

**Mogelijkheden en beperkingen van de voor-
spelling van gevaar ten behoeve van het al
of niet beëindigen van de tbs-behandeling**

Een literatuuranalyse

Ed. Leuw

april 1998



**Wetenschappelijk
Onderzoek- en
Documentatiecentrum**

Mogelijkheden en beperkingen van de voorspelling van gevaar ten behoeve van het al of niet beëindigen van de tbs-behandeling .

Een literatuur-analyse

Ed. Leuw

1. Inleiding

* 1) Voorspelling van het gevaar van (sexuele) geweldspleging is bij uitstek relevant voor de Nederlandse tbs-situatie. Zowel aan het begin, als aan het eind van de tbs-behandeling speelt de inschatting van het risico dat de veroordeelde opnieuw een ' tbs-achtig' delict zal plegen een cruciale rol. 2) Aan het begin is het één van de belangrijkste argumenten voor opname in de tbs; aan het eind vormt de inschatting van het gevaar van (sexuele) geweldsrecidive wellicht het doorslaggevend argument om al of niet de (intramurale) behandeling te beëindigen. De voorspelling van de kans op recidive is een reeds lang bestaand en betrekkelijk veelbestudeerd onderwerp binnen de criminologie. In de context van de Nederlandse tbs is echter sprake van een zeer specifiek geval van voorspelling van (gewelds)recidive. Het gaat immers om een betrekkelijk homogene, sterk uitgeselecteerde populatie waarbij de voorspelling betrekking heeft op enkele nauw gedefinieerde (strafbare) gedragingen. Voor de casus van de voorspelling van gevaar na de behandeling in de tbs kunnen we dit puntsgewijs als volgt uitdrukken:

- a. een specifiek voorafgaand misdrijf , i.c. (sexueel) geweld
- b. een specifieke psychiatrische populatie: 'geestelijk gestoorden' of 'in hun ontwikkeling' gestoorden
- c. een voorafgaande aan de behandeling hoog ingeschatte recidivekans
- d. een betrekkelijk lange voorafgaande periode van intensieve forensisch psychiatrische behandeling
- e. de aanwezigheid van veel relevante en gedetailleerde klinische kennis over de tbs-gestelde
- f. de voorspelling dient te leiden tot een beslissing met een beperkt aantal mogelijkheden: al of niet voortzetten van de (intramurale) behandeling
- g. de voorspelling heeft betrekking op een zeer specifieke vorm van recidive (i.c. recidive met ernstig -sexueel- geweld)

Gezien het feit dat de Nederlandse tbs internationaal gezien een unieke maatregel is valt op voorhand niet te verwachten dat de literatuur over de voorspelling van recidivegevaar veel informatie zal opleveren die direct relevant is voor de Nederlandse tbs-situatie. Na kennisname van de beschikbare literatuur over de voorspelling van gevaar blijkt dit inderdaad niet het geval te zijn. Het grote merendeel van het onderzoek naar voorspelling van recidive blijkt te zijn gebaseerd op in vergelijking tot de tbs veel minder homogene populaties en is veel meer gericht op het voorspellen van ruimere categorieën van delinquent gedrag. Dit betekent primair dat het bestaande predictie-onderzoek over het algemeen een beperkte relevantie heeft binnen de Nederlandse tbs-context.

Anderzijds mogen we er vanuit gaan dat het meer algemene thema van het vaststellen en voorspellen van recidive-risico's bij personen die eerder een delict hebben gepleegd ten dele ook relevant zal zijn voor zeer specifieke gevallen zoals hierboven omschreven. Zo stelt Bonta c.s. (1996) op basis van een meta-analyse van onderzoek onder gestoorde delinquenten dat "The major predictors of general and violent recidivism appear comparable for mentally disordered and nondisordered offenders." Mede om deze reden is het dus zinvol om na te gaan

2) welke aanwijzingen er in de criminologische en forensisch psychiatrische literatuur zijn te vinden voor de optimalisering van de voorspelling van recidive. Met name zal worden gezien in hoeverre deze opdracht met behulp van meetinstrumenten kan worden vervuld.

Alvorens over te gaan tot de centrale vraag naar de meest veelbelovende *methoden* van voorspelling eerst nog een enkel woord over de *inhoud* van hetgeen wordt voorspeld. Hoe kan het in de voorspelling aan de orde zijnde risico, in onderzoekstermen de criterium-variabele, worden omschreven? Er bestaat in de literatuur grote overeenstemming over het inzicht dat het voorspelde 'gevaar' niet moet worden begrepen als wel of niet criterium. "... 'risk level' should be seen as a continuous probability statement, rather than a dichotomous variable (risk or no risk), and as risk levels fluctuate over time and context, assessment of risk should be in the form of ongoing assessments, rather than on-off predictions." ()

Zoals boven al aangegeven brengt de tbs met zich mee dat de specifieke kans op het opnieuw vóórkomen van ernstige (sexuele) geweldscriminaliteit dient te worden voorspeld. Of ex-tbsgestelden zich na ontslag in alle opzichten keurig aan de wet zullen houden is voor beëindigingsbeslissing op zichzelf niet relevant.

Los van de specifieke te voorspellen gedragsvorm dient ook te worden onderscheiden tussen de *waarschijnlijkheid* dat het risico wordt gerealiseerd en mate van schade die zal ontstaan *indien* het risico wordt gerealiseerd. Logisch gezien zijn de factoren 'kans' en 'ernst' onafhankelijk van elkaar. Uit de literatuur blijken ook empirische bevindingen dat de kans op recidive en de ernst van recidive als onderscheiden te voorspellen grootheden moeten worden gezien. In de literatuur wordt zowel de waarschijnlijkheid als de ernst van het gevaar gezien als een product van (antecedente) kenmerken van de persoon en de faciliterende of inhiherende contextuele condities. Mulvey c.s. (1995) stelt dat het niet gaat om de voorspelling van 'gevaar' in algemene zin. "We instead need to know when and under what conditions violence might occur, and who needs to be protected."

2. Klinische en op test gebaseerde (statistische) voorspelling van recidive

Er worden aan de hand van de methode van voorspelling algemeen twee soorten van inschatting van het recidive-risico onderscheiden, de klinische en de statistische, op test gebaseerde¹ voorspelling. Bij de klinische voorspelling wordt door (bijvoorbeeld psychiatrische/psychologische) deskundigen een uitspraak gedaan over de waarschijnlijkheid van het vóórkomen van recidive in de toekomst. De klinische voorspelling behelst altijd een uitspraak over één bepaald individu en het gaat per definitie om het *subjectieve*, maar tegelijkertijd professionele oordeel. De klinische voorspelling wordt om deze reden ook wel getypeerd als een *intuïtief* oordeel.

Er is enig onderzoek verricht naar de elementen van klinische voorspelling. In twee gevallen heeft dit betrekking op een situatie die direct vergelijkbaar is met de Nederlandse tbs. In een factor-analytische studie werd het klinische oordeel over noodzakelijk voortgezet verblijf in de forensisch psychiatrische inrichting voorspeld. Het blijkt dat deze beslissing vooral is gebaseerd op relationele aspecten (patient is 'eerlijk', houdt rekening met anderen, niet-conflictueus etc.) en op emotionele stabiliteit. (Eucker c.s., 1994) In een ander Duits onderzoek in een forensisch psychiatrische kliniek blijkt dat de klinische beslissing met betrekking tot het beëindigen van de behandeling en de modaliteit van behandeling betrekkelijk goed kon worden voorspeld met een meetinstrument voor 'psychopathie'. Uit beide gevallen blijkt dat bij de klinische beslissing kennelijk het (intuïtieve) oordeel over de kenmerken van het sociale gedrag van de patiënt zwaar weegt. Green (1997) beschrijft het proces van de klinische besluitvorming binnen de behandelingscontext als een voortschrijdend proces waarin: "New realities emerge from negotiations between staff members. Diagnosis, therapy and recovery occur in a dialectical process constantly being reconstructed as it is informed by interpretations of everyday events." In navolging van Meehl (1954, 1986) definieert Blackburn (1993) klinische voorspelling eenvoudig als de 'subjectieve evaluation of risk'. Gendreau c.s. (1996) typeren klinische voorspelling nader als volgt: "... clinical assessment consists of an 'expert' (e.g. probation officer) interviewing an offender for the risk to reoffend. The judgement is primarily based on the 'experts' intuition, wisdom, and experience. File information is usually consulted as part of the decision making process."

Deze omschrijving lijkt onvoldoende recht te doen aan het beslissingsproces aan het eind van de tbs-behandeling, dat leidt tot het oordeel over het al of niet acceptabel zijn van het risico dat de tbs-gestelde opnieuw ernstig (sexueel) geweld zal plegen. In het geval van de klinische voorspelling bij het advies van al of niet verlenging van de tbs bestaat immers een zeer uitgebreide, multidisciplinaire en langs verscheidene wegen (bijvoorbeeld gedragsobservaties, psychiatrisch/psychologisch onderzoek, milieu onderzoek, bevindingen van verlofsituaties etc.) verworven informatie. Het lijkt daarom zinvol om bij vergelijking van klinische met statistische voorspelling de 'kwaliteit' van het klinische oordeel te betrekken. Er zijn aanwijzingen dat met name informatie ontleend aan observaties van het sociaal functioneren de effectiviteit van het klinische oordeel sterk bevordert. (Clark c.s., 1993) Met name bij de beslissingen aan het eind van de tbs-behandeling zal veel gedetailleerde informatie beschikbaar zijn over het sociale gedrag van de terbeschikkinggestelde. In de vergelijkingen van de effectiviteit van klinische versus statistische voorspelling wordt doorgaans geen rekening gehouden met de omvang en de kwaliteit van de *information base* waarop de klinische voorspelling berust.

Het laat zich aanzien dat de beslissing tot beëindiging van de (intramurale) tbs-behandeling zonder meer als een geval van klinische voorspelling kan worden getypeerd. De binnen het Ministerie van Justitie en het Trimbos-Instituut beschikbare literatuur leverde geen enkele

¹ In het Engels wordt het begrip *actuarial prediction* gebruikt.

aanwijzing op dat binnen de context van de Nederlandse tbs-maatregel meetinstrumenten een rol spelen bij het bereiken van de beslissing. Ook navraag bij enkele deskundigen leidde tot de zelfde conclusie.

Anders dan de klinische voorspelling is de statistische voorspelling niet gericht op het oordeel over een individuele persoon. De op test gebaseerde voorspelling doet een uitspraak over de kans dat een persoon, voorzover die behoort tot een groep van personen met een aantal gemeenschappelijke kenmerken, het voorspelde gedrag zal vertonen.

Naast het groepsgeoriënteerde karakter van de statistische voorspelling vormt het kenmerk *objectiviteit* het belangrijkste verschil met de klinische voorspelling. Zoals gezegd kan de laatste voorspelling worden beschreven als een individueel oordeel van één persoon over een andere persoon. De statistische voorspelling is daarentegen in principe onafhankelijk van degene die het oordeel vaststelt. In wezen is het ook niet een voorspelling over een 'persoon' maar over de abstracte drager van een aantal geoperationaliseerde kenmerken. De voorspelling van gevaar vloeit voort uit de 'som' van een aantal met behulp van meetinstrumenten vastgetelde kenmerken (predictie variabelen) en behelst een statistische kans op recidive die behoort bij de score zoals met de meetinstrumenten vastgesteld.

3. Effectiviteit van klinische versus op test gebaseerde voorspelling

Er bestaat in de criminologische en forensisch psychiatrische literatuur een sterke consensus dat over het algemeen de op test gebaseerde voorspelling veel effectiever is dan de klinische voorspelling. Gendreau c.s. (1996) heeft dit inzicht als volgt samengevat: "In 1954 Paul Meehl published a seminal book comparing the two methods. He examined 20 studies; the actuarial approach was superior in 19 instances. Subsequent reviews of larger samples of studies have confirmed Meehl's original conclusions." Ook recente vergelijkende studies van klinische en op test gebaseerde voorspelling lijken deze conclusie te bevestigen. In het zelfde artikel wordt over een in 1994 gepubliceerde meta-analyse vermeld dat de statistische voorspelling in drie kwart van de gevallen tot betere resultaten leidde dan de klinische voorspelling. De gemiddelde correlatie met het criterium van de klinische voorspelling was .08, terwijl de statistische voorspelling een correlatie van .22 opleverde. Beide correlaties zijn significant. Zowel klinische als statistische voorspellingen doen het beter dan wanneer louter op toeval zou worden voorspeld. We kunnen daaraan toevoegen dat voorspellingen met correlaties van deze orde weinig zullen kunnen bijdragen aan beslissingen op *individueel* niveau.

Op grond van de literatuur kan worden geconcludeerd dat klinische voorspelling op basis van eenmalige beoordeling (*assessment*) zeer zwak is. In dit geval doet op test gebaseerde voorspelling het zonder meer beter. Het gaat dan bijvoorbeeld om gevallen van *parole decisions* waarbij ofwel op basis van risico-inschatting door middel van tests of door middel van het oordeel van een deskundige een beslissing wordt genomen over voorwaardelijke in vrijheid stelling. Anderzijds zijn er aanwijzingen dat klinische voorspelling op basis van meer langdurige (en intensieve) ervaring met personen veel sterker is. Clark c.s. (1993) geeft aan dat onderwijzers betrekkelijk effectief problematisch gedrag van hun leerlingen kunnen voorspellen. En daarnaast werd ook geconstateerd dat "Several studies of juveniles found staff ratings of parole prognosis and of prison adjustment to be the best predictor of future recidivism." De laatst genoemde gevallen tonen overeenkomst met het type klinische voorspelling dat in de tbs-situatie zal worden gemaakt.

Het onderscheid dat in de literatuur wordt gemaakt tussen klinische voorspelling en statistische voorspelling impliceert allerm minst dat er in de praktijk een keuze moet worden gemaakt voor de ene of de andere procedure. Veel auteurs hebben gewezen op een voor de hand liggende beslissingsstrategie, waarbij klinische en op test gebaseerde voorspelling wordt gecombineerd. Dit lijkt vooral in aanmerking te komen indien, zoals bij de tbs-beoordeling het geval is, zwaarwegende en sterk op de individuele persoon toegespitste beslissingen moeten worden genomen. In de praktijk zou dit betekenen dat statistische voorspellers mede het uiteindelijke klinische oordeel bepalen. Maar het is natuurlijk ook denkbaar dat het klinische oordeel gaat fungeren als één van de elementen waarop de statistisch bepaalde beslissing is gebaseerd. Zo pleit Black (1993 p. 335) voor een geobjectiveerde beslissingsprocedure waarbij klinische beoordelingen mede een rol spelen. "...the use of decision aids [actuarial methods] which make the process of decision making more explicit, is a step in this direction. As .. demonstrated the most effective combination of actuarial and clinical methods .. is the objective reporting of clinical judgements which can be used statistically."

4. Beslistheoretische aspecten van op test gebaseerde voorspelling:

Voor de beoordeling van de mogelijke betekenis van meetinstrumenten ten behoeve van de verbetering van de beslissing over het al of niet beëindigen van de (intramurale) tbs-behandeling geldt een cruciale voorafgaande vraag. Wat is in theorie de voorspellende kracht van statistische methoden, in het bijzondere geval van de voorspellingen binnen de tbs-context? In de inleiding zijn de specifieke kenmerken van de voor de tbs relevante voorspelling genoemd. In essentie gaat het om een type voorspelling waarbij binnen een *homogene* populatie een uitspraak moet worden gedaan over een sterk *specifiek* criterium, dat in de praktijk betrekkelijk *zeldzaam* wordt aangetroffen. Deze voor de voorspelling van ernstige (sexuele) geweldsrecidive typerende kenmerken leggen sterke beperkingen op aan wat in theorie kan worden verwacht van de effectiviteit van voorspellende meetinstrumenten. Wat betreft de specificiteit van het criterium behoeft het nauwelijks commentaar dat voorspellen makkelijker is naarmate het te voorspellen criterium ruimer wordt gedefinieerd. Uit literatuur waarbij de voorspelling van algemene recidive wordt vergeleken geweldsrecidive blijkt ook in feite dat de samenhang tussen voorspeller en criterium in het laatste geval doorgaans aanmerkelijk zwakker is. Het onderzoek van Bonta c.s. (1996a) onder een grote steekproef van 3267 delinquenten is hiervan een duidelijk voorbeeld. Het gebruikte meetinstrument, de Statistical Information on Recidivism (SIR) correleerde .42 met algemene recidive, terwijl de correlatie met ernstige geweldsrecidive slechts .15 bedroeg. In alledaagse woorden betekent dit dat het meetinstrument betrekkelijk goed algemene recidive kan voorspellen, maar vrijwel niet geweldsrecidive. Uit een meta-analyse van de zelfde auteur (1996c) op basis van 64 predictie-onderzoeken blijkt dit een algemeen geldig resultaat te zijn.

Naarmate een criterium nauwer wordt gedefinieerd zal het zich ook minder voordoen. De frequentie waarmee het te voorspellen gedrag (het criterium) wordt aangetroffen, de *base rate*, is sterk bepalend voor de theoretisch te verwachten effectiviteit van een meetinstrument. De *base rate* kan worden begrepen als de toevalskans dat een bepaald verschijnsel, in een bepaalde populatie wordt aangetroffen. Zo is bijvoorbeeld uit algemeen recidive-onderzoek bekend dat rond de 50% van de veroordeelden opnieuw voor 'enig delict' veroordeeld zal worden. Op basis van toeval zou dus bij voorspelling van recidive de helft goed en de helft fout worden voorspeld. Een voorspelling kan meer effectief worden genoemd naarmate er *beter* wordt voorspeld dan op basis van het toeval mogelijk is. Uit de beslistheorie is bekend dat de mogelijkheid van effectieve voorspelling optimaal is bij verschijnselen waarbij de kans op het al of niet vóórkomen ongeveer even groot is. Bij een zeer kleine *base rate* is het per definitie bijzonder moeilijk om de voorspelling van gedrag boven de toevalskans te verbeteren: "Typically the likelihood of violent recidivism is so low that that it is hard to improve on the prediction that no one will commit a violent offense." (Harris c.s., 1993)

Met andere woorden, algemene recidive van ex-tbsgestelden is in theorie veel makkelijker te voorspellen dan ernstige (sexuele) geweldsrecidive. Uit recidive-onderzoek blijkt een *base rate* van rond 60% voor algemene recidive en een *base rate* van 10% à 15% voor ernstige 'tbs-achtige' recidive. (Leuw, 1995)

Het feit dat voorspellen moeilijker is naarmate de toevalskans van het voorspelde criterium dichter bij 0 of 1 ligt hangt samen met de 'soorten fouten' die bij toewijzingsbeslissingen gemaakt worden. In de beslistheorie worden alle mogelijke relaties tussen voorspeller en criterium beschreven in de volgende 2x2 tabel:

criterium

		+	-
	+	1(++) hits	2(+/-) valse positieven
voorspeller			
	-	3(-+) missers	4(--) ware negatieven

Het criterium zou in dit geval wel of geen *waargenomen* recidive kunnen zijn en de voorspeller wel of geen *voorspelde* recidive. Juiste voorspellingen vinden we bij de combinaties ++ en -- en onjuiste voorspellingen bij de combinaties +/- en -/. Bij een voorspelling met een *lagere* base rate, zoals bij geweldsrecidive, bestaat een onvermijdelijk *grotere* kans op het ontstaan van 'valse positieven'. In dit geval dus de *onterechte* voorspelling dat de tbs-gestelde zich opnieuw aan geweldsrecidive zal schuldig maken. Dit betekent dat het opvoeren van het percentage 'hits' (tbs-gestelden die terecht als 'gevaarlijk' worden aangewezen), onvermijdelijk een daarmee verbonden vergroting met zich meebrengt van het relatieve aantal tbs-gestelden dat ten onrechte als 'gevaarlijk' wordt bestempeld. Anders gezegd, beslissen op basis van de score op een meetinstrument brengt onvermijdelijk het dilemma met zich mee dat bij minimalisering van het aantal 'missers' maximalisering van het aantal 'valse positieven' optreedt. In termen van de beslistheorie wordt gesteld dat er een complementaire relatie bestaat tussen de *sensitiviteit* (het vermogen om 'hits' aan te wijzen) en de *specificiteit* (het vermogen om 'ware negatieven' aan te wijzen) van een test. (Dembo c.s., 1996)

In de praktijk wordt er vaak naar gestreefd om een optimaal beslissingscriterium te formuleren (de *cut-off score*) waarbij het totale percentage fouten (missers + valse positieven) zo gering mogelijk is. In wezen is een dergelijke strategie echter afhankelijk van de vraag of het onwenselijker is om delinquenten langer vast te houden of om het risico te lopen dat opnieuw een delict wordt gepleegd. Binnen de tbs-context lijkt een conservatieve strategie, minimalisering van het aantal missers, meer gebruikelijk.

Voor een infrequent voorkomend te voorspellen kenmerk zoals geweldsrecidive geldt dat een meetinstrument een grote voorspellende kracht moet hebben om de tot dusver gehanteerde klinische voorspelling, zoals die in de Nederlands tbs gebruikelijk is, te kunnen verbeteren. De volgende uitspraak van Blackburn (1993) is hier van toepassing: "... problems arise whenever the base rate is low, as it usually is for violent reoffending... 'beating' a low base rate calls for predictors whose association with the criterion is greater than is typically the case in either clinical or actuarial prediction. Efficiency may therefore depend on predicting for subgroups whose base rate is close to 50%." Het onderzoek van Harris c.s. (1993) biedt een voorbeeld dat gegeven een kleine *base rate*, zelfs bij een betrekkelijk hoge correlatie tussen voorspeller en criterium, een hoog percentage valse positieven de prijs is voor het kunnen opvoeren van het aantal hits. Er wordt gerapporteerd over een meetinstrument dat .44 correleert met geweldsrecidive. Bij een optimale *cut off score* (dat wil zeggen een score waarbij de som van de fouten minimaal is) wordt 73% van de recidivisten correct voorspeld. Dit levert echter tegelijkertijd 45% valse positieven op. Onder de op grond van de test als 'gevaarlijk' geselecteerden wordt dus bijna de helft *ten onrechte* geselecteerd.

Tenslotte verdient nog een laatste kenmerk van de TBS-populatie aandacht in verband met de te verwachten verbetering van de klinische voorspelling door het aanwenden van meetinstrumenten. Uit vrijwel alle beschikbare literatuur terzake blijkt dat voorspellen binnen

een heterogene populatie effectiever is dan voorspellen binnen een homogene populatie. Zo blijkt bijvoorbeeld bij voorselectie in een algemene groep ouders de Child Abuse Potential Inventory redelijk goed te kunnen onderscheiden tussen risico-groepen. (Milner, 1991) Het is reeds lang bekend dat verreweg de meest robuuste voorspeller van de kans op delinquent gedrag bestaat uit het eerder hebben gepleegd van criminaliteit. Binnen een groep als 'psychopathen' gediagnosticeerden bleek echter de voorgeschiedenis met betrekking tot geweldsdelinquentie geen goede voorspeller meer te zijn voor geweldsdelinquentie. (Serin, 1996) De verklaring hiervoor ligt voor de hand. Binnen een heterogene populatie bestaat een grote variatie op de predictie-variabelen, terwijl deze binnen een homogene populatie gering is. In het eerste geval kunnen daarom veel makkelijker personen binnen de populatie worden aangewezen die zich onderscheiden met betrekking tot het voorspellende kenmerk. De TBS-populatie is een bij uitstek homogene populatie. Het gaat in (vrijwel) alle gevallen om ernstige geweldsdelinquenten die (in meerderheid) lijden aan een betrekkelijk nauw te omschrijven *range* van psychiatrische symptomen. En tenslotte geldt voor de hier aan de orde zijnde beslissing aan het eind van de tbs-behandeling dat alle personen langdurig aan een zelfde soort interventie hebben blootgestaan. Oneerbiedig gezegd komt het er op neer dat een reeds bij aanvang betrekkelijk homogene populatie nog homogener is geworden doordat zij door 'de zelfde molen is gegaan'.

De conclusie van de in deze paragraaf besproken beslistheoretische overwegingen dient te zijn dat binnen de tbs-context niet meer dan een beperkte bijdrage kan worden verwacht van op test gebaseerde voorspelling van 'gevaar'.

5. Een overzicht van risicofactoren voor de voorspelling van 'gevaar'

Uit de overvloed van literatuur over de voorspelling van (gewelds)delinquentie komt een grote variëteit en een zeer groot aantal met het criterium samenhangende variabelen naar voren. In principe kunnen alle variabelen waarvan empirisch is gebleken dat ze significant samenhangen met het voorspelde gedrag dienen als predictie-factoren. Ze worden dan eenvoudig begrepen als factoren die bijdragen aan het risico van het optreden van 'gevaar'. Ten aanzien van deze risico-factoren kan een onderscheid worden gemaakt dat belangrijke consequenties heeft voor zowel het inzicht in het ontstaan van 'gevaar' als voor de mogelijkheden om het voorspelde 'gevaar' te beheersen.

Monahan c.s. (1994) spreekt van voor de verklaring en beheersing van gevaar relevante categorieën risicofactoren. Deze worden gevormd door alle voor de voorspelling van gevaar relevante groeps- of individuele kenmerken. Optimale voorspelling komt tot stand bij de kennis van zoveel mogelijk relevante risico-factoren. Deze kennis heeft zowel aspecten van (theoretische) validiteit, als van (operationele) betrouwbaarheid. Er wordt van uitgegaan dat voorspelling beter is naarmate er een breder domein van levenssterreinen mee wordt bestreken. In veel opzichten essentieel is het onderscheid dat wordt gemaakt tussen *statische* en *dynamische* risico-factoren. Statische factoren kunnen worden beschouwd als vastliggende en in principe éénmalig te meten variabelen die een bekende relatie met het te voorspellen gedrag hebben, maar waarop geen invloed kan worden uitgeoefend. Veelal zullen dit gegevens zijn uit de *personal history* van personen over wie 'gevaar' wordt voorspeld: de justitiële voorgeschiedenis, socialisatie-geschiedenis, milieu van herkomst etc. Daarnaast gaat het om als onveranderlijk beschouwde persoonlijkheidskenmerken zoals bijvoorbeeld impulsiviteit of cognitieve stijl. Dynamische risico-factoren daarentegen zijn veranderlijk en beïnvloedbaar. Monahan c.s. (1994) benoemen de statische risico-factoren als dispositioneel en historisch en de dynamische als klinisch en contextueel. De eerste zijn 'relatively enduring traits, waarvoor geldt dat "there is nothing a clinician can do to 'manage' the risk'. De laatste zijn 'current states'. Deze zijn veranderlijk. Ze zijn zowel 'natuurlijk' fluctuerend als beïnvloedbaar door (psychiatrische) interventie. "Dispositional and historical variables are relevant to risk management in the sense that management strategies can be conditioned on dispositional and historical variables".

Relevante, maar onveranderbare risico-factoren dienen ten tijde van de klinische beslissing met betrekking tot gevaar (bijvoorbeeld bij de beslissing over al of niet verlengen van de TBS) bekend te zijn. De 'current state' factoren kunnen echter per definitie niet eenmalig worden vastgesteld. Dit betekent dat het principiële onmogelijk is om op basis van een éénmalige inschatting van 'gevaar' een voor de toekomst geldige, op een enkel individu betrokken, objectieveerbare risico-inschatting te doen. Met andere woorden voor optimale individueel toepasbare risico-inschatting is het noodzakelijk om periodiek een inschatting te doen van de 'current state' factoren. Dit laatste heeft een zeer belangrijke consequentie voor de mogelijkheden van de toepassing van op test gebaseerde voorspelling in de tbs-context. Een test is per definitie een momentopname. De inschatting van 'gevaar' die op moment t1 (bijvoorbeeld bij het verlaten van de instelling) voor waar gehouden kan worden, kan op moment t2 (bijvoorbeeld een jaar na het verlaten van de instelling) volledig achterhaald zijn. Praktisch gezien betekent dit dus dat ook een optimale inschatting van 'gevaar' de eventuele noodzaak van herhaalde beoordelingen onverlet laat.

Monahan c.s. (1994) bieden in de nu volgende tabel een structuur van risico-factoren (predictie-variabelen) aan de hand waarvan voor de voorspelling van 'gevaar' gebruikte meetinstrumenten goed kunnen worden beschreven. De genoemde categorieën corresponderen met de statische of dynamische domeinen waaruit de risico-factoren afkomstig zijn.

Tabel: Een overzicht van relevante risicofactoren

Bron: J. Monahan and H.J. Steadman, *Violence and mental disorder*, Chicago 1994

1. Dispositional factors	<i>History of Crime and Violence</i>
<i>Demographic</i>	Arrests
Age	Incarcerations
Gender	Self-reported Violence
Race	Violence Toward Self
Social Class	
<i>Personality</i>	3. Contextual Factors
Personality Style	<i>Perceived Stress</i>
Anger	<i>Social Support</i>
Impulsiveness	Living Arrangements
Psychopathy	Activities of Daily Living
<i>Cognitive</i>	Perceived Support
IQ	Social Networks
Neurological Impairment	Means for Violence (i.e. guns)
2. Historical Factors	4. Clinical Factors
<i>Social History</i>	<i>Axis I Diagnosis</i>
Family History	<i>Symptoms</i>
Child rearing	Delusions
Child abuse	Hallucinations
Family deviance	Symptom Severity
Work History	Violent Fantasies
Employment	<i>Axis II Diagnosis</i>
Job Perceptions	<i>Functioning</i>
Educational History	<i>Substance Abuse</i>
<i>Mental Hospitalization History</i>	Alcohol
Prior Hospitalizations	Other Drugs
Treatment Compliance	

Vrijwel alle meetinstrumenten, in de praktijk gaat het bijna zonder uitzondering om in Noord-Amerika ontwikkelde en gebruikte instrumenten, bestaan uit een specifiek samenstel van meerdere in bovenstaande tabel opgenomen variabelen. Afhankelijk van het specifieke doel van het meetinstrument zijn variabelen uit de verschillende categorieën meer of minder vertegenwoordigd. En uiteraard neemt ook de operationalisatie van de variabelen verschillende vormen aan. In de laatste paragraaf zullen we enkele meetinstrumenten noemen die op het eerste gezicht mogelijk in aanmerking zouden komen voor gebruik in de Nederlandse tbs-praktijk.

In de recente theorievorming rond meetinstrumenten voor de voorspelling van 'gevaar' wordt een onderscheid gemaakt tussen 'risk assessment' en 'need assessment'. Tot op zekere hoogte correspondeert dit onderscheid met de bovengenoemde statische en dynamische risicofactoren. *Risk assessment* leidt tot de bepaling van het bij een persoon behorende *risk-level*. Positionering op deze dimensie wordt vooral bepaald door de traditionele (statische) predictievariabelen, zoals leeftijd, sexe, criminele voorgeschiedenis, socialisatiegeschiedenis (waaronder

de betekenis van *criminal associates*) etc. *Need assessment* is sterk tijd en situationeel gebonden. Positionering op deze dimensie wordt onder meer bepaald door behoeften en tekorten op de volgende levensgebieden: werk en inkomen, huwelijk en gezin, sociaal functioneren, middelengebruik, financiële huishouding en emotioneel-normatieve oriëntatie, attitudes (bijvoorbeeld negativisme, anti-sociale denkbeelden). In een onderzoek bleek algemene recidivisme sterk samen te hangen met de positionering op beide dimensies. In de *high-risk/high-need* groep recidiveerde 45%, terwijl in de *low-risk/low-need* slechts 6% recidiveerde. De relevantie van deze aanpak voor gestoorde geweldplegers is niet bekend. Het onderzoek had betrekking op een algemene groep delinquenten. (Motiuk, 1993)

Met betrekking tot het laatste punt is het overzichtsartikel van Gendreau c.s. (1996) relevant waarin wordt aangegeven dat ook voor geestelijk gestoorde delinquenten anti-sociale cognities (attituden) niet minder belangrijke risico-factoren zijn dan voor niet-gestoorde delinquenten. In dit artikel worden *criminogenic needs* genoemd als de belangrijkste dynamische voorspellers van 'gevaar': "... those sets of attitudes, values, beliefs and behaviors held by an offender that support a. negative attitudes towards all forms of official authority and conventional pursuits.. b. deviant values that justify aggression, hostility and substance abuse c. rationalizations from antisocial behavior that free one from moral constraints."

Het lijkt zeker plausibel dat dergelijke '*current cognitive states*' in belangrijke mate zullen bijdragen aan het bestaan van risico voor (gewelds)recidive. Helaas wordt uit het artikel niet duidelijk is hoe een betrouwbaar antwoord kan worden verkregen op vragen naar dergelijke evidente sociale (on)wenselijkheden.

6. Enkele mogelijk bruikbare meetinstrumenten

In Canada en de VS bestaat een sterk ontwikkelde traditie van het gebruik van op test gebaseerde voorspelling van het risico van (gewelds)recidive. In de ongetwijfeld nog lang niet uitputtende verzameling literatuur waarop dit overzicht is gebaseerd staat al een aantal van ca. 20 tests vermeld. Sommige van deze tests lijken een sterk ad hoc karakter te hebben. Er zijn bijvoorbeeld tests die speciaal binnen een bepaalde *parole board* worden gebruikt. Andere zijn zeer specifiek gericht op bepaalde specifieke situaties en doelgroepen, bijvoorbeeld voor de bepaling van het risico van geweld ten opzichte van kinderen of de partner in het gezin. Uit de beschikbare literatuur komen ook enkele meetinstrumenten naar voren met relatief gunstige kwalificaties in termen van relevantie voor geweldsrecidive, effectiviteit en validiteit. De vraag is nu of dergelijke meetinstrumenten bruikbaar zouden zijn voor gebruik in de Nederlandse tbs-context. En welke rol dergelijke meetinstrumenten zouden kunnen spelen bij het tot stand komen van de beslissing omtrent het al of niet voortzetten van de tbs-behandeling. Op basis van de nu beschikbare informatie kunnen we op dit moment hierover slechts enige kanttekeningen maken. In elk geval zou het aanbeveling verdienen om de mogelijkheden op dit punt nader te bestuderen en vooral ook om deze experimenterende wijze binnen de tbs-praktijk te ontwikkelen.

Er bestaat allereerst een probleem van technische aard. Zoals gezegd gaat het vrijwel zonder uitzondering om Noord-Amerikaanse meetinstrumenten. Concepten, operationalisaties en normeringen zijn ontwikkeld in en vooralsnog alleen geldig binnen deze samenlevingen. Dit wordt bijvoorbeeld geïllustreerd door enig internationaal vergelijkend onderzoek met de Psychopathy Check List Revised (PCL-R). Cooke (1997) vermeldt de experimentele toepassing van dit Amerikaanse meetinstrument in Schotland. Hierbij werden bij vergelijkbare Amerikaanse en Schotse populaties geheel verschillende scoreverdelingen gevonden. De auteur noemt het waarschijnlijk dat deze verschillen het gevolg zijn van culturele (samenlevingsgebonden) factoren. Voor Nederlands gebruik zou hier ongetwijfeld ook nog eens een semantisch (een vertaalprobleem) bijkomen.

Uit de literatuur komen enkele meetinstrumenten naar voren die lijken te voldoen aan de genoemde psychometrische kenmerken (relevantie voor geweldsrecidive, en relatief gunstige theoretische validiteit en effectieve predictie). Het gaat om de bovengenoemde PCL-R (Hare, 1991); de Violence Risk Appraisal Guide (VRAG) is een uitbreiding van de PCL-R (Rice c.s, 1997 en Gendreau c.s, 1997). De laatste auteur beveelt ook de Violence Risk Scale, de Pride In Delinquency Scale en de Psychological Inventory of Criminal Thinking Styles aan als instrumenten die specifiek zijn gericht op de voorspelling van geweldsrecidive.

Gezien de vele in dit stuk genoemde beperkingen van de te verwachten en (beslistheoretisch) mogelijke effectiviteit van op test gebaseerde voorspelling van 'gevaar' lijkt het niet waarschijnlijk dat tests de gebruikelijke klinische voorspelling in de Nederlandse tbs zullen kunnen vervangen. Wel lijkt het de moeite waard om de mogelijkheid te onderzoeken dat dergelijke tests één van de componenten van de te nemen beslissing aan het eind van de tbs-behandeling gaan vormen.