn indrukken gebruik maakte van een zevenpuntsschaal met de volgende scoringsmogelijkheden: 1 = afwezig, 2 = minimaal, 3 = enigszins, 4 = matig, 5 = sterk, 6 = zeer sterk en 7 = extreem. In een groep van 285 Nederlandse gewelddadige forensisch psychiatrische patiënten (alleen mannen) bleken interne consistentie, test-hertest betrouwbaarheid en interbeoordelaar betrouwbaarheid voldoende tot goed. De validiteit werd aangetoond door significante correlaties met metingen van altruïsme en agressief gedrag (Hornsveld, Nijman, Hollin & Kraaimaat, 2007a).

Agressie Vragenlijst. (AVL; Meesters, Muris, Bosma, Schouten & Beuving, 1996) is een Nederlandse versie van *Buss & Perry's Aggression Questionnaire* (1992). De Vragenlijst meet agressief gedrag en kent 29 items, verdeeld over vier subschalen: Fysieke agressie (bijvoorbeeld: 'Ik heb wel eens iemand die ik ken bedreigd'), Verbale agressie (bijvoorbeeld: 'Als mensen me ergeren, zeg ik soms wel wat ik van ze vind), Woede (bijvoorbeeld: 'Het kost me moeite mijn kalmte te bewaren') en Vijandigheid (bijvoorbeeld: 'Ik wantrouw onbekenden die overdreven vriendelijk tegen me doen'). Respondenten scoren de items op een vijfpuntsschaal van: 1 = helemaal mee oneens tot 5 = helemaal mee eens. Uit een onderzoek bij een grote groep Nederlandse universiteitsstudenten (Meesters, Muris, Bosma, Schouten & Beuving, 1996) bleek de interne consistentie goed. Morren en Meesters (2002) vonden bij een groep Nederlandse adolescente delinquenten (alleen mannen) ondersteuning voor de validiteit.

Inventarisatielijst Omgaan met Anderen. (IOA; Van Dam-Baggen & Kraaimaat, 2000) vraagt respondenten met betrekking tot 35 interpersoonlijke situaties aan te geven hoeveel spanning zij in die situaties zouden ervaren (sociale angst), en hoe vaak ze het beschreven gedrag daadwerkelijk zouden uitvoeren als de situatie zich zou voordoen (sociale vaardigheden). De vijf subschalen, zowel voor sociale angst als voor sociale vaardigheden, zijn: Kritiek geven (bijvoorbeeld: 'Een vriend(in) of goede kennis zeggen dat hij/zij iets doet dat u hindert'), Aandacht vragen voor de eigen mening (bijvoorbeeld: 'Een andere mening naar voren brengen dan die van het gezelschap waarin u bent'), Waardering voor een ander uitspreken (bijvoorbeeld: 'Iemand zeggen dat hij/zij iets goed gedaan heeft'), Initiatief nemen tot contact (bijvoorbeeld: 'Een gesprek beginnen met een aantrekkelijke man of vrouw') en Jezelf waarderen (bijvoorbeeld: 'Instemmen met waardering voor iets dat u hebt gedaan'). Items worden gescoord op een vijfpuntsschaal. Voor sociale angst lopen de scoremogelijkheden van: 1 = helemaal niet gespannen tot 5 = heel erg gespannen, en voor sociale vaardigheden van: 1 = nooit tot 5 = altijd. Interne consistentie en test-hertest betrouwbaarheid bleken goed. Correlaties met andere maten voor angst gaven ondersteuning aan de validiteit (Van Dam-Baggen & Kraaimaat, 2000).

Zelf-Beoordelings Vragenlijst. (Van der Ploeg, Defares & Spielberger, 2000) is een bewerking van de Spielberger State-Trait Anxiety Inventory (Spielberger, Gorsuch & Lushene, 1970) met twee subschalen die angst als toestand en als dispositie meten. In dit onderzoek werd alleen de eerste subschaal gebruikt, die 20 items omvat (bijvoorbeeld: 'Ik ben bang dingen verkeerd te doen'). Items worden gescoord op een driepuntsschaal: 1 = 'bijna nooit', 2 = 'soms' en 3 = 'vaak'. De ZBV is betrouwbaar en valide gebleken (Van der Ploeg, Defares & Spielberger, 2000).

Beck Depressieschaal. (BDI; Bouman, 1994) is een Nederlandse vertaling van de *Beck Depression Inventory* (BDI; Beck & Steer, 1993). Deze zelfrapportage Vragenlijst omvat 21 vier-keuze items waarbij respondenten moeten aangeven welke uitspraak per item het beste weergeeft hoe ze zich de afgelopen week gevoeld hebben. Een voorbeeld van een item is: 'Ik huil niet meer dan normaal' (score = 0), 'Ik huil nu meer dan vroeger' (score = 1), 'Ik huil nu voortdurend' (score = 2), 'Ik kon vroeger wel huilen, maar nu kan ik het niet meer, ook al wil ik het' (score = 3). Betrouwbaarheid en begripsvaliditeit van de BDI bleken voldoende.

Procedure

Tijdens voorlichtingsbijeenkomsten op acht afdelingen met in totaal 118 terbeschikkinggestelden werd het doel van het onderzoek uitgelegd. Benadrukt werd dat wel of niet deelnemen aan het onderzoek geen consequenties had

voor de duur van het verblijf in de instelling. Na afloop van een voorlichtingsbijeenkomst konden de patiënten zich aanmelden.

Gegevens, verkregen uit de in de paragraaf hiervoor besproken meetinstrumenten, werden geanonimiseerd en alleen gebruikt voor het onderzoek. Deelname werd vergoed met € 7. Anderhalve maand na de eerste meting werden de patiënten benaderd voor een tweede meting. Dit vervolgonderzoek werd eveneens vergoed met € 7. Zes patiënten namen niet aan de tweede meting deel. Deze patiënten verschilden niet significant in leeftijd en in psychopathiescore van de 19 patiënten die wel aan de tweede meting deelnamen.

Scoring van de SRM-AV items (en van de PFS-AV items) gebeurde door twee onderzoeksassistenten, waaronder een stagiair Klinische Psychologie van de Erasmus Universiteit Rotterdam (tweede auteur). De onderzoeksassistenten werden gesuperviseerd door een ervaren onderzoeker (eerste auteur). Voor de scores op de *Psychopathy Checklist-Revised* kon gebruik gemaakt worden van een dataset met gegevens van alle in FPC de Kijvelanden opgenomen patiënten.

RESULTATEN

De antwoorden op vragen naar de belangrijkheid van een stelling werden niet verder onderzocht. Die vragen moesten respondenten stimuleren over hun mening betreffende de stelling na te denken alvorens die op te schrijven. Voor het berekenen van de *interbeoordelaar betrouwbaarheid* werden de geschreven reacties op de items van de SRM-AV onafhankelijk van elkaar beoordeeld door de twee onderzoeksassistenten. De interne consistentie (Cronbach's) van de 31 items bedroeg bij een beoordelaar .96 en bij de andere .97. De interbeoordelaarbetrouwbaarheid werd vervolgens voor alle items bij elkaar en voor elk item afzonderlijk berekend door de scores van de twee beoordelaars met elkaar te correleren. De totaalscores leverden een significante correlatie op ($r = .72, p < .01$), maar bij de items afzonderlijk varieerden de correlaties .33 (n.s.) tot van .83 ($p < .01$). De totaalscore en de afzonderlijke itemscores van beide beoordelaars bleken niet significant van elkaar te verschillen.

Ten behoeve van een handzame vragenlijst werden niet meer dan 20 items geselecteerd en werden elf items met een correlatie van .65 of lager verwijderd. Twee van deze elf items waren afkomstig van de oorspronkelijke SRM-SF. Bij een beoordelaar bleek de *interne consistentie* van de 20-item versie .94, bedroeg de *gemiddelde inter-item correlatie* .43 en varieerden de *item-totaal correlaties* van .45 tot .83. Bij de andere beoordelaar bedroeg de *interne consistentie* van de 20-item versie .95, was de *gemiddelde inter-item correlatie* .49 en varieerden de *item-totaal correlaties* van .43 tot .88. De interbeoordelaar betrouwbaarheid nam na het verwijderen van de 11 items toe tot .78 ($p < .01$).

Om de *test-hertest betrouwbaarheid* van de SRM-AV te berekenen werden zowel de scores van de twee beoordelaars bij de eerste meting als die bij de tweede meting gemiddeld, waarna de gemiddelde scores van de twee metin-

gen met elkaar werden gecorreleerd. Er zat een periode van twee maanden tussen de eerste en de tweede meting. De test-hertest betrouwbaarheid van de SRM-AV bleek significant maar relatief laag voor zowel de 31-item versie .59 ($p < .01$) en voor de 20-item versie .54 ($p < .01$).

Voor een eerste indruk van de validiteit van SRM-AV met 20-items werd de totaalscore gecorreleerd met de scores op andere instrumenten voor persoonlijkheidstrekken en probleemgedrag. Zoals in tabel 1 te zien is, bleek de SRM-AV laag, maar positief significant te correleren met de NEO-FFI subschaal Altruïsme ($r = .48, p < .05$) en de ZBV ($r = .45, p < .05$), maar niet significant negatief met psychopathie ($r = .37, n.s.$).

Tabel 1 Correlaties tussen de SRM-AV en andere maten bij terbeschikking-gestelden ($N = 25$).

| Meetinstrument | Inhoud van de schaal | M (SD) | r |
|----------------|----------------------|----------------|------|
| Leeftijd | | 37.96 (11.81) | .30 |
| PCL-R | Psychopathie | 21.82 (7.76) | -.37 |
| | Interpersoonlijk | 3.15 (2.48) | -.12 |
| | Affectief | 6.15 (1.71) | -.23 |
| | Leefstijl | 5.72 (2.67) | -.33 |
| | Antisociaal | 5.17 (2.81) | -.39 |
| NEO-FFI | Neuroticisme | 33.17 (6.88) | .03 |
| | Extraversie | 39.75 (5.44) | .15 |
| | Openheid | 36.75 (5.42) | .32 |
| | Altruïsme | 42.13 (4.58) | .48* |
| | Consciëntieusheid | 45.71 (5.60) | .23 |
| ZAV | Woede | 16.40 (4.56) | .06 |
| PFS-AV | Vijandigheid | 27.26 (5.51) | .02 |
| AVL | Agressie | 74.56 (17.00) | -.01 |
| IOA | Angst | 63.25 (27.08) | .16 |
| | Vaardigheden | 125.26 (19.26) | -.29 |
| ZBV | Angst | 33.04 (7.61) | .45* |
| BDI | Depressie | 13.20 (9.02) | .38 |

* $p < .05$; ** $p < .01$. Noot. PCL-R = Psychopathy Checklist-Revised; NEO-FFI = NEO Five Factor Inventory; ZAV = Zelf-Analyse Vragenlijst; PFS-AV = Aangepaste Versie van de Picture-Frustration Study; AVL = Agressie Vragenlijst; IOA = Inventarisatielijst Omgaan met Anderen; ZBV = Zelf-Beoordelings Vragenlijst; BDI = Beck Depressie Inventarisatielijst.

ONDERZOEK 2: D SRM-AV MET 20 ITEMS

In een vervolgonderzoek werden de psychometrische eigenschappen van de SRM-AV met 20 items (zie Bijlage 2) onderzocht bij een grotere groep terbeschikkinggestelden en bij een groep voetbalsupporters. Voor een vergelijking met relatief laag opgeleide 'normale' mannen moest via de kennissenkring van één van de auteurs een beroep gedaan worden op de supporters van een voetbalclub. Een drietal bedrijven met relatief laag opgeleid personeel, die gevraagd waren om aan het onderzoek mee te doen, weigerden 'omdat het om TBs'ers ging'.

METHODE

Deelnemers

Het vervolgonderzoek werd uitgevoerd bij 80 terbeschikkinggestelden (inclusief de 25 patiënten van de pilootstudie) met een gemiddelde leeftijd van 37.8 jaar ($SD = 9.8$; bereik: 19-62 jaar). De gemiddelde leeftijd van de 40 mannelijke voetbalsupporters was 28.4 jaar ($SD = 9.3$; bereik: 18-62 jaar).

Meetinstrumenten en procedure

De meetinstrumenten die werden gebruikt in de pilootstudie werden ook in het vervolgonderzoek afgenomen, op grond van de ervaringen in de pilootstudie met uitzondering van de IOA en BDI. Alle terbeschikkinggestelden deden op basis van vrijwilligheid mee aan het onderzoek. Deelname werd vergoed met € 7. Bij de voetbalsupporters was het onderzoek van te voren aangekondigd in het clubblad en werden tijdens een wedstrijddag folders met informatie uitgedeeld. Na afloop van het invullen van de vragenlijsten ontvingen de supporters € 10.

Scoring van de meetinstrumenten gebeurde door een stagiair Klinische Psychologie van de Erasmus Universiteit Rotterdam (derde auteur), onder supervisie van een ervaren onderzoeker (eerste auteur). Voor de scores op de *Psychopathy Checklist-Revised* werd opnieuw gebruik gemaakt van een dataset met gegevens van alle in FPC de Kijvelanden opgenomen patiënten.

RESULTATEN

De *interne consistentie* (Cronbach's α) van de SRM-AV met 20 items bleek in de groep terbeschikkinggestelden .89 en in de groep voetbalsupporters .92. Voor het bepalen van de *factorstructuur* werden de data van de patiënten en supporters samen ($N = 120$) geanalyseerd volgens de hoofdassen methode met oblimin rotatie. Dit leverde vijf factoren op die als volgt konden worden geïnterpreteerd: factor 1 = zelf fatsoenlijk omgaan met anderen, factor 2 = anderen aanspreken op hun wangedrag, factor 3 = helpen van hulpbehoevenden, factor

Tabel 2 Gemiddelden, standaarddeviaties en factorladingen van de SRM-AV items ($N = 120$).

| Item | M | SD | Factorladingen | | | | |
|------|------|------|----------------|------------|------------|------------|------------|
| | | | 1 | 2 | 3 | 4 | 5 |
| 1 | 2.67 | 1.30 | .84 | .16 | .14 | .34 | .13 |
| 2 | 2.46 | 1.19 | .69 | .13 | .31 | .54 | .19 |
| 4 | 2.62 | 1.28 | .69 | -.55 | .35 | .36 | .35 |
| 3 | 3.15 | 1.31 | .67 | .11 | .48 | .55 | .29 |
| 20 | 3.15 | 1.31 | .63 | .45 | .24 | .26 | .57 |
| 13 | 2.75 | 1.26 | .62 | -.03 | .53 | .34 | .35 |
| 19 | 2.80 | 1.43 | .55 | .49 | .38 | .43 | .33 |
| 17 | 2.47 | 1.38 | .34 | .69 | .35 | .42 | .12 |
| 16 | 2.60 | 1.48 | .22 | .67 | .38 | .34 | .42 |
| 15 | 3.44 | 1.40 | .31 | .13 | .89 | .28 | .23 |
| 14 | 3.28 | 1.39 | .22 | .28 | .83 | .23 | .28 |
| 11 | 2.52 | 1.31 | .31 | .20 | .37 | .82 | .24 |
| 10 | 2.40 | 1.27 | .41 | .13 | .37 | .80 | -.04 |
| 12 | 1.79 | 1.00 | .48 | .06 | .26 | .71 | .47 |
| 9 | 2.51 | 1.29 | .35 | .19 | .05 | .70 | .37 |
| 7 | 2.08 | 1.22 | .53 | -.24 | .29 | .56 | .45 |
| 8 | 2.42 | 1.44 | .48 | .20 | .22 | .54 | .28 |
| 5 | 2.98 | 1.30 | .44 | -.23 | .44 | .51 | .22 |
| 6 | 3.05 | 1.52 | .22 | .07 | .33 | .30 | .78 |
| 18 | 2.70 | 1.52 | .39 | .42 | .29 | .25 | .65 |

Dikgedrukt = factorladingen groter dan .50. Factor 1 = zelf fatsoenlijk omgaan met anderen, factor 2 = anderen aanspreken op hun wangedrag, factor 3 = helpen van hulpbehoevenden, factor 4 = fatsoenlijk gedrag van anderen verwachten en factor 5 = voorkomen van levensgevaar bij anderen.

4 = fatsoenlijk gedrag van anderen verwachten en factor 5 = voorkomen van levensgevaar bij anderen (zie Tabel 2).

Voor het vaststellen van de *validiteit* werden de totaalscores op de SRM-AV gecorreleerd met de scores op andere meetinstrumenten. In de groep terbeschikkinggestelden werd alleen een significant positieve correlatie gevonden met het NEO-FFI domein Altruïsme. In de groep voetbalsupporters bleek de SRM-AV significant positief te correleren de NEO-FFI domeinen Extraversie, Altruïsme en Consciëntieusheid. Een significant negatieve correlatie werd gevonden met PFS-AV vijandigheid (zie Tabel 3).

Tabel 3 Toetsing van verschillen tussen variabelen en hun correlaties met de SRM-AV bij terbeschikkinggestelden (N = 80) en voetbalsupporters (N = 40).

| Meetinstrument | Inhoud van de schaal | SRM-AV | | | |
|----------------|----------------------|----------------|-----------------|--------------------|-------|
| | | Patiënten | Supporters | | |
| Leeftijd | | M (SD) | Verschillen | | |
| | | | t(119) = 5.03** | | |
| PCL-R | Psychopathie | 37.77 (9.79) | 28.38 (9.31) | .22 | .20 |
| | Interpersoonlijk | 21.82 (7.76) | --- | -.07 | --- |
| | Affectief | 3.15 (2.48) | --- | .15 | --- |
| | Leeftijd | 6.15 (1.71) | --- | -.15 | --- |
| | Antisociaal | 5.72 (2.67) | --- | -.08 | --- |
| NEO-FFI | Neuroticisme | 5.17 (2.81) | --- | -.17 | --- |
| | Extraversie | 32.89 (8.13) | 28.85 (7.00) | F(1.119) = 6.01** | -.22 |
| | Openheid | 40.01 (5.23) | 44.55 (6.29) | F(1.119) = 10.30** | .16 |
| | Altruïsme | 37.35 (5.82) | 37.13 (5.86) | F(1.119) = 1.63 | .14 |
| | Conscientieusheid | 42.06 (5.95) | 40.68 (4.86) | F(1.119) = 10.36** | .31** |
| ZAV | Woede | 45.23 (5.45) | 44.38 (6.59) | F(1.119) = 1.81 | .14 |
| | Vijandigheid | 16.84 (4.76) | 19.15 (6.38) | F(1.119) = 7.26** | -.02 |
| PFS-AV | Agressie | 19.65 (6.91) | 24.28 (7.09) | F(1.119) = 5.37** | -.08 |
| AVL | Angst | 75.64 (17.85) | 77.93 (14.77) | F(1.119) = 4.56* | -.04 |
| IOA | Vaardigheden | 65.31 (21.84) | --- | --- | .10 |
| | Angst | 118.78 (18.43) | --- | --- | -.14 |
| ZBV | Angst | 33.13 (8.31) | 30.98 (7.45) | F(1.119) = 2.91 | .19 |
| BDI | Depressie | 13.06 (9.64) | --- | --- | .14 |

*p < .05; **p < .01. Noot: PCL-R = Psychopathy Checklist-Revised; NEO-FFI = NEO Five Factor Inventory; ZAV = Zelf-Analyse Vragenlijst; PFS-AV = Aangepaste Versie van de Picture-Frustration Study; AVL = Agressie Vragenlijst; IOA = Inventarisatielijst Omgaan met Anderen; ZBV = Zelf-Beoordelings Vragenlijst; BDI = Beck Depressie Inventarisatielijst

Het verschil in SRM-AV score tussen de terbeschikkinggestelden ($M = 52.2$, $SD = 14.9$) en voetbalsupporters ($M = 56.4$, $SD = 16.9$) bleek, gecontroleerd voor leeftijd, significant [$F(2,118)$, 3.71, $p < .05$] maar klein in verhouding tot de standaarddeviatie. Normen voor de twee onderzochte groepen zijn te vinden in bijlage 3.

ALGEMENE DISCUSSIE

We ontwikkelden een instrument voor het meten van moreel besef, omdat tot op heden een betrouwbaar en valide instrument ontbrak dat onderzocht was bij Nederlandse populaties. De 11 items van het oorspronkelijke *Sociomorele Reflectie Meetinstrument* (SRM-SF) werden daartoe aangevuld met 20 nieuwe items. De antwoorden op de in totaal 31 items werden bij 25 forensisch psychiatrische patiënten onafhankelijk van elkaar gescoord door twee onderzoeks-assistenten, die daarbij werden gesuperviseerd door een ervaren onderzoeker. Na verwijdering van 11 items met onvoldoende interbeoordelaar betrouwbaarheid omvatte de SRM-AV 20 items. Van deze 20-item versie bleek de interne consistentie goed, de interbeoordelaar betrouwbaarheid goed en de test-hertest betrouwbaarheid redelijk. Gezien de beperkte test-hertest betrouwbaarheid is daarom een terughoudend gebruik van de SRM-AV voor de evaluatie van de module Moreel redeneren geboden. Met behulp van factoranalyse konden de 20 items als volgt gerubriceerd worden: Zelf fatsoenlijk omgaan met anderen, Anderen aanspreken op hun wangedrag, Helpen van hulpbehoevenden, Fatsoenlijk gedrag van anderen verwachten en Voorkomen van levensgevaar bij anderen. Naar verwachting kon de convergente validiteit van de SRM-AV worden aangetoond door significante positieve correlaties met de persoonlijkheidstrekken altruïsme en consciëntieusheid, en een negatieve correlatie met vijandigheid. De discriminante validiteit bleek uit niet-significante correlaties met agressief gedrag, (sociale) angst en sociale vaardigheden. De validiteit werd verder ondersteund door de bevinding dat de voetbalsupporters significant hoger op de SRM-AV scoorden dan de terbeschikkinggestelden. De bij de terbeschikkinggestelden verwachte negatieve relatie van de SRM-AV met psychopathie werd echter niet gevonden. De verschillen tussen beide groepen waren weliswaar significant maar zodanig klein dat met behulp van de SRM-AV in de klinische praktijk geen conclusies kunnen worden getrokken over verschillen in moreel besef tussen terbeschikkinggestelden en 'normalen', in casu voetbalsupporters.

Het onderzoek naar de psychometrische eigenschappen van de SRM-AV werd gekenmerkt door een aantal beperkingen. Zo namen de terbeschikkinggestelden op basis van vrijwilligheid deel aan het onderzoek. Onduidelijk is daarom of de uitkomst representatief is voor alle opgenomen forensisch psychiatrische patiënten. Bovendien kon nog geen vergelijking worden gemaakt met een groep gewelddadige patiënten op een forensisch psychiatrische

polikliniek. Een dergelijk vervolgonderzoek is dan ook gepland en zal in de nabije toekomst plaats vinden. Een tweede beperking is dat de groep terbeschikkinggestelden vergeleken werd met een groep voetbalsupporters. Drie grote bedrijven met relatief laag opgeleid personeel weigerden aan een onderzoek mee te doen, vanwege de te verwachten overlast of omdat de onderzoekers verbonden waren aan een forensisch psychiatrische instelling. Het aantal onderzochte voetbalsupporters zal overigens in vervolgonderzoek uitgebreid worden met een tweede groep supporters van een andere voetbalclub.

De test-hertest betrouwbaarheid van de SRM-AV was redelijk, maar niet bijzonder goed. Onderzoek van andere meetinstrumenten, zoals de *Agressie Vragenlijst* (Hornsveld, Muris, Kraaimaat, & Meesters, 2009), de *Observation Scale for Aggressive Behavior* (OSAB; Hornsveld, Nijman, Hollin & Kraaimaat, 2007b) en de *Rehabilitation Evaluation Hall And Baker* (REHAB; Nijman, Van Nieuwenhuizen & De Kruyk, 2004) leverden bij terbeschikkinggestelden echter vergelijkbare resultaten op. Mogelijk wordt de teleurstellende test-hertest betrouwbaarheid bij deze instrumenten verklaard door het over het algemeen impulsieve gedrag van terbeschikkinggestelden.

De relevantie van de gevonden factorstructuur dient verder te worden onderzocht. De structuur bleek echter overeen te komen met aspecten van moreel besef die bij het ontwerpen van het nieuwe instrument als uitgangspunt waren gekozen. De factoren vertoonden bovendien overeenkomst met vier van de vijf morele categorieën van Haidt (2007), te weten 'harm/care', 'fairness/reciprocity', 'ingroup/loyalty' en 'authority/respect'. Forensisch psychiatrische patiënten zijn over het algemeen geneigd zich niet met anderen te bemoeien, maar kunnen anderzijds ook sentimenteel zijn als het om situaties gaat waarin de baas spelen niet relevant is. Voor de exploratieve factoranalyse moest echter de totale groep van patiënten en supporters worden gebruikt. We raden daarom aan om de factoranalyse opnieuw uit te voeren bij grotere groepen patiënten en normalen.

Opvallend was dat bij de groep terbeschikkinggestelden er geen omgekeerde relatie werd gevonden tussen scores op de SRM-AV en op psychopathie, gemeten met de PCL-R. Een mogelijke verklaring voor dit resultaat is dat de SRM-AV morele cognities meet en niet moreel handelen. Psychopaten lijken wel te weten wel 'hoe het hoort' (Greene, 2005), maar handelen als het er op aan komt toch in hun eigen, korte termijn belang. Bovendien bleken de antwoorden van sommige patiënten met een hoge psychopathiescore hoog te scoren op de SRM-AV. Bijgevoegde normtabel moet daarom bij individuele patiënten met de nodige voorzichtigheid gebruikt worden. De SRM-AV is in de eerste plaats bedoeld voor de evaluatie van groepen patiënten die een behandeling ter verbetering van moreel besef gevolgd hebben.

De inhoud van de items doet bovendien veronderstellen dat reacties op de gestelde vragen mede bepaald worden door de context waarin het instrument wordt afgenomen. Turiel (1998) benadrukte culturele verschillen in

moreel gedrag. Wellicht verklaart dit waarom van de oorspronkelijke 11 items (Basinger et al., 1992) twee items niet bruikbaar bleken voor de Nederlandse SRV-AV. De items van de SRV-AV zijn naar onze mening dan ook plaats- en tijdgebonden. Over de inhoud van bijvoorbeeld item 19 (Hoe belangrijk is het dat lesbische vrouwen niet gediscrimineerd worden?) werd in het onderhavige onderzoek over het algemeen tolerant geoordeeld, maar het is de vraag of dit ook geldt voor andere Westerse landen (bijvoorbeeld: de Verenigde Staten) en of die tolerantie over 10 jaar in Nederland nog aanwezig is. Een dergelijk item kan dan een minder betrouwbare of valide bijdrage gaan leveren aan de totaalscore van de SRV-AV.

Samengevat lijkt de SRM-AV een veelbelovend instrument voor het meten van moreel besef bij terbeschikkinggestelden. Een eerste evaluatie wijst uit dat de psychometrische eigenschappen voldoende tot goed zijn, doch verder onderzoek is geboden, met name naar de test-hertest betrouwbaarheid en naar het onderscheidend vermogen van de SRM-AV tussen forensisch psychiatrische patiënten en normalen.

SUMMARY

The development of the Adapted Version of the Sociomoral Reflection Measure (SRM-AV)

We have developed a new instrument for the measurement of moral awareness in forensic psychiatric patients, since a reliable and valid instrument investigated in a Dutch population has been lacking up to now. The original *Sociomoral Reflection Measure-Short Form* functioned as a starting point. This instrument comprises eleven statements of which respondents have to indicate how important they consider the issue which is mentioned in each statement, and then to write down the arguments for their choice. Twenty new statements were added to these eleven original statements. The psychometric properties of the extended measure were studied in a sample of 80 violent forensic psychiatric inpatients and 40 supporters of a soccer club. In order to realize a manageable instrument, we removed 11 from the 31 items with the lowest interrater reliability. Interrater reliability and internal consistency of the new instrument with the remaining 20 items (SRM-AV) were good. The test-retest reliability over a period of two months was significant but relatively low. Validity could be demonstrated by meaningful relations with personality traits such as agreeableness and conscientiousness. The patients had significantly lower scores than the supporters, but the difference in scores was so small that in clinical practice the SRM-AV as yet cannot be used for discriminating moral awareness between forensic psychiatric patients and normals. A reversed relation between psychopathy and moral awareness was not found. As yet, the SRM-AV seems to measure moral aware-

ness reliably and valid in violent forensic psychiatric inpatients. However, further investigation is advised in other populations.

Key words: SRM-AV; forensic psychiatry; moral awareness; psychometric properties.

LITERATUUR

- American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Association.
- Basinger, K.S., Gibbs, J.C., & Fuller, D. (1995). Context and the measurement of moral judgement. *International Journal of Behavioral Development*, 18, 537-556.
- Beck, A.T., & Steer, R.A. (1993). *Manual for the Beck Depression Inventory*. San Antonio, TX: Psychological Corporation.
- Bouman, T.K. (1994). De 'Beck Depression Inventory' (BDI). *Gedragstherapie*, 27, 69-71.
- Brand, E.F.J.M., & Emmerik, J.L. van (2001). *Handboek Forensisch Psychiatrische Profielen: Handleiding FP40*. Den Haag: Ministerie van Justitie.
- Buss, A.H., & Perry, M. (1992). The aggression questionnaire. *Journal of Personality and Social Psychology*, 63, 452-459.
- Carlo, G., Eisenberg, N., & Knight, G.P. (1992). An objective measure of adolescents' prosocial moral reasoning. *Journal of Research on Adolescence*, 2, 331-349.
- Colby, A., & Kohlberg, L. (1987). *The measurement of moral judgement* (Vol.1). Cambridge, UK: Cambridge University Press.
- Colby, A., Kohlberg, L., Spechier, B., Hewer, A., Candee, D., Gibbs, J., & Power, C. (1987). *The measurement of moral judgement* (Vol. 2). Cambridge, UK: Cambridge University Press.
- Costa, P.T. Jr., & McCrae, R.R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and the Five Factor Inventory (NEO-FFI): Professional manual*. Odessa, FL: Psychological Assessment Resources Inc.
- Dam-Baggen, C.M.J. van, & Kraaimaat, F.W. (2000). *Inventarisatielijst Omgaan met Anderen (IOA). Handleiding (2^e geheel herziene druk)*. Lisse: Swets Test Services.
- 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.
- Gibbs, J.C. (2003). *Moral development & reality: Beyond the theories of Kohlberg and Hoffman*. Thousand Oaks, CA: Sage.
- Gibbs, J.C., Basinger, K.S., & Fuller, D. (1992). *Moral maturity: Measuring the development of sociomoral reflection*. New Jersey: Erlbaum.
- Greene, J.D. (2005). Cognitive neuroscience and the structure of the moral mind. In P. Laurence, S. Carruthers, & S. Stich (Eds.), *The innate mind: Structure and contents*. New York: Oxford University Press.
- Haidt, J. (2007). The New Synthesis in Moral Psychology. *Science*, 316, 998-1002.
- Hare, R.D. (1991). *The Hare Psychopathy Checklist-Revised*. Toronto, CA: Multi-Health Systems.
- Hare, R.D., & Neumann, C.S. (2006). The PCL-R assessment of psychopathy: Development, structural properties, and new directions. In C. J. Patrick (Ed.), *Handbook of psychopathy* (pp. 58-88). New York: Guilford.
- Hoekstra, H.A., Ormel, J., & Fruyt, F. de (1996). *Handleiding NEO Persoonlijkheidsvragenlijsten*. Lisse: Swets Test Services.
- 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. (2007a). 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., Nijman, H.L.I., Hollin, C.R., & Kraaimaat, F.W. (2008). Aggression Control Therapy for violent forensic psychiatric patients: Method and clinical practice. *International Journal of Offender Therapy and Comparative Criminology*, *52*, 222-233. DOI: 10.1177/0306624X07303876.
- Hornsveld, R.H.J., Nijman, H.L.I., Hollin, C.R., & Kraaimaat, F.W. (2007b). Development of the Observation Scale for Aggressive Behavior (OSAB) for Dutch forensic psychiatric inpatients with an antisocial personality disorder. *International Journal of Law and Psychiatry*, *30*, 480-491.
- Kohlberg, L. (1984). Moral stages and moralization: The cognitive-developmental approach. In L. Kohlberg (Ed.), *Essays on moral development: The psychology of moral development* (pp. 170-205). San Francisco: Harper & Row.
- 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.
- Nijman, H.L.I., van Nieuwenhuizen, Ch. Van, & Kruyk, C. de (2004). Behavioral changes during forensic psychiatric (TBS) treatment in the Netherlands. *International Journal of Law and Psychiatry*, *27*, 79-85.
- Ploeg, H.M. van der, Defares, P.B., & Spielberger, C.D. (1982). *Handleiding bij de Zelf-Analyse Vragenlijst, ZAV*. Lisse: Swets & Zeitlinger.
- Ploeg, H.M. van der, Defares, P.B., & Spielberger, C.D. (2000). *Handleiding bij de Zelf Beoordelings Vragenlijst, ZBV*. Lisse: Swets & Zeitlinger.
- Rosenzweig, S. (1978). *Aggressive behavior and the Rosenzweig Picture-Frustration Study*. New York: Praeger.
- Spielberger, C.D. (1980). *Preliminary manual for the State-Trait Anger Scale (STAS)*. Tampa, FL: University of South Florida.
- Spielberger, C.D., Gorsuch, R.L., & Lushene, R.E. (1970). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.
- Stams, G.J., Brugman, D., Dekovic, M., Rosmalen, L. van, Laan, P. van der, & Gibbs, J.C. (2006). The moral judgment of juvenile delinquents: A meta-analysis. *Journal of Abnormal Child Psychology*, *24*, 697-713.
- Timmerman, I.G.H., Vastenburg, N.C., & Emmelkamp, P.M.G. (2001). The Forensic Inpatient Observation Scale (FIOS): Development and validity. *Criminal Behavior and Mental Health*, *11*, 144-162.
- Turiel, E. (1998). The development of morality. In W. Damon (Ed.), *Handbook of child psychology*, fifth edition, volume 3 (pp. 863-932). New York: Wiley.
- Vermeulen, L. (2007). *Moreel redeneren en de ontwikkeling van een nieuwe vragenlijst: De Sociomorele Reflectie Vragenlijst-Aangepaste Versie*. Rotterdam, Erasmus Universiteit Rotterdam (masterscriptie).
- Vertommen, H., Verheul, R., Ruiters, C. de, & Hildebrand, M. (2002). *De herziene versie van Hare's Psychopathie Checklist (PCL-R) [The revised version of Hare's Psychopathy Checklist (PCL-R)]*. Lisse, NL: Swets Test Publishers.

Bijlage 1. Scoring van de aangepaste versie van het sociomorele reflectie meetinstrument

(SRM-AV).

Onrijp niveau

Fase 1: unilateraal en fysicistisch; score: 1

- Rechtvaardiging vanwege een autoriteitsfiguur
- Rechtvaardiging vanwege de meest kenmerkende rol van de betreffende persoon of personen
- Stellige beweringen in absolute termen
- Zwart-wit redeneringen
- Rechtvaardiging vanwege straf

Overgangsfase 1-2; score: 2

Fase 2: Uitwisselend en instrumenteel, moraliteit door interactie met anderen; score: 3

- Quid Pro Quo
- Volstreekte gelijkheid of ongelijkheid in relaties
- Hantering van verworven vrijheden als concrete rechten
- Rechtvaardiging op grond van eigen voorkeuren
- Beroep op pragmatische behoeften
- Rekenen op praktische voordelen of verplichtingen in de toekomst

Overgangsfase 2-3; score: 4

Rijp niveau

Fase 3: Wederzijds en prosociaal, score: 5

- Besef van de betekenis van interpersoonlijke relaties
- Sterke betrokkenheid met het emotionele welbevinden van anderen
- Besef van consequenties bij schending van normatieve verwachtingen
- Besef van prosociale bedoelingen als kenmerk van normaal sociaal gedrag
- Normatieve en prosociale waarden voor relaties in het algemeen
- Hanteren van waarden die refereren aan het geweten

Overgangsfase 3-4; score: 6

Fase 4: Systematisch en standaard, score: 7

- Beroep op waarde als vereiste van het functioneren van de maatschappij of van maatschappelijke instellingen
 - Beroep op basale maatschappelijke rechten en waarden
 - Beroep op overwegingen als reputatie en integriteit
 - Rechtvaardiging van normatieve waarden omdat alternatieven schadelijk zijn voor de maatschappij
 - Rechten en plichten die de maatschappij het individu verschuldigd is
 - Maatschappelijke normen ten aanzien van het individuele geweten
-

Bijlage 2. De 20 Items van de SRM-AV1

1. Hoe belangrijk is het dat je beloften aan je vrienden nakomt?
2. Wat vind je van het nakomen van beloften aan mensen in het algemeen? Hoe belangrijk is het dat mensen beloften nakomen aan iemand die ze nauwelijks kennen?
3. Wat vind je van het nakomen van een belofte aan kinderen? Hoe belangrijk is het dat ouders beloften aan hun kinderen nakomen?
4. Hoe belangrijk is het dat mensen in het algemeen eerlijk zijn?
5. Hoe belangrijk is dat kinderen hun ouder(s) soms helpen?
6. Hoe belangrijk is het dat iemand blijft leven, zelfs als die persoon dat zelf niet wil?
7. Hoe belangrijk is het dat mensen niet stelen?
8. Hoe belangrijk is het dat men de wet gehoorzaamt?
9. Hoe belangrijk is het dat rechters mensen die de wet overtreden een straf opleggen?
10. Hoe belangrijk vind je het dat anderen je met respect behandelen?
11. Stel dat twee kinderen een ander kind pesten. Hoe belangrijk is het ervoor te zorgen dat die kinderen leren respect voor elkaar te hebben?
12. Hoe belangrijk vind je het dat anderen niet over je roddelen?
13. Hoe belangrijk is het dat mensen hun mening op een directe manier geven?
14. Stel dat iemand met een handicap de straat niet over durft te steken. Hoe belangrijk is het die persoon met oversteken te helpen?
15. Hoe belangrijk is het mensen met een lichamelijke handicap te helpen als dat nodig is?
16. Stel dat je vriend zijn vriendin in het bijzijn van anderen uitscheldt voor 'kutwijf'. Hoe belangrijk is het dat die anderen hem daarop aanspreken?
17. Stel dat twee mannen een andere man mishandelen terwijl omstanders toekijken. Hoe belangrijk is het dat de omstanders die twee mannen aanspreken op hun gedrag?
18. Stel: je merkt dat een vriend van je in harddrugs dealt. Hoe belangrijk is het dat dealen verboden is?
19. Stel: twee lesbische vrouwen zoenen elkaar op straat. Hoe belangrijk is het dat lesbische vrouwen niet gediscrimineerd worden?
20. Stel dat een kind dagelijks geconfronteerd wordt met ruziemakende ouders. Hoe belangrijk is het dat ouders rekening met hun kinderen houden door hen niet bij die ruzies te betrekken?

Bijlage 3. Voorlopige normen SRM-AV voor mannelijke terbeschikkinggestelden en voetbalsupporters.

| Ruwe score | Terbeschikkinggestelden (N = 80) | Voetbalsupporters (N = 40) | Ruwe score |
|------------|----------------------------------|----------------------------|------------|
| 29 | 1 | 1 | 29 |
| 31 | 1 | 1 | 31 |
| 33 | 1 | 1 | 33 |
| 35 | 2 | 1 | 35 |
| 37 | 2 | 2 | 37 |
| 39 | 2 | 2 | 39 |
| 41 | 3 | 3 | 41 |
| 43 | 3 | 3 | 43 |
| 45 | 4 | 4 | 45 |
| 47 | 5 | 4 | 47 |
| 49 | 5 | 4 | 49 |
| 51 | 6 | 5 | 51 |
| 53 | 6 | 5 | 53 |
| 55 | 7 | 5 | 55 |
| 57 | 7 | 6 | 57 |
| 59 | 7 | 7 | 59 |
| 61 | 8 | 7 | 61 |
| 63 | 8 | 7 | 63 |
| 65 | 9 | 8 | 65 |
| 67 | 9 | 8 | 67 |
| 69 | 9 | 8 | 69 |
| 71 | 9 | 8 | 71 |
| 73 | 9 | 9 | 73 |
| 75 | 9 | 9 | 75 |
| 77 | 10 | 9 | 77 |
| 79 | 10 | 9 | 79 |
| 81 | 10 | 9 | 81 |
| 83 | 10 | 10 | 83 |
| 85 | 10 | 10 | 85 |
| 87 | 10 | 10 | 87 |
| 89 | 10 | 10 | 89 |
| 91 | 10 | 10 | 91 |
| 93 | 10 | 10 | 93 |