



**Universiteit
Leiden**
The Netherlands

In de oude fout : over het meten van recidive en het vaststellen van het succes van strafrechtelijke interventies

Wartna, B.S.J.

Citation

Wartna, B. S. J. (2009, March 5). *In de oude fout : over het meten van recidive en het vaststellen van het succes van strafrechtelijke interventies*. *Onderzoek en beleid WODC*. Boom Juridische uitgevers, Den Haag. Retrieved from <https://hdl.handle.net/1887/13596>

Version: Not Applicable (or Unknown)

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/13596>

Note: To cite this publication please use the final published version (if applicable).

1 Recidive

Introductie

Recidiveren betekent 'terugvallen'. Een recidivist valt terug in crimineel gedrag. Hij vervalt in de oude fout, wordt ook wel gezegd, of: 'hij is in herhaling vervallen'. Ook buiten de criminologie is het begrip bekend. In hun onderzoek naar financiële regelingen gebruiken Kersten et al. (1997) de term in de betekenis van 'het hebben van nieuwe schulden'. In de medische wetenschap spreekt men van recidive wanneer een schijnbaar genezen ziekte zich opnieuw vertoont. Er is hier sprake van een parallel, want ook criminaliteit kan worden opgevat als een persoonlijke aandoening die blijft sluimeren en onder bepaalde omstandigheden weer de kop opsteekt. Maar de parallel schiet tekort. Recidive heeft binnen de criminologie een algemenere betekenis. Het begrip past niet alleen in het medische model, maar hoort thuis in elke theorie over de oorzaken van delinquent gedrag. Het meten van recidive is voor de criminologie letterlijk van centraal belang. Tarling, het gewezen hoofd van het onderzoeksbureau van het Britse Home Office, typeerde recidiveonderzoek als 'one of the key tools of criminology' (Lloyd, Mair & Hough, 1994). Het vaststellen van herhalings-criminaliteit is een empirische toets voor alle mogelijke ideeën over het ontstaan, het aanhouden en het tegengaan van crimineel gedrag. De laatste jaren staat het recidivebegrip weer sterk in de belangstelling van de politiek. In 2002 lanceerde het kabinet-Balkenende II het deltaplan *Naar een veiliger samenleving*, beter bekend als het 'Veiligheidsprogramma'. Het plan bevat een groot aantal maatregelen met als doel de misdaad in ons land tegen te gaan en de veiligheid en de veiligheidsgevoelens van de burgers te versterken. Veel aandacht gaat daarbij uit naar de 'veelplegers'. Veelplegers zijn de recidivisten per uitstek. Het zijn personen die frequent in aanraking komen met politie en justitie. Een betrekkelijk kleine groep, het gaat naar schatting om circa 12% van alle daders, is verantwoordelijk voor een onevenredig groot deel van de criminaliteit. Meer in het algemeen, constateert ook het kabinet, is het zo dat de meeste verdachten al eens eerder met justitie in aanraking kwamen. Meer dan de helft van de geregistreerde criminaliteit wordt gepleegd door een recidivist. Indien deze groep, met de veelplegers voorop, kan worden afgestopt, zal de veiligheid in het land aanzienlijk toenemen en komen de doelstellingen van het kabinet dichterbij. In het vervolgprogramma *Veiligheid begint bij Voorkomen* (2007), uitgebracht door het vierde kabinet-Balkenende, worden de aspiraties op dit punt geconcretiseerd. Het kabinet wil in de periode 2002-2010 de recidive onder ex-gedetineerden en jeugdige daders met procentpunten verlagen. Met deze 'recidivedoelstelling' maakt het kabinet zich kwetsbaar, want eenvoudig lijkt de opdracht niet. Het expliciteren van ambities en het werken met streefgetallen past echter in de nieuwe, open stijl van overheidshuishouding, die bekend is geworden onder de slagzin 'van beleidsbegroting naar beleidsverantwoording (VBTB)' (Ministerie van Financiën,

2000). Ministers dienen jaarlijks begrotingsvoorstellen in waarin zij per beleidsterrein duidelijk maken wat ze willen, hoe ze dat denken te bereiken en wat het kost. Via jaarverslagen en andere rapportages leggen zij verantwoording af over het gevoerde beleid. Recidive fungeert daarbij steeds vaker als prestatie-indicator. Ook Justitie moet kunnen laten zien wat haar beleid heeft opgeleverd. Op het terrein van de strafrecht-toepassing – een terrein waarop in 2006 naar schatting 6,7 miljard euro werd uitgegeven (Moolenaar, 2007), wordt recidive, of liever: het uitblijven daarvan, omarmd als een geldige en efficiënte kwaliteitsmaat. Met recidivecijfers kunnen de prestaties van de organisatie of van organisatieonderdelen eenvoudig in kaart worden gebracht. Tenminste, dat is de veronderstelling. De vraag is echter onder welke voorwaarden dit het geval is. Wanneer vormt recidive een bewijs van de effecten van het beleid?

Tabel 1.1 Percentages recidivisten onder volwassen daders met een strafzaak afgedaan in de periode 1997-2003; naar type afdoening

	1997	1998	1999	2000	2001	2002	2003
Onvw. vrij ≤ 6 mnd	62,3	63,3	61,5	64,5	63,2	63,4	63,2
Onvw. vrij > 6 mnd	47,5	48,3	47,0	48,3	48,8	49,6	48,3
Taakstraf	33,0	33,4	32,2	36,0	37,2	34,9	33,5
Voorw. vrij	32,5	33,0	32,2	34,5	34,9	36,5	35,7
Geldstraf	19,7	19,9	20,2	20,7	21,4	22,3	22,3
Beleidssepot	24,7	24,2	24,5	25,6	25,9	26,2	27,1
Totaal	26,2	26,5	26,4	27,8	28,2	28,5	28,1

Bron: Rijksbegroting 2007

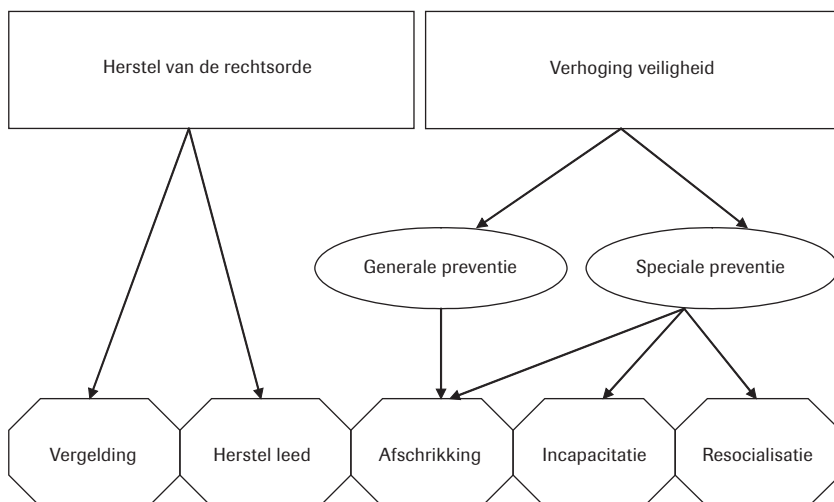
De betekenis van recidivecijfers wordt nog wel eens overschat. In de justiebegroting van 2007 is tabel 1.1 te vinden onder de titel 'prestatiegegevens op het terrein van rechtshandhaving, criminaliteitsbestrijding & terrorismebestrijding', een van de terreinen waarop het kabinet jaarlijks zijn plannen ontvouwt. De cijfers geven aan welk deel van de mensen die in een bepaald jaar een bepaalde straf kregen, aangemerkt kan worden als een recidivist. We zien in de onderste rij dat in totaal ruim een kwart van de volwassen daders heeft gerecidiveerd. Er is sprake van een lichte toename; het percentage recidivisten is tussen 1997 en 2003 met bijna 2% gestegen. Betekent dit dat het justitiebeleid heeft gefaald? Dat is moeilijk te bepalen. Ligt de recidive van Nederlandse burgers überhaupt wel binnen de invloedssfeer van de overheid? Wanneer een organisatie of een overheidsinstelling wordt afgerekend op haar prestaties, moet men ervan uit kunnen gaan dat de prestaties worden gemeten in 'maakbare grootheden'. Zelfs de beste coach kan een groep stenen niet leren voetballen; er moet wel iets te beïnvloeden zijn, anders kan men niet zinvol spreken van

succes of, omgekeerd, van falen. Een algemene doelstelling op het gebied van recidive is een loffelijk streven, maar om de cijfers omlaag te krijgen zal men op zijn minst een ruw idee moeten hebben langs welke weg dit zou kunnen worden gerealiseerd.

1.1 Strafdooelen en strategieën om recidive tegen te gaan

De wens om de terugval van daders te beteugelen is natuurlijk niet nieuw. Het reduceren van recidive valt samen met ‘speciale preventie’, een van de klassieke strafdooelen. Figuur 1.1 geeft de doeleinden van strafrechtelijke straffen en maatregelen – samen de ‘sancties’ – weer. Het schema is ontleend aan een tekst van Jonkers uit 1975 (Buruma, 1999). Hij onderscheidt doelen op legislatief, applicatief en executief niveau. Hier is van belang dat de hogere doelen van het strafrecht in de praktijk van de strafrechttoepassing – dus bij het opleggen en uitvoeren van de sancties – een verschillende invulling krijgen.

Figuur 1.1 Functies van het strafrecht en doelen van straffen en maatregelen



In Nederland geldt de ‘verenigingstheorie’. Straffen worden opgelegd om te vergelden, om recht te zetten wat is fout gedaan, maar ook om herhaling te voorkomen. Het toepassen van het strafrecht heeft daarmee twee maatschappelijke functies: het herstellen van de rechtsorde en het beveiligen van de samenleving. Het plegen van een strafbaar feit maakt inbreuk op het rechtsgevoel. Het kwaad kan worden hersteld door de dader een straf op te leggen die in verhouding staat tot de schade die is toegebracht. Dat

kan grofweg op twee manieren. Door leedtoevoeging – de dader moet dan dus letterlijk boeten voor wat hij heeft gedaan ('het hem betaald zetten') – of door reparatie van het leed dat het slachtoffer of de slachtoffers is aangedaan. Het eerste valt samen met het klassieke, abstracte idee van vergelding. De tweede strategie hangt samen met meer recente denkbeelden over *restorative justice*, de benadering van het recht waarin herstel van de feitelijke relaties tussen alle betrokkenen – dader, slachtoffer en de gemeenschap – centraal staat (Braithwaite, 2002).

Het tweede hoofddoel, de beveiliging van de maatschappij, wordt door het strafrecht op verschillende manieren gediend. Op de eerste plaats zorgt het bestaan en het toepassen van strafbepalingen ervoor dat veel mensen geen delicten plegen. Zij houden zich aan de wet. Personen die de wet overtreden, worden gestraft om anderen te laten zien welke normen er gelden en om de strafdreiging reëel te houden. Het straffen van daders moet anderen er dus van weerhouden delicten te plegen. Naast deze 'generale preventie' wil men via het strafrecht ook voorkomen dat de dader opnieuw een delict pleegt. Dit strafdoel heet 'speciale preventie'. Het kan grofweg op drie manieren worden bereikt: door afschrikking, door incapacitatie en door resocialisatie. Van afschrikking is sprake indien de dader een nieuwe overtreding vermijdt uit vrees voor de negatieve gevolgen. Het verschil met generale preventie is dat hij die negatieve gevolgen zelf al eens aan den lijve heeft ondervonden. Incapacitatie betekent 'onschadelijkmaking' of 'uitschakeling'. De sanctie die wordt opgelegd maakt het de dader onmogelijk, of althans lastiger, om opnieuw een delict te plegen. Resocialisatie ten slotte, houdt in dat de dader veranderingen doormaakt en in het vervolg beter in staat en bereid is zich aan de geldende normen te houden.

Samen met vergelding en leedherstel vormen afschrikking, incapacitatie en resocialisatie de vijf elementaire strategieën die een rechter in individuele gevallen tot zijn beschikking heeft om de hoofddoelen van het strafrecht te bereiken. Bij het bepalen van de sanctie maakt de rechter – expliciet of impliciet – een keuze uit deze mogelijkheden. Kiest hij voor speciale preventie, dan stelt hij de dader centraal. Het gaat dan immers om het voorkómen van recidive. Bij de overige twee strategieën liggen de zaken anders. Daar spelen de belangen van het slachtoffer en van de burgers in het land een doorslaggevende rol en is de vraag of de straf de kans op herhaling groter of kleiner maakt, hoogstens van secundair belang. Een straf waarbij 'opsluiting' het oogmerk was, is een succes als de opsluiting geëffectueerd is en zonder incidenten is verlopen. Met andere woorden, recidive is alleen een maat voor het welslagen van een sanctie indien men speciale preventie voor ogen heeft.

Nu is het niet zo dat elke sanctie slechts één strafdoel kent. Meestal, en misschien wel altijd, geldt er een combinatie van doelstellingen. Accentverschillen zijn er wel; de straffen en maatregelen die in Nederland worden opgelegd, hebben geen uniform karakter. Sommige zijn meer

bedoeld als vergelding, bij andere ligt de nadruk meer op de resocialisatie van de dader. Een korte vrijheidsstraf uitgevoerd in een gesloten inrichting met een sober regime, is gericht op vergelding en afschrikking. Bij een lange gevangenisstraf komt daar nog het element van incapacitatie bij. Afhankelijk van de concrete invulling doen zich ook binnen de sancties variaties voor. Het feitelijke karakter van de voorwaardelijke vrijheidsstraf bijvoorbeeld, wordt sterk bepaald door de inhoud van de gestelde voorwaarden. Sommige bijzondere voorwaarden, zoals het opleggen van straatverboden, zijn gebaseerd op het idee van incapacitatie. Andere, zoals de verplichting van de dader om deel te nemen aan therapie, zijn gericht op resocialisatie. Er vindt dan immers behandeling plaats. Er zijn weinig straffen in Nederland die reparatie van het aangedane leed als uitgangspunt hebben. We kennen de schadevergoeding, maar dat is een bijkomende straf en geen hoofdstraf. Daarnaast zijn er procedurele voorzieningen, zoals slachtofferhulp en het spreekrecht voor slachtoffers en nabestaanden van ernstige delicten. Van de hoofdstraffen heeft alleen de werkstraf het element van reparatie van leed in aanleg in zich. Bij de introductie van de 'alternatieve sancties' in de jaren '80 van de vorige eeuw leefde nog het idee dat de straf een weerspiegeling moest zijn van het delict. Maar de directe verbinding tussen de schade die met het delict werd veroorzaakt en de verplichte werkzaamheden is gaandeweg verdwenen. De straf is een bulkstraf geworden, primair bedoeld als vergelding en afschrikking. Bij volwassenen staat de werkstraf na de boete op de tweede plaats en bij minderjarige daders is het momenteel de meest gebruikte sanctie; in meer dan de helft van de jeugdstrafzaken wordt een werkstraf opgelegd. Bij geen van beide leeftijdsgroepen is schade- of leedherstel nog een belangrijk oogmerk. Het element van resocialisatie speelt wel een rol, zij het indirect. De werkstraf vervangt een vrijheidsstraf en zorgt er zo voor dat 'detentieschade' wordt voorkomen. Zo staat een werkstraf resocialisatie in elk geval niet in de weg.

Vanuit het oogpunt van vergelding wordt de gevangenisstraf alom gezien als de zwaarste sanctie die het Nederlandse strafrecht kent. De gevangenisstraf geldt als een 'ultimum remedium', het laatste redmiddel dat vanwege de nadelige bijwerkingen zo lang mogelijk moet worden vermeden. Het idee dat vrijheidsbeneming schadelijk zou zijn voor de dader en mogelijk zelfs leidt tot verhoging van de kans op recidive, geldt niet voor alle vormen van detentie. Dat hangt weer af van de invulling van de straf. Vrijheidsstraffen zijn primair gericht op vergelding, incapacitatie en (generale) afschrikking, maar tijdens de detentie kan er worden gewerkt aan de resocialisatie van de dader. Bij volwassen gedetineerden gebeurt dat tegenwoordig in het kader van een 'penitentiair programma'. In de laatste fase van de detentie wordt de dader meer vrijheden gegund en vindt er training en begeleiding plaats om hem voor te bereiden op de terugkeer naar de vrije maatschappij. Daarnaast kennen we natuurlijk de terbeschikkingstelling (tbs). Deze maatregel is primair gericht op de

beveiliging van de maatschappij. Maar tijdens de bewaring vindt behandeling plaats van de psychische stoornis die een rol speelde bij het plegen van het delict. Net als bij de verschillende penitentiaire programma's is bij de terbeschikkingstelling resocialisatie dus het uiteindelijke doel. In beide gevallen streeft men tijdens de tenuitvoerlegging van de sanctie naar verbetering van de uitgangspositie van de justitiabele. De detentie moet uiteindelijk een positieve uitwerking hebben in plaats van dat ze schade berokkent.

1.2 De recidivemeting als toets van de speciaal preventieve werking van interventies

Of de sanctie nu wordt opgelegd om de dader af te schrikken, om hem 'onschadelijk' te maken of te resocialiseren, de speciaal preventieve werking van een sanctie kan door middel van een recidivemeting worden vastgesteld. Sinds de jaren '70 van de vorige eeuw wordt in Nederland gebruikgemaakt van recidivecijfers om het succes van strafrechtelijke interventies uit te drukken. Het onderzoek naar de terugval van justitiabelen is niet altijd even populair geweest. In de jaren '80 vonden slechts enkele studies plaats. In de jaren '90 echter deed zich in de criminologie, meer bijzonder in de penologie, een kentering voor in het denken over de mogelijke effecten van strafrechtelijke interventies. Samen met de eerder besproken ontwikkelingen in het openbaar bestuur heeft dit er toe geleid dat de belangstelling voor dit type onderzoek weer is toegenomen.

De jaren '80 staan bekend als de 'ijstijd' van de penologie. Het was de periode waarin het geloof in de resocialisatiegedachte – het idee dat daders door of tijdens hun straf kunnen verbeteren – naar een dieptepunt was gedaald. In 1975 publiceerden Lipton, Martinson en Wilks een inventarisatie van het effectonderzoek dat tussen 1945 en 1967 in de Verenigde Staten was uitgevoerd. Uit hun overzichtsstudie bleek dat geen enkele strafrechtelijke interventie, van welke aard of intensiteit dan ook, een verschil tot stand had kunnen brengen in de recidive na afloop van de straf. Hun conclusie beruiste op de analyse van de uitkomsten van meer dan 200 studies. Naar later bleek was op het onderzoek wel het een en ander aan te merken, maar de teerling was geworpen. De stelling dat via het strafrecht de kansen van justitiabelen kunnen worden gekeerd, werd verlaten en maakte plaats voor het idee van 'just deserts': het strafrecht moet straffen en verder niet. In deze benadering gaat de aandacht uit naar het vinden van de juiste strafmaat gebaseerd op de ernst van het delict en de mate van schuld. De beveiligingsgedachte speelt daarbij een ondergeschikte rol. Op individueel niveau is hoogstens de 'uitschakeling' van daders nog van belang.

De 'nothing works'-doctrine heeft het denken over interventiestrategieën voor justitiabelen bijna twintig jaar in haar greep gehad. In de jaren '90

deed zich echter een omslag voor. Eerder al had Palmer (1975) het materiaal van Lipton, Martinson en Wilks aan een heranalyse onderworpen, maar nu drongen er vanuit Canada en de Verenigde Staten ook nieuwe geluiden door over de vooruitzichten voor *rehabilitation*, de Amerikaanse term voor resocialisatie. Gendreau en Ross (1979, 1987) waren de eersten die successen meldden. Andrews en Bonta (1995, 2003) wezen erop dat interventies wel degelijk effectief kunnen zijn, mits zij:

- 1 zijn afgestemd op het recidiverisico van de justitiabelen;
- 2 gericht zijn op de criminogene factoren en situaties die horen bij de betreffende groep;
- 3 en aansluiten op de leerstijl van de deelnemers.

Deze drie beginselen – verkort weergegeven als *risk*, *needs* en *responsivity* – golden lange tijd als de drie basisprincipes van een benadering die bekend is komen te staan onder de naam ‘what works’. What works is een sterk op de uitvoeringspraktijk gerichte beweging waarin men met empirisch onderzoek tracht na te gaan onder welke condities (strafrechtelijke) interventies uitzicht bieden op speciale preventie. Inmiddels is de lijst van *do's and don'ts* belangrijk uitgebreid. Het ‘compendium on effective correctional programming’ uit 2001, een soort handleiding voor justitiemedewerkers, telt maar liefst achttien basisprincipes waaraan de interventies voor justitiabelen dienen te voldoen.

In Nederland hebben de inzichten uit de what works-literatuur iets later ingang gevonden, maar ook hier was vanaf het begin van de jaren '90 sprake van een nieuw elan. Door de geluiden van overzee groeide de overtuiging dat door het bieden van hulp en begeleiding de kansen van justitiabelen kunnen worden gekeerd. Binnen de strafrechttoepassing kwam weer ruimte om te experimenteren. De resocialisatiegedachte, of het idee van ‘maatschappelijke reïntegratie’ zoals dat in die tijd vaak werd genoemd, stond daarbij voorop. Op tal van terreinen werden proefnemingen met preventieprojecten gestart en een groot aantal recidivestudies was daarvan het gevolg. Na de eeuwwisseling zette deze situatie zich voort, zij het onder een ander gesternte. Uit politieke overwegingen – er moest worden bezuinigd en er was groeiende onrust over de omvang van de criminaliteit – verschoof het ideologische brandpunt van het strafrechtelijke beleid opnieuw in de richting van vergelding en incapacitatie. Het is echter niet zo dat de resocialisatiegedachte totaal verdween. Achter de schermen werd doorgewerkt aan de randvoorwaarden voor ‘effectieve gedragsinterventies’. Resocialisatie is in Nederland al enige tijd geen algemeen strafdoel meer, maar als er aan wordt gewerkt, dan gebeurt dat steeds uitdrukkelijker op basis van de inzichten uit de internationale literatuur en wordt het resultaat gemeten in termen van de terugval van de deelnemers. Kortom, ook in Nederland is de recidivemeting inmiddels een erkend en veelgebruikt evaluatie-instrument.

1.3 Probleemstelling

Onder welke voorwaarden vormen recidivecijfers het bewijs voor de effecten van strafrechtelijke interventies? Dat is in het kort de probleemstelling van dit proefschrift. Een recidivemeting is weliswaar het geëigende middel om het speciaal preventieve effect van een strafrechtelijke interventie te meten, het is echter niet zo dat met het genereren van cijfers over de terugval van justitiabelen er onmiddellijk een haarscherp beeld ontstaat van de werking van een sanctie. De cijfers in tabel 1.1, in de openingsparagraaf van dit hoofdstuk, vormen hiervan een illustratie. De tabel maakt duidelijk dat van alle sancties de korte, onvoorwaardelijke vrijheidsstraffen de hoogste recidivepercentages laten zien. Meer dan 60% van de kortgestraften recidiveert. Afgaand op deze cijfers zou men geneigd kunnen zijn te denken dat lange vrijheidsstraffen effectiever zijn dan korte. En dat het gelet op de lagere recidive nog beter zou zijn om plegers van misdrijven een geldstraf of een beleidssepot op te leggen. Maar iedereen die dit proefschrift leest, zal beseffen dat dit een denkfout is. Beleidssepots en boetes worden opgelegd aan andere personen en in andere situaties dan korte of lange vrijheidsstraffen. Een directe vergelijking van de recidivepercentages is dus niet op zijn plaats. Een deel van de hoogteverschillen in de tabel kan wellicht worden toegeschreven aan de effecten van de sancties, maar we weten niet welk deel. Misschien is het wel zo dat het over het geheel genomen niets uitmaakt welke sanctie er wordt opgelegd. Er zijn criminologen die uitgaan van deze hypothese van 'gelijke werking' (zie o.m. Kerner, 1996). Een straf zou positief kunnen uitwerken op de een, en negatief op de ander. Per saldo is er tussen de sancties dan misschien wel geen verschil te zien.

Los van de interpretatie of de zeggingskracht van de cijfers, speelt hier echter ook een ander probleem: hoe wordt de recidive eigenlijk gemeten? Hoe zijn bijvoorbeeld de cijfers van tabel 1.1 tot stand gekomen? Er wordt gesproken van 'recidivisten', maar wat wordt daar precies onder verstaan? Wie heeft de metingen verricht en op welke wijze vond het onderzoek plaats? Het bepalen van het succes van strafrechtelijke interventies is technisch gesproken bepaald geen sinecure. Het voorbeeld in box 1.1 geeft aan hoe belangrijk het is dat de recidive zorgvuldig wordt gemeten. Een detail, een klein verschil in de meetmethode kan verstrekken gevolgen hebben. Als we werkelijk willen weten 'wat werkt', zullen we ernst moeten maken met de wijze waarop de uiteindelijke uitkomsten van de strafrechtelijke interventies worden vastgesteld.

Box 1.1 Den Engh en recidive: meten is weten?

In 2003 vroeg een kamermeerderheid de Minister van Justitie haast te maken met de uitbreiding van plaatsen in rijksinrichting Den Engh en op de Glen Millsschool. In beide inrichtingen ligt het accent op groepsdiscipline en respect voor hiërarchische verhoudingen. De stevige aanpak in beide inrichtingen sprak veel politieke partijen aan. Men was blij dat er een antwoord was gevonden op het hardnekkige probleemgedrag van groepen jonge delinquenten die vooral in de grote steden voor veel overlast zorgden. De cijfers waren bovendien bemoedigend. Uit onderzoek onder de eerste jongeren die de 'sociogroepsstrategie' in Den Engh hadden afgerond, was gebleken dat slechts 9% binnen één jaar recidiveerde (Jonker, 2004). Een percentage dat veel lager is dan het percentage onder ex-pupillen uit behandelinrichtingen in het algemeen. Toen de recidivemeting onder de eerste deelnemers van de sociogroepsstrategie werd herhaald, leverde dit echter een teleurstelling op. De uitkomsten van de eerste meting konden niet worden gerepliceerd. Niet 9 maar 28% bleek binnen één jaar na vertrek te hebben gerecidiveerd; vergelijkbaar met de terugval onder jongeren uit andere behandelinrichtingen.

Het verschil in de uitkomsten kon worden teruggevoerd op wat in eerste instantie slechts een detail leek. Bij de eerste meting was het moment waarop de observatieperiode startte, niet voor iedereen hetzelfde. Bij de deelnemers van de sociogroepstherapie viel het begin samen met de afloop van het feitelijke onderwijsprogramma, bij de jongeren uit andere behandelinrichtingen begon de meting pas op het moment dat men uit de inrichting werd ontslagen. De personen uit de vergelijkingsgroep waren langer 'at risk' geweest dan de deelnemers aan sociogroepstherapie en daarom viel het recidivepercentage in de vergelijkingsgroep bij het eerste onderzoek hoger uit. Bij de tweede meting was de onderzoeksmethode voor beide groepen identiek (Wartna, Kalidien & Essers, 2006).

In dit proefschrift zal worden ingegaan op de problematiek van het meten van herhalingscriminaliteit en het gebruik van recidive als maat voor het succes van strafrechtelijke interventies. De dissertatie bestaat uit zes hoofdstukken. In dit eerste hoofdstuk is het belang van het recidivebegrip vanuit verschillende invalshoeken belicht. Voor politiek en beleid passen de pogingen om de recidive te beperken naadloos in het streven naar meer veiligheid. Voor bestuur en beheer vormen recidivecijfers een kengetal. Het meten van de recidive schept de mogelijkheid de effecten van het beleid in kaart te brengen en uitvoeringsorganisaties ter verantwoording te roepen. In de penologie is recidive verbonden met 'speciale preventie', een van de klassieke strafdoelen. En voor de criminologie in het algemeen is een recidivemeting een toetssteen voor theorieën over het ombuigen van criminele carrières.

Ontwikkelingen die zich op het terrein van het openbaar bestuur en binnen de criminologie hebben voorgedaan, verklaren waarom de belangstelling voor recidivemetingen de laatste jaren weer is toegenomen. Voor een deel werken wetenschap en overheid aan dezelfde opdracht: de terugval onder justitiabelen moet omlaag. Maar het meten van recidive is niet zonder problemen en de meetresultaten kunnen gemakkelijk verkeerd worden geïnterpreteerd. Om onoordeelkundig gebruik te voorkomen dienen voorwaarden te worden geformuleerd. En zo komen we tot de formele probleemstelling van dit proefschrift: *hoe kunnen recidivemetingen worden vormgegeven en welke plaats nemen zij in bij de evaluatie van strafrechtelijke interventies?* Op deze vraag moet de dissertatie antwoord bieden. Op de inhoud van de overige hoofdstukken ga ik hier kort in, zodat de lezer in grote lijnen weet wat hij mag verwachten.

1.4 Opbouw van de dissertatie

Hoofdstuk 2 geeft een overzicht van de belangrijkste methodologische ontwikkelingen die zich op het gebied van het recidiveonderzoek hebben voorgedaan. Recidivegegevens worden binnen de criminologie op vier manieren gebruikt, wij concentreren ons op de toepassing in evaluatieonderzoek. Vanaf de 19e eeuw worden er recidivemetingen gedaan om het succes van de aanpak van delinquenten aan te tonen. Engeland kent een lange onderzoekstraditie op dit terrein. Deze zal worden beschreven. Maar ook andere ontwikkelingen komen aan bod: de experimentele onderzoeksdesigns die overkwamen uit de Verenigde Staten, de opkomst van risicotaxatie-instrumenten, het gebruik van duurmodellen voor de analyse van de gegevens, de toepassing van meta-analyse en meest recent: standaardisering en monitoring.

Onder invloed van de *what works*-benadering vindt veel onderzoek plaats naar de speciaal preventieve werking van strafrechtelijke interventies. De wetenschappelijke tijdschriften staan vol artikelen waarin wordt nagegaan of deelname aan een bepaald programma de kans op recidive binnen een bepaalde doelgroep kleiner maakt. Meta-analyse is een methode om de uitkomsten van de losse effectstudies te overzien. Door te werken met een overkoepelende maat, de *effect size*, worden de methodologische verschillen tussen de afzonderlijke studies als het ware uitgeschakeld en kan de meta-analist meer algemene uitspraken doen over 'wat werkt, voor wie en onder welke omstandigheden'. In Nederland is een poging gedaan de resultaten van 29 internationale overzichtsstudie en meta-analyses op een rij te zetten (Wartna, Baas & Beenackers, 2004). Ook hierop zal in hoofdstuk 2 kort worden ingegaan.

Op de methode van de meta-analyse is veel kritiek, vooral uit de hoek van sociologen als Pawson en Tilly, die een minder mechanistisch en meer kwalitatieve methode van evaluatie voorstaan. Centraal in hoofdstuk 2

staat echter een andere kwantitatieve methode van research synthese die naast de meta-analyses in opkomst is: standaardisering en monitoring. In een toenemend aantal landen wordt de recidive telkens op dezelfde manier berekend. Standaardisering van de meetmethode maakt het even-goed mogelijk vergelijkingen te maken tussen de uitstroomresultaten van verschillende strafrechtelijke interventies. De ontwikkelingen op dit terrein, in Europa, de Verenigde Staten en Canada, worden beschreven. In Nederland is er op dit vlak met de totstandkoming van de zogeheten 'Recidivemonitor' een belangrijke stap gezet. De Recidivemonitor is een project waarin periodiek gestandaardiseerde metingen worden verricht onder grote groepen justitiabelen die uiteenlopende sancties hebben ondergaan.

Hoofdstuk 3 gaat uitgebreid in op de meetmethode die in het kader van de Recidivemonitor werd ontwikkeld. Voordat een standaard tot stand kwam, is het recidivebegrip eerst conceptueel 'gefileerd'. Er is antwoord gezocht op vragen als: welke facetten kent het begrip? Met welke zaken krijgt men bij het meten van de recidive te maken? En: welke gegevensbronnen zijn er voor het doen van recidiveonderzoek beschikbaar? Om tot een vaste aanpak te komen zijn op tal van onderdelen keuzes gemaakt, die in hoofdstuk 3 zullen worden toegelicht. Tevens wordt aandacht besteed aan de methoden en technieken die binnen de Recidivemonitor worden gebruikt om de verzamelde gegevens te analyseren.

Als in hoofdstuk 3 het meetinstrument is beschreven, kunnen in hoofdstuk 4 de meetresultaten worden getoond. Ingegaan wordt op de uitkomsten van recidiveonderzoek onder jeugdige en volwassen daders die wegens een misdrijf in aanraking kwamen met Justitie. Meer specifiek wordt een beeld geschetst van de terugval onder ex-gedetineerden, onder jongeren uitgestroomd uit een justitiële jeugdinstelling en onder ex-terbeschikkinggestelden. Ook wordt aandacht besteed aan soortgelijk onderzoek uit het buitenland. Een aantal landen heeft net als Nederland recidivecijfers op nationale schaal tot de beschikking. Het populatieonderzoek van de Recidivemonitor levert inzicht op in de *base rates* van tal van dadergroepen. Het laat niet zien welke sancties bij welke dadergroepen moeten worden ingezet om de recidive zo laag mogelijk te houden. Het is dus niet zo dat de hoogte van de recidive model staat voor de effectiviteit van de sanctie. Dit komt omdat ook andere factoren, zoals populatiekenmerken en registratie-effecten, van invloed zijn op de cijfers. Door de ruwe cijfers met behulp van statistische technieken te corrigeren kan de invloed van een deel van deze exogene factoren worden geneutraliseerd. Maar omdat dit slechts beperkt mogelijk is, zullen algemene recidivecijfers nooit een hard bewijs vormen van de effectiviteit van de opgelegde straffen en maatregelen. Daar is aanvullend onderzoek voor nodig.

En met die vaststelling zijn we aangeland in hoofdstuk 5. Dit hoofdstuk plaatst het gebruik van recidivecijfers terug in de meeromvattende context van het onderzoek naar de effectiviteit van strafrechtelijke interventies. Recidivemetingen moeten aan tal van voorwaarden voldoen, maar daarmee zijn we er nog niet. Onderzoek naar de effecten van strafrechtelijke interventies bestaat uit meer dan een recidivemeting. Een recidivemeting is het sluitstuk van een reeks van onderzoeksactiviteiten, een reeks die hier zal worden aangeduid als de 'evaluatieketen'. We willen niet alleen vaststellen óf een interventie werkt. We willen ook weten waarom een strafrechtelijke interventie erin slaagt de recidive te beperken en hoe dit dan precies in zijn werk gaat. Aan de fase waarin de uiteindelijke *outcome* wordt gemeten gaan daarom andere typen van evaluatieonderzoek vooraf. De 'evaluatieketen' is een opeenvolging van onderzoeksactiviteiten waarin eerst de achtergronden van de interventie worden doorgelicht en de uitvoeringspraktijk aan een analyse wordt onderworpen, alvorens de uitkomsten worden gemeten in termen van de recidive van de deelnemers aan het programma. Het idee van de evaluatieketen vormt het voor de hand liggende compromis tussen de experimentele benadering die ten grondslag ligt aan de kwantitatieve meta-analyses en het kwalitatieve, *theory-driven* onderzoek dat door Pawson en Tilly wordt bepleit. In hoofdstuk 5 wordt beschreven hoe de ideale opbouw van het evaluatieonderzoek er uit ziet.

In hoofdstuk 6, de slotbeschouwing, belanden we met beide voeten op de grond en wordt het een en ander in een actueel kader geplaatst. Wie de wereld van de strafrechttoepassing kent, weet dat de ideale onderzoeksomstandigheden zelden kunnen worden gerealiseerd. Onderzoek waarin sprake is van random toewijzing van kandidaten aan nauw omschreven condities komt niet vaak voor en heeft, als het voorkomt, doorgaans een beperkt bereik. In het slothoofdstuk zal dan ook geen eenzijdig pleidooi worden gehouden voor het uitvoeren van meer gerandomiseerde studies. Wetenschappers pleiten al jaren voor onderzoek met het sterkst denkbare design, maar hun oproep vindt op het terrein van de strafrechttoepassing tot dusver weinig gehoor. In de praktijk blijkt het toch steeds weer erg lastig te zijn om dergelijk onderzoek te realiseren. De stelling hier is dan ook dat we op dit terrein ook moeten uitkijken naar mogelijkheden om 'slecht' onderzoek beter te maken. Misschien levert dat wel meer op dan het wachten op de weinige momenten dat we de *gold standard* van het evaluatieonderzoek kunnen toepassen. Dit proefschrift is geschreven vanuit het motto: 'If you can't find proof, why not use the evidence?' Wacht niet op het onomstotelijke bewijs, maar maak gebruik van de aanwijzingen die er zijn. Alle typen onderzoek, in alle fasen van de evaluatieketen, kunnen aanwijzingen opleveren die bijdragen tot een rationeler sanctiebeleid. Juist nu de voornemens op het terrein van de sanctietoepassing in een aantal nota's zijn geformuleerd, is er behoefte aan een

samenhangend en ‘nuchter’ onderzoeksprogramma waarin ook het meten van recidive de plaats krijgt die het verdient. In het slothoofdstuk wordt nog eens geschetst hoe dit onderzoeksprogramma eruit kan zien.

2 Kleine geschiedenis van het recidiveonderzoek

Een overzicht van de belangrijkste methodologische ontwikkelingen¹

In dit hoofdstuk gaan we in op de methodologische ontwikkelingen die zich – internationaal en in Nederland – op het gebied van het recidiveonderzoek hebben voorgedaan. Een duik in de geschiedenis biedt de gelegenheid de thema's die voor dit proefschrift van belang zijn nader te introduceren. Het toont ons de herkomst van de problemen waarmee men op dit vlak te maken heeft.

2.1 Het gebruik van recidivegegevens

Recidivegegevens kunnen, afhankelijk van de probleemstelling van het onderzoek, op verschillende manieren worden gebruikt (Wartna, 1999):

- 1 als beschrijvend kenmerk of onderdeel van een daderprofiel;
- 2 in onderzoek naar het verloop van criminele carrières;
- 3 in etiologisch onderzoek naar de oorzaken van criminaliteit; en
- 4 bij de evaluatie van strafrechtelijke of buitengerechtelijke interventies.

De vier typen onderzoek liggen in elkaars verlengde. Bij beschrijvend recidiveonderzoek is het al dan niet vóórkomen van herhaald crimineel gedrag een kenmerk waarmee een groep personen kan worden gekarakteriseerd. Bij onderzoek naar criminele carrières vormt de recidive de afhankelijke variabele. Vaak staat dit type onderzoek in het teken van de voorspelling van herhaald crimineel gedrag. Bij etiologisch onderzoek bestudeert men de oorzaken van herhaald crimineel gedrag. Een bekend voorbeeld is de Pittsburg Youth Study, waarin gegevens over de vroegtijdige ontwikkeling van een groot aantal stadskinderen in verband worden gebracht met criminaliteitspatronen op latere leeftijd (Loeber et al. 1998). Het vierde type recidiveonderzoek – het evaluatieonderzoek – gaat weer een stap verder. Nu gaat het om het bepalen van het effect van een interventie, een tussenkomst in de omstandigheden die normaal gesproken tot recidive leiden. De interventie grijpt dus aan op wat wordt gezien als de oorzaak van crimineel gedrag. De recidive fungeert bij dit type onderzoek als uitkomstmaat. Hoe lager de recidive, des te groter het veronderstelde effect van de interventie.

In dit proefschrift gaat het om het vierde type onderzoek, om de inzet van recidivegegevens bij de evaluatie van strafrechtelijke interventies. In paragraaf 2.2 staan we eerst echter kort stil bij het onderzoek naar criminele carrières, in het bijzonder bij de predictie van recidive. De ontwikkelingen

¹ Dit hoofdstuk bevat fragmenten uit artikelen verschenen in het *Tijdschrift voor Criminologie* (Wartna, 1999, 2000).

op dit terrein zijn namelijk ook van belang voor het effectonderzoek. Als men weet wie er zal recidiveren, kan men immers proberen in te grijpen. Paragraaf 2.3 voert ons terug naar de eerste recidivemetingen bedoeld om het succes van de aanpak van justitiabelen vast te stellen. Engeland en Amerika kennen beide hun eigen onderzoekstraditie, die kort zal worden besproken. We gaan daarbij onder meer in op het probleem van het *design* van het onderzoek en eindigen met de bespreking van moderne analysemethoden. Paragraaf 2.4 behandelt het Nederlandse effectonderzoek. In vogelvlucht wordt nagegaan wat veertig jaar aan recidivemetingen heeft opgeleverd aan kennis over de speciaal preventieve werking van onze straffen en maatregelen. In paragraaf 2.5 bespreken we de opkomst van meta-analyse. Deze techniek is voor de criminologie op dit moment van groot belang. Bij meta-analyse creëert men een gemeenschappelijke overkoepelende maat waarin het succes van de onderzochte strafrechtelijke interventies wordt uitgedrukt. Het is een simpele manier om de uitkomsten van afzonderlijke, methodologisch uiteenlopende evaluatiestudies te overzien. Een volgende stap zou kunnen zijn om bij evaluaties steeds dezelfde uitkomstmaat te gebruiken. Men hoeft dan nadien geen overkoepelende maat te construeren, want de onderzoeksresultaten zijn al onderling vergelijkbaar. In steeds meer landen wordt het recidiveonderzoek herhaald uitgevoerd en is de onderzoeksmethode gestandaardiseerd. Deze ‘standaardisering en monitoring’ is het onderwerp van paragraaf 2.6. Paragraaf 2.7 bevat enkele slotopmerkingen.

2.2 Predictie van recidive: hoe groot is de kans op terugval?

Het vroegste onderzoek naar het verloop van ‘criminele carrières’ vinden we terug in het begin van de vorige eeuw, bij de sociologen van de Chicago-school (Van de Bunt, 1988). Het is sterk kwalitatief van aard en verbonden met het symbolisch interactionisme. Vertegenwoordigers van deze onderzoekstraditie zien de ontwikkeling van een deviante loopbaan als een resultante van een reeks van sociale interacties (Sutherland, 1937; Shaw, 1967; Lemert, 1972). Na het verschijnen van het werk van Blumstein et al. (1986) veranderde het carrièreonderzoek van karakter.² De opstellers van dit lijvige Amerikaanse overheidsrapport deelden het begrip criminele carrière in vier dimensies in:

- 1 deelname; het onderscheid tussen daders en mensen die geen delicten (zullen) plegen;
- 2 de frequentie van de gepleegde delicten;
- 3 de ernst van de gepleegde delicten; en
- 4 de duur van de periode waarin de delicten worden gepleegd.

² Een kwalitatieve benadering waarbij de nadruk ligt op de processen die tot feitelijke recidive leiden, vinden we nog terug bij Zamble en Quinsey (1997) in hun boek *The criminal recidivism process*.

De meetkundige benadering van Blumstein c.s. kreeg in tal van landen navolging en is sterk verbonden met begrippen als ‘risicotaxatie’ en ‘incapacitatie’, de strategie die is gericht op het uitschakelen van vermoedelijke recidivisten (zie figuur 1.1). De grondgedachte is dat door het blootleggen van de regelmatigheden die zich in het verloop van criminele carrières voordoen, het mogelijk wordt te voorspellen welke daders opnieuw in de fout zullen gaan. Op basis van de inschatting van het recidiverisico zou men bepaalde daders langer in de gevangenis kunnen houden, om zo de criminaliteit te kunnen beteugelen.

Ook in Nederland werd eind jaren '80 kwantitatief onderzoek gedaan naar het verloop van criminele carrières. Een goed voorbeeld hiervan is te vinden bij Block en Van der Werff (1991). Zij beschrijven de criminele loopbanen van ‘de meest actieve en gevaarlijke daders in Nederland’. Block en Van der Werff stellen vast dat in Nederland circa 5% van de daders die in 1977 werden veroordeeld, verantwoordelijk is voor bijna 15% van alle voor deze cohort geregistreerde justitiecontacten in de daaropvolgende periode van zes jaar. Zij noemen deze daders ‘loopbaancriminelen’. Het zijn de daders die vrij snel op elkaar voor ten minste twee relatief ernstige delicten werden vervolgd. Mogelijkheden om hen vroegtijdig te herkennen, dus nog voordat hun criminele loopbaan van start is gegaan, zijn er in de ogen van Block en Van der Werff niet. Bij aanvang van hun carrière onderscheidde de latere loopbaancriminelen zich niet sterk genoeg van de daders die niet of slechts met lichte delicten recidiveerden. Selectieve onschadelijkmaking – het uitdelen van zwaardere gevangenisstraffen aan hen die met een zekere waarschijnlijkheid zullen recidiveren – is volgens hen dus niet goed uitvoerbaar. Het zou betekenen dat veel personen ten onrechte zouden worden gedetineerd.

Dezelfde conclusie vinden we – in Nederland – terug bij Nijboer (1975), Van der Werff (1986) en meest recent bij Nieuwbeerta en Blokland (2006)³. Duidelijk is dat crimineel gedrag – net als veel andere vormen van menselijk gedrag – slechts beperkt voorspelbaar is. Niettemin zien we de laatste decennia ook in Nederland een sterke opmars van het gebruik van formele procedures voor risicotaxatie. Steeds vaker worden er tijdens de strafrechttoepassing standaardvragenlijsten of checklists afgenomen waarmee voor elke justitiabele een individuele risicoscore kan worden berekend. In de strafrechtsketen is op verschillende momenten een inschatting nodig van het recidivegevaar. Bij de voorgeleiding, om te bepalen of de verdachte in voorlopige hechtenis moet worden genomen; bij de strafoplegging, als de rechter kiest voor een bepaalde sanctie, en bij de tenuitvoerlegging,

3 Deze onderzoekers van het Nederlands Studiecentrum Criminaliteit en Rechtshandhaving (NSCR) heranalyseerden de gegevens van Van der Werff en schatten de kosten en baten van verschillende scenario's van selectieve onschadelijkmaking. De kosten die nodig zijn om een bepaald deel van de daderbevolking langer te detineren, blijken in hun analyses steeds groter te zijn dan de opbrengsten in termen van criminaliteitsreductie. Zie voor de ethisch-juridische aspecten van selectieve incapacitatie, onder meer de oratie van Moerings (2003), getiteld *Straffen met het oog op veiligheid: een onderneming vol risico's*.

wanneer wordt beslist welke invulling de straf krijgt. Vroeger berustte de taxatie van het risico doorgaans op het oordeel van gedragskundigen; psychologen of psychiaters die op grond van hun expertise en na gesprekken met de cliënt een advies uitbrachten aan de rechter of het OM. Vanuit de psychologie is echter al enige tijd bekend dat in dit soort situaties het gebruik van statistische modellen betere voorspellingen oplevert dan het klinische oordeel van experts (zie o.m. Mossman, 1994; Trout & Bishop, 2002; Aegisdottir et al., 2006). Bij gebruik van een formeel instrument wordt de voorspelling gebaseerd op verbanden aangetroffen in een onderzoekspopulatie, terwijl het klinisch oordeel van de expert noodzakelijkerwijs berust op beperkte ervaringen en een subjectieve blik.

In de tbs-sector, waar op grond van een risico-inschatting moet worden bepaald of een patiënt met verlof kan of kan worden ontslagen, heeft dit inzicht geleid tot de ontwikkeling van verschillende predictie-instrumenten waarbij vaste dimensies worden nagelopen om tot een risicoscore te komen (Philipse, 2005; De Vogel, 2005; Hildebrand et al., 2006; Brand & Van Emmerik, 2006). Deels gaat het dan om vertalingen van buitenlandse instrumenten, deels om zelfgeconstrueerde vragenlijsten. Ook in andere sectoren van de Nederlandse strafrechtspleging zijn inmiddels standaard-instrumenten in ontwikkeling, zoals de RISc (Vinke et al., 2004) en de QuickScan (De Ruiters & De Jong, 2006). Het zijn voorbeelden van methoden van 'gestructureerde klinische risicotaxatie'. Een expert gebruikt de wetenschappelijk gefundeerde checklist als leidraad om tot een eigen inschatting te komen. Komt de risicoscore tot stand volgens een vast algoritme, dan spreken we van een 'actuarieel' taxatie-instrument.

De voorspelkracht van risicotaxatie-instrumenten kan worden uitgedrukt in een samenvattende maat, de AUC (Hanley & MacNeil, 1982; Hosmer & Lemeshow, 1989). De AUC neemt een waarde aan tussen 0 en 1. Een score van 0,5 betekent dat een test heeft uitgewezen dat het instrument in de helft van de gevallen correct voorspelt wie van twee willekeurig gekozen personen de hoogste recidivekans heeft. Dat is geen bijzondere prestatie, want met het opwerpen van een munt bereikt men hetzelfde resultaat. Een AUC van 1 komt niet voor, het zou betekenen dat recidivisten en non-recidivisten vooraf perfect van elkaar zouden kunnen worden onderscheiden. In de praktijk valt de AUC van de meeste instrumenten uit tussen 0,65 en 0,80 (Dolan & Doyle, 2000; Philipse, 2005). Zeker niet perfect dus, maar wel een verbetering ten opzichte van het toeval en kennelijk voldoende om in de praktijk van de strafrechttoepassing behulpzaam te zijn bij tal van plaatsingsbeslissingen.

2.3 Recidive als maat voor effectiviteit

Dit proefschrift handelt slechts zijdelings over de predictie van recidive. Centraal staat de toepassing van recidivegegevens bij de evaluatie van

strafrechtelijke interventies. Het gebruik van recidive als uitkomstmaat is gestoeld op het principe van speciale preventie. In hoofdstuk 1 zagen we dat een strafrechtelijke ingreep tegelijk soms verschillende doelen kent. Puur bezien vanuit het oogpunt van speciale preventie is een interventie effectief als het de betrokkenen weerhoudt van het plegen van hernieuwde criminaliteit. Dit effect kan worden bereikt door:

- 1 'afschrikking': de betrokkene laat herhaling van het bestrafte gedrag achterwege uit vrees voor de consequenties die een nieuwe overtreding zal hebben;
- 2 'incapacitatie': de dader is door het opleggen van beperkingen niet in staat of in de gelegenheid om het gewraakte gedrag opnieuw te vertonen;
- 3 'resocialisatie': belemmeringen worden weggenomen en de betrokkene krijgt vaardigheden en mogelijkheden aangereikt waardoor hij meer bereid en beter in staat is om herhaling van het delictgedrag te voorkomen.

In feite is recidive geen succesmaat, maar een indicatie van mislukking (Maltz, 1984). Het vóórkomen van nieuwe criminaliteit wijst uit dat de doelstelling van speciale preventie niet is gehaald en dat de gevolgde strategie van afschrikking, incapacitatie of resocialisatie niet – althans niet uitputtend – heeft gewerkt.

Al in de 19e eeuw werden recidivemetingen verricht ter evaluatie van strafrechtelijke interventies. Radzinowicz (1945), de eerste Britse hoogleraar criminologie, wijst op een overheidsrapport waarin verslag wordt gedaan van de 'after-conduct' van 77 jongens die in de periode 1818-1825 werden vrijgelaten uit een jeugdgevangenis in Warwick County. De helft van de jongens keerde terug, sommige van hen meer dan eens (Command Paper, 1847). Later die eeuw hielden Engelse tuchtscholen en heropvoedingskampen bij hoe het de ex-bewoners na de invrijheidsstelling verging. De inrichtingen onderscheidden vijf mogelijkheden: 'doing well', 'doubtful', 'condemned', 'unknown' en 'dead'. Meer dan drie kwart van de kinderen werd ingedeeld in de eerste categorie (Command Paper, 1896). In een rapport van de Raad van Europa uit 1967 blikt dr. R. Hood terug op de effectstudies die na de Tweede Wereldoorlog in Engeland zijn uitgevoerd. Hij constateert dat het succes van een sanctie niet kan worden vastgesteld met behulp van recidivecijfers uit een *single treatment study*. Er dient contrast te worden aangebracht met een groep die de bewuste behandeling niet heeft ondergaan:

... 'without an assessment of what rate of reconviction could have been expected had the offenders had no treatment at all, all this is simply a statement of faith' (Hood, 1967: 77).

Met andere woorden, er is vergelijkingsmateriaal nodig om de meerwaarde van een interventie te bepalen. Met zijn verzuchting verwoordt Hood de zorg om een adequaat *onderzoekdesign*. Hij bespreekt een aantal studies waarin het benodigde vergelijkingsmateriaal werd gegenereerd door gebruik te maken van een vorm van statistische controle. De personen uit de onderzoeksgroep werden vergeleken met personen die een andere sanctie kregen, maar die, afgaand op een aantal belangrijke voorspellers, voorafgaand aan de straf hetzelfde recidiverisico kenden. Engeland kent een traditie in de toepassing van regressietechnieken bij de evaluatie van strafrechtelijke interventies. Deze traditie voert terug naar het werk van Mannheim en Wilkins (1955) en de Amerikaanse criminologen Sheldon en Eleanor Glueck (1930), die predictietafels gebruikten om recidive te voorspellen. Hammond was de eerste die voorspelde recidivepercentages vergeleek met feitelijke recidivepercentages, om zo een inschatting te kunnen maken van de effecten van de opgelegde sancties (Home Office, 1964). De voorspelde recidive is de recidive die men gelet op de achtergronden van de betrokkenen mag verwachten. Valt de feitelijk geobserveerde recidive anders uit, dat wil zeggen hoger of lager uit dan verwacht, dan komt dat – volgens deze redenering, althans – door de sanctie die zij hebben ondergaan. Dertig jaar later passen Lloyd, Mair en Hough (1994) hetzelfde idee toe in hun studie onder personen veroordeeld tot een gevangenisstraf, een taakstraf of bepaalde vormen van reclasseringstoezicht ('probation'). Ook zij constateren dat de hoogte van de recidive varieert. Ze hangt samen met factoren zoals de leeftijd en de sekse van de daders en het aantal eerdere veroordelingen. Lloyd, Mair en Hough gebruiken logistische regressie om de invloed van de instroomverschillen op deze kenmerken te neutraliseren en zo de 'netto-effecten' van de verschillende sancties in te schatten. Logistische regressie is bedoeld voor de voorspelling van *dichotome* uitkomstvariabelen (wel of geen nieuwe delicten). Ze is beter geschikt voor gebruik bij recidiveonderzoek dan normale, lineaire regressietechnieken (Hosmer & Lemeshow, 2000).

Experimentele designs

Terwijl in Engeland wordt gewerkt aan verfijning van methoden voor statistische controle van instroomverschillen, richt de aandacht in Amerika zich meer op de mogelijkheden van het experiment. In 1979 verschijnt in Amerika het rapport van het *Panel on Research on Rehabilitative Techniques* (Sechrest, White & Brown, 1979). Deze commissie heranalyseerde het werk van Martinson (1974) en Lipton, Martinson en Wilks (1975), de auteurs die verantwoordelijk worden gehouden voor het adagium *nothing works*: geen enkele interventie helpt.

Lipton, Martinson en Wilks bestudeerden 231 behandelprogramma's op het terrein van de strafrechttoepassing. Zij stelden vast dat zo goed als geen enkele poging een merkbaar effect had gehad op het recidiveniveau. Hun conclusie leidde tot een verhit debat tussen voor- en tegenstanders

van resocialisatie. De commissie, samengesteld uit leden die in deze kwestie nog geen standpunt hadden ingenomen, hield de evaluaties waarop Lipton, Martinson en Wilks zich hadden gebaseerd nog eens tegen het licht. Zij merkt op dat uit het overzicht van Lipton, Martinson en Wilks niet blijkt of de interventies wel goed waren uitgevoerd. Bovendien stelt de commissie vast dat de auteurs erg mild waren geweest bij de inclusie van effectstudies. Bij de meeste studies ontbrak een vergelijkingsgroep, of vertoonde het onderzoeksdesign andere zwakke plekken. Bij andere studies was het onderzoeks-aantal te klein. Kortom, de kwaliteit van de evaluaties liet veel te wensen over.⁴ De conclusie van het rapport uit 1979 luidde dat het gezien het gebrek aan adequaat onderzoek simpelweg nog niet bekend was welke interventies een recidiveverlagend effect hadden:

‘A great deal of money has been spent on research of rehabilitation, and all we are left with is no conclusions, weak conclusions, and, all too frequently but unpredictably, wrong conclusions. What is now needed is research with the most rigorous design and of the highest quality’ (Sechrest, White & Brown, 1979: 60).

In een apart hoofdstuk van het rapport wordt een aantal onderzoeksvormen besproken waarmee de effecten van interventies kunnen worden blootgelegd (Rezmovic, 1979). Het design van het onderzoek bepaalt het causale gehalte van de uitspraken die op basis van de uitkomsten van het onderzoek kunnen worden gedaan. De *interne validiteit* is de zekerheid waarmee een effect op de uitkomstmaat kan worden toegeschreven aan de uitgevoerde interventie. Het ‘ware’ of gerandomiseerde experiment wordt allereerst gezien als het best mogelijke design. Bij een gerandomiseerde studie worden de onderzoekspersonen op basis van toeval aan een behandel- of controleconditie toegewezen: een groep die wel en een groep die niet de bewuste straf of maatregel ondergaat. Randomisering geeft de best mogelijke garantie dat beide groepen vergelijkbaar zijn (zie o.m. Campbell, 1969; Boruch, 1975; Cook & Campbell, 1976, 1979; Weisburd, 2000, 2003; Farrington, 2003). Doet er zich tussen de twee groepen een verschil voor op de uitkomstmaat, in ons geval de recidive, dan kan dit effect zonder problemen worden toegeschreven aan de onderzochte interventie. Dat is immers de enige factor waarop de behandelgroep stelselmatig van de controlegroep verschilt.

Rezmovic stelt dat het gebruik van experimenten moet worden bevorderd, maar ziet in dat gerandomiseerde toewijzing om ‘juridische, ethische, politieke of logistische’ redenen niet altijd haalbaar zal zijn (ibid.: 166). In zo’n geval wordt het ‘quasi-experiment’ als het beste alternatief

⁴ Eerder kwam Logan (1972) al tot dezelfde conclusie. Logan bestudeerde honderd evaluaties van programma’s en interventies gericht op recidivereductie en constateerde dat geen van de studies voldeed aan de belangrijkste van de door hem opgestelde kwaliteitscriteria.

beschouwd. In het quasi-experimentele design worden de resultaten in de onderzoeksgroep afgezet tegen de uitkomsten van een groep personen die op zo veel mogelijk relevante kenmerken vergelijkbaar is. Matching op achtergronden die gerelateerd zijn aan de kans op recidive, lijkt sterk op de vorm van statistische controle die in de voorgaande sectie werd besproken. Een belangrijk verschil is dat bij 'observatiestudies' de invloed van instroomverschillen achteraf wordt gecorrigeerd, terwijl bij het quasi-experiment van tevoren een controlegroep wordt geformeerd. Dit laatste is een voordeel, omdat het de onderzoekers in staat stelt bij de controlegroep dezelfde informatie te verzamelen als bij de onderzoeksgroep. De mogelijkheden om de vergelijkbaarheid van beide groepen te controleren, zijn daardoor groter. Het voordeel van observatiestudies in vergelijking met experimenten in het algemeen, is dat ze goedkoper zijn en dat het onderzoek zelf geen inbreuk doet op het te meten gedrag. Ze zijn 'non-obtrusive'. Als mensen weten dat ze deel uitmaken van een onderzoek, gedragen zij zich misschien anders dan normaal en dat kan het meten van het ware effect van de interventie in de weg staan (Cook & Campbell, 1979: 53 e.v.).

Het meten en analyseren van de recidive

Het rapport van het *Panel* gaat niet alleen in op het design van het onderzoek, maar ook op andere methodologische kwesties, zoals de duur van de observatieperiode. Hoe lang moet men personen volgen om de effecten van een interventie te kunnen meten? Ook de bron van de recidivegegevens (de daders zelf, of cijfers van politie of justitie) en de operationele definitie van de recidive (welke delicten tellen mee?) zijn bepalend voor het meetresultaat.

Maltz (1984) behandelt dezelfde thema's, in zijn monografie *Recidivism*. Hij wijst op het verschil tussen *de facto* en *de jure* recidive, tussen het feitelijke criminele gedrag en wat daarvan in de administratie van politie en justitie terecht komt. Lang niet alle delicten worden door de instanties geregistreerd, maar toch vormen officiële cijfers vaak de basis voor het recidiveonderzoek. Maltz somt op welke operationele definities van recidive tot dan toe in het Amerikaanse onderzoek zijn gebruikt. De meeste studies kwalificeren recidive als een 'terugkeer naar de gevangenis'. Andere duiden recidive met 'overtredingen van reclasseringsvoorwaarden'. En ook een 'nieuwe veroordeling' of een 'nieuwe arrestatie' gold vaak als criterium. De laatste definitie geniet overigens zijn voorkeur. In de Verenigde Staten worden gegevens over arrestaties landelijk bijgehouden, gegevens over veroordelingen niet. Elke staat registreert enkel de eigen strafzaken.

Naast de keuze van de bron van onderzoeksgegevens vestigt Maltz ook aandacht op de methoden waarmee de verzamelde recidivegegevens kunnen worden geanalyseerd. Een belangrijke innovatie op dit gebied was de introductie van 'survival'- of 'failure time'-modellen. Bij dit type

technieken is niet de recidive op zich, maar de duur tot aan de recidive de afhankelijke variabele. Men analyseert met andere woorden niet alleen het vóórkomen van nieuwe delicten, maar ook de snelheid waarmee de personen in het onderzoek terugvallen in de criminaliteit. De term 'survival' is afkomstig uit de biostatistiek, de techniek zelf komt uit de hoek van de operations research, een onderzoeksveld waar men zich bezighoudt met het testen van de levensduur van gebruiksvoorwerpen, zoals gloeilampen. Partanen (1969) was vermoedelijk de eerste die de techniek in een criminologische context heeft toegepast, op de voet gevolgd door Carr-Hill en Carr-Hill (1972).⁵ De statisticus Cox (1972) introduceerde een regressietechniek voor duurvariabelen waarmee op eenvoudige wijze kan worden ingeschat welke invloed individuele kenmerken hebben op de kans dat er op enig moment in de observatieperiode recidive optreedt. De techniek kan worden ingezet voor de voorspelling van recidive, maar ook bij de evaluatie van strafrechtelijke interventies. Met Cox regressie kan worden gecontroleerd of de onderzochte sanctie een lagere recidivekans geeft, of dat het gemeten verschil in de uitkomst moet worden toegeschreven aan instroomverschillen tussen de experimentele en de controlegroep. Het is een vorm van statistische controle die ook van waarde kan zijn bij quasi-experimenten.

2.4 Effectstudies in Nederland: veertig jaar recidiveonderzoek

De eerste effectstudies in Nederland vonden plaats eind jaren '60, begin jaren '70 van de vorige eeuw (Jonkers et al., 1969; Jessen, 1969; Nijboer, 1971). Dit was relatief vroeg; op andere beleidsvelden duurde het tot de jaren '90 voordat er evaluatieonderzoek werd gedaan naar de maatschappelijke effecten van het overheidsbeleid (zie o.m. Hoogerwerf, 1977; Leeuw, 1992). De oprichting van het Wetenschappelijk Onderzoek- en Documentatiecentrum (WODC), in 1974, droeg hieraan bij. Het verschaftte een infrastructuur om dergelijk onderzoek in onafhankelijkheid uit te voeren. De eerste evaluaties op het gebied van de strafrechttoepassing hadden betrekking op gevangenisregimes (Fiselier, 1970; Nagel, 1970; Nijboer, 1971; Dijksterhuis, 1973; Van der Linden, 1978) en de aanpak van dronken rijders (Van der Werff, 1971; Steenhuis, 1972). Het eerste algemene recidiveonderzoek werd uitgevoerd door Van der Werff (1978). Zij zette de terugvalpercentages na de belangrijkste hoofdstraffen op een rij. Het onderzoek was gebaseerd op een steekproef uit de groep personen die in 1966 werden veroordeeld. Van der Werff verzamelde gegevens over de voortzetting van hun strafrechtelijke carrière gedurende een follow-up-periode van zes jaar. In die tijd bleek 41% van de veroordeelden te hebben

⁵ Voor de techniek in het algemeen geldt Kalbfleisch en Prentice (1980) als standaardwerk. Later verscheen Schmidt en Witte (1988). Een Nederlandstalige introductie is te vinden bij Wartna (2000).

gerecidiveerd. Ruim tien jaar later werd het onderzoek gerepliceerd en was het overall percentage opgelopen tot 51% (Van der Werff, 1986). Per strafsoort liep de recidive echter behoorlijk uiteen. Van der Werff beseftte dat haar cijfers niet model staan voor de differentiële effectiviteit van de diverse sancties. De veroordeelden in de diverse sanctiegroepen verschilden onderling qua leeftijd, sekse en de omvang van hun strafrechtelijk verleden en deze kenmerken hangen sterk samen met de kans op recidive. Door het uitvoeren van een lineaire regressie en het berekenen van partiële correlaties trachtte Van der Werff de invloed van deze kenmerken te corrigeren, om zo het netto-effect van de verschillende strafsoorten in te kunnen schatten. Haar voorzichtige conclusie luidde dat 'het speciaal preventieve effect van onvoorwaardelijke vrijheidsstraffen over het algemeen niet groter is dan dat van extramurale straffen, zoals boetes en voorwaardelijke straffen' (1978: 26).

Tot voor kort waren de onderzoeken van Van der Werff de enige twee algemene recidivestudies die in Nederland zijn verricht.⁶ Alle andere effectstudies beperken zich tot de evaluatie van één bijzondere interventie of strafsoort. In de jaren '80 vond niet veel evaluatieonderzoek plaats. Het was de periode waarin het geloof in de mogelijkheden tot resocialisatie zich ook in Nederland op een dieptepunt bevond. Veel hoop om de recidivekans via het strafrecht te verlagen was er niet. Men zocht naar buitenjustitiële oplossingen, zoals blijkt uit de experimenten met het 'lik-op-stuk'-beleid (Koolen, Wiebrens & Van den Berg, 1989) en de introductie van de Halt-afdoening (Kruissink & Verwers, 1989).

De tbs vormde hierop een uitzondering. Op dit terrein werd juist veel onderzoek gedaan. Vanaf het midden van de jaren '80 werd een reeks van studies opgezet waarin de recidive van alle uitgestroomde terbeschikkinggestelden volgens een vast stramien werd vastgesteld (Van Emmerik, 1985, 1989; Leuw 1995, 1999). Steeds werden de ex-terbeschikkinggestelden die tijdens een periode van vijf jaar werden ontslagen, voor een periode van minimaal drie tot maximaal acht jaar gevolgd. Hoewel hier in feite sprake is van beschrijvend onderzoek en strikt genomen niet van evaluatie, noemen we deze reeks van studies hier toch. Wat het interessant maakt, is dat het de omvang van de recidive over een langere tijd in kaart heeft gebracht. De uitkomsten van het vierde onderzoek leren ons dat de terugval onder terbeschikkinggestelden in de periode 1974-1994 is afgenomen (Leuw, 1999). Wat de betekenis is van dit resultaat, weten we niet goed. Want ook hier speelt het probleem van de vergelijkbaarheid.

6 Bedoeld wordt onderzoek gebaseerd op een landelijke steekproef waarin zowel voor volwassenen als voor jeugdigen in algemene zin de recidive na verschillende sanctiesoorten wordt bepaald. Bij Van Dijk (1971) vinden we de eerste verwijzing naar dit type onderzoek. Hij bespreekt de opzet van een recidivemeting onder alle personen die in 1947 voor het eerst werden veroordeeld. Om de effectiviteit van sancties als geldboetes en vrijheidsstraffen te bepalen, zouden de uittreksels uit de justitiële documentatie van maar liefst 23.000 personen worden gecodeerd en ingevoerd. De grootschaligheid is het onderzoek noodlottig geworden, want het werd na enige jaren stopgezet. Zo bleek uit persoonlijke communicatie met de auteur.

Alleen als men zou mogen aannemen dat de tbs-populatie onveranderd is gebleven, zegt deze uitkomst iets over het succes van de maatregel. Alleen dan wijst de daling van de recidive na de tbs op een verbetering van de behandelresultaten. In paragraaf 4.2.5 komen we op deze materie terug.

Box 2.1 Dagdetentie en recidive: meten is weten?

Begin jaren '90 vond een evaluatie plaats van een experimentele executiemodaliteit: 'dagdetentie'. Bij dagdetentie mogen langgestrafte gedetineerden de laatste zes weken van hun gevangenisstraf 's avonds en in het weekend alvast op hun thuisadres verblijven. Overdag melden zij zich bij de laag beveiligde inrichting, waar zij een divers programma volgen met activiteiten gericht op resocialisatie. Dagdetentie is bedoeld als 'voorbereiding op de terugkeer naar de vrije maatschappij' en zou moeten helpen 'de overgang tussen gesloten situatie en thuissituatie te vergemakkelijken' (MvJ, 1990). De suggestie om met recidiveonderzoek na te gaan of deze doelstellingen werden bereikt, vond geen genade bij de commissie die de evaluatie van het experiment begeleidde. Men betwijfelde of de gunstige werking van dagdetentie met recidiveonderzoek zou kunnen worden blootgelegd.

De commissieleden kregen gelijk. Jaren later – de jonge onderzoeker zette door, de recidivemeting kwam er uiteindelijk toch – bleek dat de eerste deelnemersgroep dezelfde terugval kende als langgestrafte gedetineerden die in vrijheid werden gesteld vanuit een 'open inrichting'. In dit type inrichting geldt alleen weekendverlof en wordt overdag buiten de inrichting gewerkt. De meerwaarde van de nieuwe executiemodaliteit kon met het onderzoek dus niet worden aangetoond, althans niet in termen van speciale preventie. Toch werd de formule na het experiment in Rotterdam uitgebreid. In 1998 kende ons land 11 inrichtingen voor dagdetentie met een totaal van 175 plaatsen.* Inmiddels is deze modaliteit omgezet in een zogenoemd 'penitentiair programma'.

* Met dank aan Paul Linckens (DJI).

In de jaren '90 werd veel meer onderzoek gedaan naar de effecten van strafrechtelijke interventies dan in het decennium daarvoor. Nadat vanuit Canada en de Verenigde Staten successen waren gemeld en er weer optimistische geluiden konden worden opgevangen over de aanpak van justitiabelen (o.m. Palmer, 1975; Cullen & Gilbert, 1982; Gendreau & Ross, 1979, 1987; Andrews et al., 1990), keerde ook in Nederland het geloof in de resocialisatiegedachte terug. Men wilde weer nieuwe dingen proberen op het terrein van de strafrechttoepassing. Een stoet van beleidsexperimenten was het gevolg, met in het kielzog daarvan een ongekende hoeveelheid evaluatiestudies. Van der Laan (1991) bestudeerde de opkomst van de 'alternatieve sancties' voor jeugdigen. Hij vergeleek de recidive van ruim 1.000 jongeren die in plaats van een boete of een vrijheidsstraf repara-

tie- en onderhoudswerkzaamheden kregen te verrichten, met de recidive van 1.140 traditioneel gestrafte jeugdigen. De recidive onder de alternatief gestraften was iets lager dan onder traditioneel gestraften (69 tegen 74% in een periode van 3½ tot 5½ jaar). Maar de ernst van het gepleegde uitgangsdeldict was dat gemiddeld gesproken ook. Kortom, de vergelijking ging niet helemaal op, want de controlegroep was 'zwaarder' dan de experimentele groep. Alleen in het 'zwaardere segment' van beide groepen bleek de recidivekans van de alternatief gestraften kleiner te zijn dan voor de traditioneel gestraften. Ten opzichte van (on)voorwaardelijke vrijheidsstraffen doen de alternatieve sancties het volgens Van der Laan dus beter, maar in vergelijking met de boetes vielen de resultaten tegen.

In een rapport met de veelzeggende titel 'Appels en peren' doet Spaans (1994) verslag van een recidivemeting onder 900 (jong)volwassen dienstverleners en bijna 1.000 kortgestraften. Hij constateert dat dienstverleners weliswaar minder, minder vaak, minder snel en minder ernstig recidiveerden, maar ook dat de controlegroep van kortgestraften een uitgebreider strafrechtelijk verleden kent en dat een vergelijking dus eigenlijk niet op zijn plaats is. Wel blijkt dat als men de analyses beperkt tot het zware segment, dat wil zeggen personen met meer dan tien eerdere justitiecontacten, de dienstverleners het beter doen dan hun tegenhangers met een korte vrijheidsstraf.

Naast de omvangrijke studies van Van der Laan en Spaans werden in de jaren '90 onderzoeken verricht naar de effecten van tal van nieuwe interventies. Er zijn evaluaties van cursussen voor daders van alcoholgerelateerde criminaliteit (Bovens, 1991; Den Otter, 1991; Bovens & Timmerman, 1992; Leuw & Brouwers, 1992, 1995; Slis & Boosten, 1995), van speciale projecten op het gebied van de jeugdreclassering (Spaans & Doornhein, 1991; Duipmans, 1996), van experimentele werk- en leerstraffen (Bunt & Breukelaar, 1991; Duipmans, 1993), van maatschappelijke integratieprogramma's en arbeidstoeleidingsprojecten (Essers et al., 1995; Hoff & Lagendijk, 1995; Wartna et al., 1996) en van projecten voor 'Harde Kern'-jongeren (Kleiman & Terlouw, 1997). In geen van deze studies werden op het punt van recidive spectaculaire uitkomsten gemeld. De verschillen met de controlegroep – als daar in de opzet van het onderzoek al sprake van is – zijn over het algemeen niet groot of laten een gemengd beeld zien. Dit geldt overigens niet alleen voor het onderzoek uit de jaren '90, maar ook voor de studies die hebben plaatsgevonden na de eeuwwisseling. Na 2000 ging aandacht uit naar de aanpak van criminele verslaafden (Vermeulen, Hendriks & Zomerveld, 2000; Bieleman et al., 2002; Koeter & Bakker, 2007) en jongeren uit jeugdinrichtingen (Heiden-Attema & Wartna, 2000; Bink, 2004; Wijkman, Van der Geest & Bijleveld, 2006). Ook deze meer recente studies hebben over het algemeen geen grote effecten aan het licht gebracht, met uitzondering wellicht het onderzoek van Nelissen (2003) naar de uitkomsten van een resocialisatieprogramma voor gedetineerden.

Bijna veertig jaar wordt in Nederland recidiveonderzoek gedaan. In de periode 1969-2008 zijn meer dan honderd studies verricht waarin op basis van empirische gegevens met betrekking tot de terugval van justitiabelen uitspraken worden gedaan over de uitkomsten van strafrechtelijke interventies (Wartna & Alberda, in voorbereiding). Veel aanwijzingen dat de ene interventie effectiever is dan een andere, heeft het onderzoek niet opgeleverd. Nergens in Nederland is de invoering van nieuwe sancties of experimentele executiemodaliteiten gepaard gegaan met een spectaculaire vermindering van de recidive. Is daarmee vast komen te staan dat het nutteloos is om naar resocialisatie te streven? Zeker niet, want als één ding opvalt wanneer men het binnenlandse onderzoek op dit terrein overziet, is dat aan de meeste studies wel het een en ander schortte. De meeste evaluaties vonden plaats in de proeffase van de interventie. De interventies waren dus nog in ontwikkeling toen de effecten werden opgetekend. De gehanteerde onderzoeksmethoden liepen sterk uiteen. Sommige onderzoekers gebruikten zelfrapportagegegevens om de recidive te meten, anderen gingen af op politie- en justitiecijfers. Soms bedroeg de observatieperiode zes maanden, soms duurde de follow-up zes jaar. Bij de meeste studies ontbrak het bovendien aan geschikt vergelijkingsmateriaal, vaak was er niet eens sprake van een controlegroep. Ook de doorlooptijd van het onderzoek was niet zelden te kort. Veel studies gaan daardoor mank aan te lage onderzoeks aantallen. De statistische power, het vermogen om een werkelijk bestaand effect te meten, is dan te laag.

Men kan zich afvragen of de werking van de onderzochte sancties wel betrouwbaar is vastgesteld. De problemen die het *Panel on Research on Rehabilitative Techniques* in de jaren '70 signaleerde, gelden misschien wel des te sterker voor het Nederlandse onderzoek. Als het gaat om de kennis van de werking van onze strafrechtelijke interventies zijn we in ons land in feite niet veel verder dan het punt waarop men zich bijna dertig jaar geleden in Amerika bevond. Het citaat op pagina 33 lijkt nog onverminderd van kracht; over speciale preventie weten we eigenlijk niets definitief. Is een taakstraf vanuit het oogpunt van speciale preventie effectiever dan een gevangenisstraf? Wordt er met het opleggen van bijzondere voorwaarden of het uitvoeren van penitentiaire programma's recidive voorkomen? We weten het niet.

2.5 Meta-analyse: een wijde blik op de uitkomsten van effectstudies

Een bijkomend probleem bij de evaluatie van strafrechtelijke interventies is dat er geen consensus bestaat over de wijze waarop de recidive moet worden berekend. Als de uitkomstmaat steeds anders is opgebouwd, kunnen de resultaten van studies niet zomaar bij elkaar worden opgeteld. Bij een groeiend aantal effectstudies wordt het dan al snel lastig om de uitkomsten van het onderzoek te overzien. Logan (1972) probeerde daar

wat aan te doen. Hij deelde de honderd effectstudies die hij bekeek aan de hand van de *aard* van de uitkomsten in een beperkt aantal categorieën in: de interventies waren ‘succesvol’, ‘niet succesvol’, of de werking bleef ‘onbekend’. Maar dit loste het probleem niet echt op, want de vraag hoe de interventies binnen één categorie zich tot elkaar verhielden, bleef onbeantwoord.

Indelingsproblemen als deze gelden overigens niet alleen voor de criminologie. In veel wetenschappen waarin de effectiviteit van interventies wordt onderzocht geldt dat de gehanteerde methoden van invloed zijn op het meetresultaat en dat verschillen in de gehanteerde methoden een directe vergelijking van de meetresultaten in de weg staan. In een debat over de effecten van psychotherapie ontwikkelde Gene V. Glass (1978) een kwantitatieve methode om de resultaten van 375 uiteenlopende effect-evaluaties samen te vatten. Hij ontwierp een overkoepelende maat, de *effect size*, die los van de exacte definitie van de gebruikte uitkomstmaten, uitdrukking geeft aan de richting en de omvang van de verbanden die in de afzonderlijke studies zijn gevonden. De effect size is in tegenstelling tot testen voor statistische significantie niet afhankelijk van de groepsgrootten. Ook minder omvangrijke studies kunnen dus bijdragen aan het beeld. Een belangrijk voordeel, zeker in kleine landen als Nederland, waar de onderzoeksgroepen welhaast per definitie minder groot zijn.

Glass (1976) noemde zijn methode ‘meta-analyse’. Hij berekende de gemiddelde effectgrootte van de 375 studies en liet zien dat psychotherapie overall wel degelijk gunstige effecten had (Smith & Glass, 1977). Het duurde even voordat de methode haar intrede deed in de criminologie. Garrett (1985) paste meta-analyse toe op een groep studies naar de effecten van ‘residential treatment’, gevolgd door Andrews et al. (1990) en Lipsey (1992), die meer dan 400 justitiële behandelprogramma’s bekeek. Inmiddels kennen we de Campbell Collaboration, een organisatie die bevordert dat er systematische reviews op sociaalwetenschappelijk terrein worden uitgevoerd en zo een verzamelplaats is geworden van de zogenoemde *what works*-literatuur.⁷

Een veelgebruikte maat voor de berekening van de effect sizes van vergelijkende recidivestudies is de *odds ratio*. De odds of wedverhouding is het quotiënt van twee waarschijnlijkheden: de kans dat een gebeurtenis (hier: de recidive) optreedt gedeeld door de kans dat die gebeurtenis niet optreedt. De odds ratio is de verhouding tussen de odds in de experimentele groep en de odds in de controlegroep. Zij kan als volgt worden uitgedrukt:

$$ES_{or} = (p_c/1-p_c)/(p_e/1-p_e),$$

7 Zie www.campbellcollaboration.org/. Met Wilson schreef Lipsey (2001) een handboek over meta-analyse.

waarin p_c staat voor de proportie recidivisten in de controlegroep en p_e voor de proportie recidivisten in de experimentele groep. De odds ratio neemt waarden aan vanaf 0. Een waarde kleiner dan 1 wil zeggen dat de controlegroep minder recidivisten herbergt en de interventie dus een averechts effect heeft gehad. Is de odds ratio groter dan 1, dan is er sprake van een positief effect: de experimentele groep telt minder recidivisten dan de controlegroep.

De odds ratio (en elke andere maat voor de berekening van effect sizes) stelt meta-analisten in staat de uitkomsten van effectstudies waarin de recidive op verschillende wijze is berekend, naast elkaar te zetten en onderling te vergelijken. Vaak gebeurt dit door middel van een zogeheten forest plot (zie figuur 2.1), waarmee in één oogopslag duidelijk wordt hoe de vlag er voor staat op het onderzoeksterrein dat met de geselecteerde effectstudies wordt afgebakend.

Figuur 2.1 Voorbeeld van een forest plot

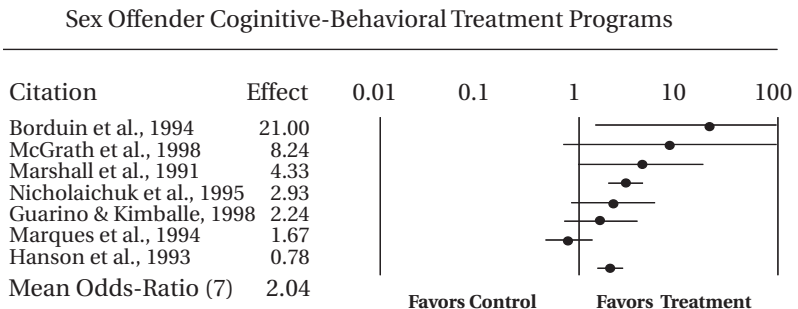


Figure 8.2. Forest plot showing odds ratio and 95 percent confidence interval for studies of sex offender cognitive-behavioral treatment programs.

Bron: MacKenzie (2006)

Het kwantificeren en samenvatten van de gemeten effecten, dat is het sterke punt van meta-analyse. De effect size maakt het mogelijk de speciaal preventieve werking van sancties en executiemodaliteiten onderling te vergelijken. Kritiek is er echter ook. Eysenck (1978), een bekende persoonlijkheidspsycholoog, was een van de eerste verklaarde tegenstanders van meta-analyse. Hij noemde de techniek een oefening in 'mega-silliness'. De beperking van de methode zit hem in de selectie van de primaire studies. De analyse is zo sterk als de keuze van de afzonderlijke effectevaluaties. Sharpe (1997) bespreekt drie bedreigingen van de validiteit van meta-analyse, kwesties waarop elke meta-analist een antwoord moet hebben. Het eerste is het *apples & oranges*-probleem. Meta-analyse kan gemakkelijk worden misbruikt om te vergelijken wat onvergelijkbaar is. De onderzoeker moet duidelijk omschrijven wat de gemene deler is van de

studies die hij groepeerd en zelfs dan dient hij rekening te houden met de rol van ‘moderatoren’. Dit zijn aspecten waarop de gekozen studies van elkaar verschillen, maar die wel gecorreleerd zijn aan de gemeten effect sizes (zie o.m. Weisburd, Lum & Petrosino, 2001). Het kan daarbij gaan om verschillen in de uitvoering van de onderzochte interventies, maar ook in de gehanteerde onderzoeks aanpak of tussen de groepen die met het onderzoek werden gevolgd. Moderatoren beïnvloeden de onderzoeksresultaten.⁸ Zo vond Lipsey (2003) dat interventies die deel uitmaakten van een speciaal pilottraject, over het algemeen betere resultaten te zien gaven dan interventies die onderdeel vormden van de reguliere gang van zaken. Kennelijk is naast de inhoud van de interventies ook de context waarin een interventie wordt uitgevoerd van belang. Bij startende projecten zet iedereen zijn beste beentje voor. Behoort het programma eenmaal tot een vaste routine, dan is het enthousiasme minder.

Het bestaan van moderatoren staat het trekken van inhoudelijke conclusies in de weg. Maar daar is wel een mouw aan te passen, want men kan ze in de meta-analyse betrekken. Als ze worden benoemd en gemeten, kunnen ze als extra onafhankelijke variabelen worden meegenomen in de analyse. Een goed voorbeeld van een dergelijke aanpak zien we bij Landenberger en Lipsey (2005). Zij onderzochten de rol van een aantal moderatoren bij programma’s voor cognitief-gedragsmatige therapieën en gebruiken regressie om hun invloed bloot te leggen. In de tweede generatie meta-analyses wordt het verband tussen interventies en de gemeten effect size via statistische technieken van storende invloeden ontdaan, waardoor het netto-effect van het type interventie beter kan worden ingeschat. In paragraaf 2.2 zagen we hoe verschillen tussen onderzoeksgroepen statistisch werden gecontroleerd. Nu blijkt dat zelfs ‘instroomverschillen’ tussen effectstudies via regressietechnieken teniet kunnen worden gedaan.

Een tweede klip die een meta-analist moet zien te omzeilen is het *file drawer*-probleem. Vergelijkende studies met niet-significante uitkomsten verdwijnen vaak in een spreekwoordelijke la. Ze worden minder snel gepubliceerd dan studies met aansprekende resultaten. Als de onderzoeker er niet in slaagt om ook evaluaties op te nemen die buiten de internationale tijdschriften bleven, kan het zijn dat het effect van de onderzochte interventie wordt overschat. Elke meta-analist zal dus zijn best moeten doen om ‘grijze’ literatuur op te sporen. Ook beleidsrapporten, congrespapers, dissertaties en dergelijke moeten in het onderzoek worden opgenomen. De kans bestaat dat deze niet-gepubliceerde studies van mindere kwaliteit zijn. Opname brengt dus wel een gevaar met zich mee.

De derde bedreiging van de validiteit van meta-analyses wordt aangeduid met ‘*garbage in, garbage out*’ (Eysenck, 1978). Als de primaire studies van

8 Naast moderatoren worden ook mediators onderscheiden. Een moderator beïnvloedt de relatie tussen een interventie en de uitkomst. Een mediator maakt onderdeel uit van het mechanisme dat tot de uitkomst leidt en *verklaart* zo het verband tussen interventie en uitkomst (Baron & Kenny, 1986). Een voorbeeld van een mediator is het IQ van het kind bij het meten van het succes van een leer methode.

slechte kwaliteit zijn, kan een meta-analyse dat niet verhelpen. Sommige auteurs werpen daarom een drempel op en laten alleen studies toe met een experimenteel design. Anderen scoren de kwaliteit van de evaluaties en nemen deze score mee als moderator variabele in de regressieanalyse. De meta-analist gaat dan na of de score van invloed is op de effect size. Het voordeel van deze laatste strategie is dat er meer studies in de analyse kunnen worden opgenomen en het beeld van de effecten van een interventie dus completer is.

Sinds de eerste aanzetten van Glass heeft meta-analyse zich als onderzoeksmethode verder ontwikkeld. Sharpe (1997) geeft aan welke oplossingen er zijn bedacht voor de problemen die inherent zijn aan het gebruik van deze techniek. De kritiek van Pawson en Tilly (1997) gaat verder en lijkt fundamenteeler van aard. Net als de aanhangers van de Campbell Collaboration staan zij *evidenced based* overheidsbeleid voor. Ook zij ijveren voor systematic reviews om de effecten van de gepleegde interventies vast te stellen. Maar Pawson en Tilly beperken zich niet tot de studies met een experimenteel of quasi-experimenteel design. Bij hun analyse worden ook andere bronnen betrokken. Hun opvatting van 'oorzakelijkheid' verschilt met die van de aanhangers van de Campbell Collaboration. De experimentele traditie is in hun ogen in essentie een black box-benadering. Er is input (de interventie) en er is output (de uitkomst), maar wat er zich tussen beide afspeelt doet er weinig toe. Volgens Pawson en Tilly gaan een kwantitatieve meta-analyse en de experimentele traditie waarop zij berust voorbij aan de sociale en maatschappelijke context waarin de interventies worden ingezet. Een interventie is niet een onveranderlijk geheel met een vaststaande werking, maar een dynamisch onderdeel in een reeks van processen die afhankelijk van uiteenlopende omstandigheden al dan niet uitmonden in hernieuwd crimineel gedrag. Zij kiezen voor een benadering waarin de aandacht uitgaat naar de inhoud van de interventies en de condities waaronder deze verondersteld worden te werken. Bij een 'realistische evaluatie', zoals zij hun benadering in navolging van de wetenschapsfilosoof Harré (1985) noemen, gaat het om het blootleggen van onderliggende theoretische noties en het ontrafelen van de mechanismen die met de interventies in werking worden gezet. De vraag 'Wat werkt?' wordt niet beantwoord door de berekening van een score die is ontdaan van alle contextuele informatie. De vraag waar het in hun benadering werkelijk om gaat, is: 'Wat werkt, voor wie en onder welke omstandigheden?' Die vraag vergt een aanpak waarbij niet alleen studies van de hoogst mogelijke wetenschappelijke kwaliteit in de analyse worden betrokken, maar waarbij alle mogelijke bronnen behulpzaam kunnen zijn. Ook in zijn laatste boek laat Ray Pawson (2006) er geen misverstand over bestaan. Hij vindt meta-analyse geen toereikende methode voor de accumulatie van kennis:

‘... such a process of simplification, standardization and aggregation does not produce any ... enduring truths. ... One needs to be clear about just what is omitted, ... meta-analysis eliminates most of the evidence that is capable of telling us how interventions work and how we might account for their differential effectiveness. What is detritus for meta-analysis is, however, explanatory gold dust for the realist.’ (p. 72)

De realistische school van Pawson en Tilly wordt vaak gezien als het spiegelbeeld van de onderzoekscultuur van de experimentele wetenschapstraditie. Toch zal blijken dat de standpunten van beide kampen gemakkelijk met elkaar kunnen worden verzoend. De ‘zachte’, kwalitatieve benadering van Pawson en Tilly vormt alleen op papier een contrast met de ‘harde’, wetenschappelijk lijn van Farrington, Wilson en Lipsey. In de praktijk vullen zij elkaar perfect aan. Meer hierover in hoofdstuk 5.

Box 2.2 What works? Een voorlopige stand van zaken

Met het aantal primaire effectstudies groeit ook het aantal meta-analyses. Wartna, Baas en Beenackers (2004) maakten een samenvatting van de uitkomsten van 29 buitenlandse overzichtsstudies, samen goed voor een pool van meer dan duizend afzonderlijke evaluatiestudies. In navolging van MacKenzie (2002) deelden zij de programma's en interventies in met behulp van de drie hoofdstrategieën voor speciale preventie: afschrikking, incapacitatie en resocialisatie (zie ook hoofdstuk 1, figuur 1.1). Langs deze weg gingen de auteurs na welke resultaten bepaalde varianten van deze strategieën te zien hebben gegeven, in termen van recidivereductie welteverstaan. Andere strafdoelen, zoals vergelding en generale preventie, werden niet meegenomen. De interventies die in de 29 overzichtsstudies aan bod kwamen, hadden bovendien alleen betrekking op volwassen, mannelijke justitiabelen.

Afschrikking door het opleggen van zwaardere of langere straffen

Interventies die gebaseerd zijn op afschrikking door middel van het opleggen van zware of lange straffen lijken niet te werken, dat wil zeggen: ze verlagen de recidive niet (46, 47, 88)*. Dit houdt overigens *niet* in dat afschrikking op zich niet effectief kan zijn. Er is bijvoorbeeld onderzoek waaruit blijkt dat het opleggen van boetes – ook deze interventie stoelt op afschrikking – in vergelijking met andere sancties een lagere recidive te zien gaf (28). Wel heeft onderzoek aangetoond dat van het (verder of langduriger) inperken van de bewegingsvrijheid op zich geen speciaal preventieve werking uitgaat. Puur beredeneerd vanuit het oogpunt van recidivereductie zijn (lange) gevangenisstraffen dus niet effectief.

Afschrikking door verhoogd toezicht

Arrangementen van vrijheidsbeperking waarin door het uitoefenen van extra controle en toezicht (afschrikking) of gedeeltelijke insluiting (incapacitatie)

getracht wordt toekomstig delictgedrag te voorkomen, hebben tot dusver weinig succes gehad. Het gaat hier om vormen van intensief reclasseringstoezicht al dan niet in combinatie met aanvullende vormen van toezicht (77, 79, 92) of elektronisch toezicht c.q. huisarrest (80, 81, 91, 93). Ook het onderzoek naar varianten van halfopen inrichtingen laat op zijn best gemengde resultaten zien (64).

Afschrikking door verhoogd toezicht gecombineerd met behandeling

Wel zijn er aanwijzingen dat als men de beperking van de bewegingsvrijheid combineert met elementen van behandeling, interventies effectief kunnen zijn. Dit geldt zowel bij drugsverslaafde daders in sterk gestructureerde therapeutische gemeenschappen (36, 37, 54, 55, 57, 58) en *drug courts* (95, 96) als voor personen die (herhaaldelijk) zijn veroordeeld voor het rijden onder invloed (90) en gedetineerden in halfopen inrichtingen (71). Bij deze interventies is er naast het verhoogde toezicht ruimte voor counseling, andere vormen van therapie en het bieden van hulp en steun.

Incapacitatie of onschadelijkmaking

Onderzoek naar verschillende vormen van incapacitatie of onschadelijkmaking laat gemengde resultaten zien. Naar de effecten van incapacitatie door insluiting is goed beschouwd nog weinig onderzoek gedaan. Probleem bij de evaluatie van deze strategie is dat vooraf vaak niet goed kan worden ingeschat welke daders in welke mate en voor hoe lang zouden recidiveren als zij niet zou worden opgesloten. Generale incapacitatie, het opsluiten van daders ongeacht hun recidiverisico na afloop van de detentie, is effectief, maar alleen tegen zeer hoge kosten (38). De doelmatigheid van *selectieve* insluiting is afhankelijk van de mate waarin men erin slaagt de daders met het hoogste recidiverisico vroegtijdig te identificeren (39). Het toedienen van methadon aan drugsverslaafden is een indirecte vorm van onschadelijkmaking. Het methadon neemt de behoefte aan opiaten weg, waardoor de verslaafden minder geneigd zouden zijn tot het plegen van delicten om de drugs te kunnen bekostigen. Deze strategie is onderzocht en blijkt over het algemeen niet effectief (27, 53).

Resocialisatie door educatie

Onderwijsprogramma's voor gedetineerden kunnen effectief zijn. De meeste onderzoeken naar basiseducatie of meer specifieke scholing laten gunstige (11, 63) of veelbelovende (30, 31, 34, 42, 43) resultaten zien.

Resocialisatie via arbeidstoeleiding

Programma's voor gedetineerden gericht op het verkrijgen van werk in aansluiting op de detentie, geven een gemengd beeld. De meeste overzichtsstudies melden positieve (32, 48, 66) of veelbelovende resultaten (44, 49, 67), maar er is ook onderzoek waarin onder de deelnemers geen recidivereductie kon worden aangetoond (72) of waarvan de resultaten onduidelijk blijven (33). Uit de grootste metastudie op dit terrein (49) komt

naar voren dat kortere arbeidstoeleidingsprogramma's effectiever lijken te zijn dan langere.

Resocialisatie via het bijbrengen van discipline

In Amerika is veel onderzoek gedaan naar de effecten van *boot camps*, inrichtingen voor jeugdige en (jong)volwassen gedetineerden waarin een strakke discipline heerst en het accent van het activiteitenprogramma ligt op het bijbrengen van orde en tucht, sport en lichamelijke arbeid. Nergens blijkt dat dergelijke programma's tot recidivevermindering hebben geleid (29, 52, 59).

Resocialisatie door behandeling

Er is veel buitenlands onderzoek naar de effecten van behandelprogramma's. De meeste overzichtsstudies laten gunstige resultaten zien. Vooral programma's die cognitief-gedragsmatig van aard zijn leiden vaak tot recidivevermindering, zowel bij volwassen daders in het algemeen (2, 3, 12, 13, 20, 21, 22, 26, 86) als bij zedendelinquenten (14, 15, 41, 82, 87) en drugsverslaafde delinquenten (8, 24). Het gaat hier om interventies die gericht zijn op de ontwikkeling en training van sociale en cognitieve vaardigheden. Er zijn ook overzichtsstudies waarin de resultaten van dergelijke interventies niet onverdeeld gunstig, maar in elk geval veelbelovend waren (23, 25, 35, 51, 61, 62, 70, 85). Programma's gericht op louter praktische vaardigheden (7) of gedragstherapieën die afhankelijk zijn van een systematische (lees: voortdurende) beloningsstructuur (18,19), lijken minder kans van slagen te hebben. Behandelprogramma's voor drugsverslaafden die bestaan uit groepstherapeutische gesprekken zonder verdere structuur, werken niet (60).

Bovenstaande opsomming is in feite een samenvatting van samenvattingen en geeft uiteraard alleen globaal aan hoe strafrechtelijke interventies moeten worden ingericht als men streeft naar reductie van herhaalde criminaliteit. Overzichtsstudies en meta-analyses leveren niet meer dan een spoor op, een aanwijzing in welke richting men het wel of juist niet moet zoeken. MacKenzie (2006) geeft een recenter overzicht van de what works-literatuur. Zij beperkt zich niet tot interventies voor volwassen mannen, maar gaat ook in op het succes van programma's voor andere categorieën van daders, zoals jeugdigen, plegers van huiselijk geweld, zedendelinquenten en criminele drugsverslaafden.

* Bijlage 1 bevat een tabel met het ruwe materiaal waarop de uitspraken in deze box zijn gebaseerd. De nummers tussen haakjes corresponderen met de rijen van de tabel. Daarin is terug te vinden wat er over de werking van bepaalde typen interventies bekend is. De tabel wordt voorafgegaan door een leeswijzer.

2.6 Monitoring: overzicht door standaardisering

De scherpe kritiek van Pawson en Tilly heeft niet verhinderd dat meta-analyse is uitgegroeid tot een dominante methodologie binnen de criminologie. Het aantal artikelen in internationale tijdschriften waarin de

resultaten van een kwantitatieve meta-analyse worden besproken, neemt hand over hand toe.⁹ Dit heeft alles te maken met het herwonnen geloof in resocialisatie. Interventies kunnen werken, zo blijkt uit onderzoek, en meta-analyses zijn nodig om dat te laten zien. Naast de kwantitatieve meta-analyses van de aanhangers van de Campbell Collaboration en de kwalitatieve research syntheses van hun rivalen van de realistische school, dient zich echter nog een andere methode aan die kan bijdragen aan de integratie van de uitkomsten van evaluatiestudies. In meer en meer landen wordt recidiveonderzoek tegenwoordig op nationale schaal toegepast. De recidive wordt in uiteenlopende studies op overeenkomstige wijze gemeten, dat wil zeggen volgens een onveranderlijke definitie en over een gelijke observatieperiode. Het voordeel van het gebruik van een vaste meetmethode is evident. De uitkomsten van afzonderlijke evaluaties komen automatisch in een bredere context terecht. Verschillen in meetresultaten kunnen niet worden toegeschreven aan de wijze waarop de uitkomsten werden geproduceerd. Het is dus niet langer nodig om, zoals bij een kwantitatieve meta-analyse, achteraf een overkoepelende maat te ontwikkelen om de uitkomsten van verschillende effectstudies te vergelijken. Door de recidivemetingen te standaardiseren is vergelijkbaarheid, althans in methodologische zin, vooraf gegarandeerd.

Standaardisering gaat gepaard met schaalvergroting. Recidiveonderzoek vindt steeds vaker plaats op nationaal niveau. Figuur 2.2 laat zien welke Europese landen anno 2006 beschikten over landelijke cijfers. Schaalvergroting is een direct gevolg van de automatisering van de registratiesystemen van politie en justitie. De ruwe gegevens zijn beter bereikbaar geworden. De dataverzameling en -verwerking is enorm versneld. In Engeland deden deze ontwikkelingen zich het eerst voor. Britse onderzoekers hebben veel betekend voor het beleidsgerichte recidiveonderzoek. In Engeland en Wales werd tot voor kort standaard gebruikgemaakt van de *Offenders Index*, een database met justitiegegevens die in 1991 werd geautomatiseerd.¹⁰ We zien dit bijvoorbeeld terug in het eerder aangehaalde werk van Lloyd, Mair en Hough (1994), waarin een beschrijving wordt gegeven van de recidive in een landelijke steekproef van bijna 18.000 veroordeelde daders. De auteurs splitsen de cijfers uit naar soort afdoening (gevangenisstraf, taakstraf, reclasseringstoezicht en reclasseringstoezicht met speciale voorwaarden) en schatten de relatieve effecten van de sancties met behulp van een logistisch model. In geen van de sanc-

9 In Nederland ontbreekt het nog aan een kwantitatieve meta-analyse. Het WODC werkt aan een bestand waarop een dergelijke analyse kan plaatsvinden, zie verder hoofdstuk 5. Wel zijn er overzichtsstudies en -artikelen verschenen over de effectiviteit van strafrechtelijke interventies (D'Anjou, De Jonge & Van der Kaaden, 1975; Van den Heuvel, 1991, 1995; Hessing & Van Koppen, 1997; Doek & Bac, 1998; Beenackers, 2000, 2001; Wartna, Baas & Beenackers, 2004) en van preventieprojecten in het algemeen (Polder & Van Vlaardingen, 1992; Wittebrood & Van Beem, 2004; Klein Haarhuis et al., 2005; Van Noije & Wittebrood, 2008).

10 Inmiddels is men overgestapt op het gebruik van de *Police National Computer*, omdat deze pleegdata bevat en het tijdstip van de recidive dus nauwkeuriger kan worden vastgesteld.

tiegroepen was het verschil tussen de verwachte en de feitelijke recidive erg groot. De auteurs sluiten niet uit dat er lokale verschillen zijn in de effectiviteit van de sancties, maar op landelijk niveau vallen de cijfers zo ongeveer uit zoals mocht worden verwacht afgaand op de achtergronden van degenen die de sancties ondergingen.

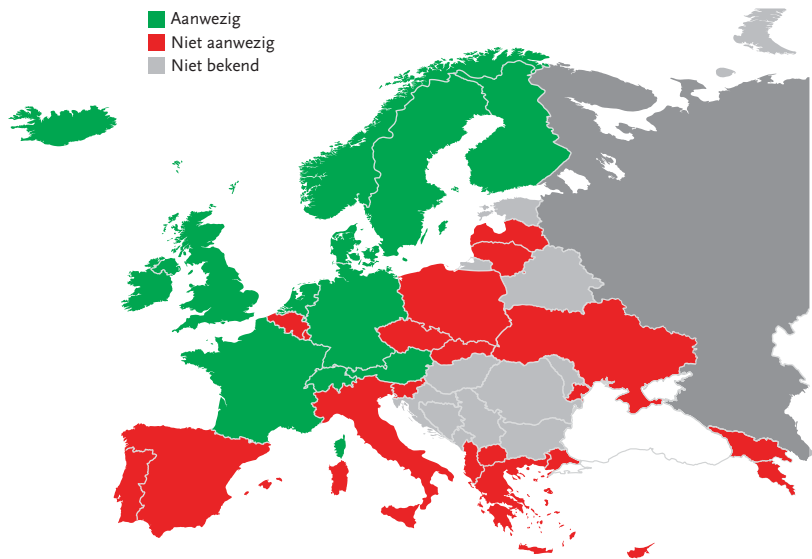
Niet alleen het evaluatieonderzoek, ook het predictieonderzoek profiteerde van de automatisering van de gegevensbronnen. Tot dat moment was ook dit type onderzoek afhankelijk van handmatige invoer van de ruwe gegevens, de predictie van recidive was gebaseerd op kleine steekproeven en beperkte zich tot specifieke daderpopulaties. De beschikbaarheid van de Offenders Index veranderde dat. Op basis van een grote, landelijke steekproef ontwierpen Copas, Marshall en Tarling (1998) een algemeen risicotaxatie-instrument. Met deze schaal, de *offender group reconviction scale* (OGRS), kan voor elke verdachte op grond van een beperkt aantal achtergrondkenmerken – de sekse, de leeftijd, het type delict en de omvang van het strafrechtelijke verleden van een verdachte – een inschatting worden gemaakt van het recidivegevaar. Net als Lloyd, Mair en Hough pasten Copas, Marshall en Tarling logistische regressie toe. De schaal geeft voor elk individu de statistische kans dat hij binnen twee jaar opnieuw in de fout gaat. De berekende score is gelijk aan het geschatte recidivepercentage in de groep daders met dezelfde combinatie van achtergronden als de verdachte. Medewerkers van de reclassering maken inmiddels standaard gebruik van de OGRS in hun rapportages aan de rechter. Het is een voorbeeld van een actuarieel instrument. De OGRS is eenvoudig en snel toe te passen. Hij geeft weliswaar een grove inschatting van het recidiverisico; men moet naast de voorspellers die zijn opgenomen in het model altijd ook de individuele omstandigheden van de verdachte in aanmerking nemen. Overall levert de schaal echter een goede voorspelling op.¹¹

De beschikbaarheid van digitale gegevens heeft ook de omloopsnelheid van het recidiveonderzoek vergroot. In een aantal landen is inmiddels sprake van *monitoring*, het herhaald meten van de terugval van grote groepen daders. Met een zekere regelmaat publiceert de onderzoeksafdeling van het Engelse ministerie van Justitie rapporten waarin de ontwikkeling van de landelijke recidive in kaart wordt gebracht. Door de standaardmetingen te herhalen kunnen trends in de cijfers worden bestudeerd. Men werkt daarbij met 'gecorrigeerde' cijfers. Er wordt rekening gehouden met verschuivingen in de daderpopulatie op kenmerken als leeftijd, sekse, type delict en delictgeschiedenis; dezelfde variabelen als in het voorspellingsmodel van de OGRS. De bijgestelde cijfers leveren een

11 De predictieve validiteit van (varianten van) de OGRS is in diverse settings onderzocht. De gerapporteerde AUC-waarden liggen tussen 0,72 en 0,81 (Maden et al., 2005; Francis, Soothill & Humphreys, 2005).

indruk op van de netto-ontwikkeling van de recidive. De Britse regering gebruikt ze bij de verantwoording van het strafrechtelijke beleid. In 2002 en 2004 werden ook in Engeland streefcijfers opgesteld. Men nam zich voor ten aanzien van zowel jeugdige als volwassen daders het landelijke recidiveniveau in een paar jaar tijd met 5 tot 10% te verlagen (Command Paper, 2005).

Figuur 2.2 Veertien Europese landen met recidivecijfers op nationale schaal



Bron: Wartna & Nijssen, 2006

Ook in de Scandinavische landen en in Schotland en Noord-Ierland vinden inmiddels periodiek nationale recidivemetingen plaats. In Australië en Canada wordt gewerkt aan standaardisering (Payne, 2007; Farrington & Davies, 2007). In de Verenigde Staten deden Langan en Levin (2002) grootschalig onderzoek onder ex-gedetineerden. Deze studie droeg echter geen landelijk karakter; het bleek te lastig te zijn om de ruwe gegevens uit de verschillende regio's op een handzame manier aan elkaar te koppelen. De teller bleef daardoor op vijftien van de vijftig staten staan. In Nederland heeft men van problemen met de dataverzameling hoegenaamd geen last. Het ministerie van Justitie beschikt over een onderzoeksbestand dat vergelijkbaar is met de Engelse Offenders Index. Sinds 2005 kent ons land de zogeheten Recidivemonitor, een project waarin op basis van dit onderzoeksbestand herhaalde metingen worden verricht onder bijna alle personen die recent met justitie in aanraking zijn gekomen. In het volgende

hoofdstuk gaan we uitgebreid op de ontwikkeling van de Recidivemonitor in.

2.7 Slot: we weten weinig, maar wel steeds meer

Van de jeugdige gedetineerden uit het Engelse Warwick County van de 19e eeuw tot aan het monitoren van alle personen met een justitiecontact: meer dan 150 jaar wordt er nu met behulp van recidivecijfers evaluatieonderzoek gedaan. Wat weten we in Nederland van de speciaal preventieve werking van onze straffen en maatregelen? Niet erg veel. Het onderzoek dat in ons land is verricht, kent vele tekortkomingen en biedt weinig houvast. De kennis die we hebben opgebouwd komt voornamelijk uit het buitenland. Kritisch beschouwd lijkt het belangrijkste dat in ons land op dit vlak is gebeurd, de automatisering van de gegevensverzameling te zijn. Misschien komt het recidiveonderzoek hierdoor in een stroomversnelling terecht.

Waarom weten we nog zo weinig over speciale preventie? Komt dit misschien omdat de resocialisatie van justitiabelen in onze samenleving van ondergeschikt belang is? Is daarom de omvang en de kwaliteit van het Nederlandse effectonderzoek achtergebleven? Recht doen, het onrecht vergelden en een ongestoorde tenuitvoerlegging; dat is wat wij primair van het strafrecht verwachten. Resocialisatie van justitiabelen is hoogstens theoretisch van belang. Feitelijk echter worden de kansen om via het strafrecht te corrigeren wat in de tijd daarvoor fout is gegaan, gering geacht. En het voorbeeld van de evaluatie van 'dagdetentie' in box 2.1 leert dat wanneer men niettemin kansen tot resocialisatie ziet, men niet altijd wil weten of de interventie werkt. Soms is de theorie beter.

Een andere verklaring voor het gebrek aan kennis op dit terrein is dat het lastig is om goed onderzoek te doen naar de effectiviteit van strafrechtelijke interventies. Kernprobleem bij het uitvoeren van dit soort evaluaties is het vinden van een *counterfactual*: wat zou er zijn gebeurd als deze daders de interventie niet hadden ondergaan? We hebben gezien dat door de jaren heen verschillende middelen zijn ingezet om dit probleem het hoofd te bieden. Verschillende varianten van statistische controle, quasi-experimentele designs en gerandomiseerde toewijzing; alle bedoeld om de interne validiteit van het onderzoek te verhogen. In veel Nederlands onderzoek ontbreekt het echter aan geschikt vergelijkingsmateriaal, waardoor causale uitspraken over de effecten van de interventie onmogelijk zijn. Het onderzoek naar de speciaal preventieve werking van strafrechtelijke interventies kent meer problemen. Wat te denken van de kritiek van Pawson en Tilly? Zij wijzen op het belang van het bestuderen van de context waarin de interventie plaatsvindt en stellen dat zolang niet wordt gekeken naar de mechanismen die aan de interventie ten grondslag liggen, nooit duidelijk zal worden wat haar effecten zijn. Als dit laat-

ste klopt, dan is dit misschien wel de doodsteek voor het experimentele model. Is het vergelijken van uitkomsten in twee groepen wel de geschikte manier om de werking van een interventie vast te stellen? In een experiment gaat het inderdaad niet om de vraag *hoe* een interventie werkt. Het gaat puur om de vaststelling *dat* het werkt. Maar betekent dit dan ook dat we met een experiment de verkeerde vraag stellen? Nee. De tweede generatie meta-analyses waarin multivariate technieken worden gebruikt om de effect sizes te analyseren, laat zien dat ook in de experimentele traditie plaats wordt ingeruimd voor contextvariabelen. Meta-analyses vatten de uitkomsten van experimenten samen, maar ze doen meer dan dat. Ze maken het mogelijk systematisch onderzoek te doen naar de rol van moderatoren.

Het wordt steeds duidelijker aan welke voorwaarden strafrechtelijke interventies moeten voldoen om succesvol te zijn. De lijst van kwaliteitscriteria die op basis van de literatuur wordt samengesteld, breidt zich uit. Lange tijd hanteerden de kopstukken van de what works-benadering drie basisprincipes: *needs*, *risk* en *responsivity* (zie o.m. Andrews, Bonta & Hoge, 1990; Andrews et al., 1990; Andrews & Bonta, 2003). Inmiddels formuleert Andrews (2001) er achttien (zie ook NIC, 2004).¹² We weten misschien weinig over de effectiviteit van strafrechtelijke interventies, maar de kennis neemt wel toe. Er raakt steeds meer bekend over hoe interventies moeten worden ingericht om de kansen op terugval zo klein mogelijk te houden. Meta-analyses spelen daarbij een beslissende rol en dat geldt dus ook voor de vergelijkende studies die daar onderdeel van uitmaken.

De kritiek van Pawson en Tilly is deels achterhaald. De context waarin interventies worden uitgevoerd, is inderdaad bepalend voor het succes en deze wordt daarom ook als factor meegenomen in het onderzoek. Wel kan men zich afvragen of dit voldoende systematisch gebeurt. In de effectstudies die onderdeel uitmaken van meta-analyses wordt in de regel weinig aandacht besteed aan de inhoud en opzet van de interventie waarvan de effecten worden onderzocht. De typische publicatievorm van een effectstudie, een artikel in een wetenschappelijk tijdschrift, wijdt doorgaans slechts enkele alinea's aan de ideeën die aan de interventie ten grondslag liggen en de wijze waarop deze in de praktijk worden uitgewerkt. De aandacht gaat vooral uit naar de *outcome*, de effecten van de interventie in termen van de recidive van de deelnemers na afloop van hun deelname aan het bewuste programma. Deze preoccupatie met de uiteindelijke

12 Het behoeftebeginsel stelt dat effectieve interventies zich richten op de behandeling van factoren die samenhangen met het *criminele* gedrag. Het risicoprincipe houdt in dat de intensiteit van de interventie evenredig moet zijn met kans dat er recidive optreedt. Daders met een laag recidiverisico zijn alleen gebaat bij een minimale interventie, daders met een hoog recidiverisico vergen een meer intensieve aanpak. Het responsiviteitsprincipe dicteert dat de opzet van het programma moet aansluiten bij de leerstijl en motivatie van de deelnemers. In het compendium voor 'effective correctional programming' noemt Andrews (2001) in totaal achttien eisen. Naast *needs*, *risk* en *responsivity* wordt onder meer ingegaan op het belang van program integrity en de aanwezigheid van nazorgfaciliteiten (zie verder hoofdstuk 5).

effecten van de interventie is wel tekenend voor het experimentele model. Men wil snel door naar de uitkomsten van de recidivemeting, want daar gaat het uiteindelijk om. Maar dat is dan ook precies wat het is: een preoccupatie. Aangemoedigd door de opdrachtgevers van het onderzoek die 'snel duidelijkheid willen', richten onderzoekers zich vaak te gehaast op 'waar het allemaal om draait'. De vraag hoe de interventie werkt, komt er vaak bekaaid af. Daar hebben Pawson en Tilly een punt. Er is binnen de experimentele traditie echter geen dwingende reden om deze vraag te negeren. Sterker nog, Farrington, die bekendstaat als de exponent van de experimentele benadering, erkent het belang van een goede beschrijving van de interventie:

'It is important in all research to consider what might be the causal chain between an intervention and an outcome and to measure possible mediating factors' (Farrington, 2003: 161).

De experimentele traditie met haar 'starre manier van meten' en haar 'nadruk op interne validiteit' verzet zich niet tegen een kwalitatieve benadering waarin de werking van de interventie conceptueel wordt geanalyseerd. Integendeel, een dergelijke analyse dient vooraf te gaan aan het meten van de outcome en dát – zo luidt hier de stelling – wordt te weinig en te weinig systematisch gedaan. In hoofdstuk 5 zal worden besproken hoe beide benaderingen, de kwantitatieve en kwalitatieve benadering, met elkaar kunnen worden gecombineerd. Dan gaan we dieper in op de voorwaarden voor goed evaluatieonderzoek. Meta-analyses zijn zo goed als de effectstudies waarop zij zijn gebaseerd. Het is dus zaak om de kwaliteit van de evaluaties te verhogen. Hoofdstuk 5 verschaft hiervoor een kader. Voordat we hieraan toekomen, richten we ons echter meer bepaald op het meten van de recidive. Zoals aangegeven, kent ons land sinds kort de Recidivemonitor, een project waarin periodiek gestandaardiseerde metingen plaatsvinden. Kleine en selecte steekproeven behoren tot het verleden, met dit instrument kunnen *alle* justitiabelen onderzoeksmatig worden gevolgd. In hoofdstuk 3 gaan we in op de totstandkoming van de Recidivemonitor. Het hoofdstuk daarna, hoofdstuk 4, bevat een greep uit de meetresultaten. Voor een aantal dadergroepen gaan we na hoe de recidive zich de laatste jaren heeft ontwikkeld.

3 Meten en maten van recidive

Opzet en verantwoording van de Recidivemonitor

Het WODC verricht recidiveonderzoek op nationale schaal, de metingen vinden steeds op dezelfde wijze plaats. Dit hoofdstuk beschrijft de procedures die in het kader van de Recidivemonitor zijn ontwikkeld. Eerst echter wordt het recidivebegrip verder ontleed. Paragraaf 3.1 gaat in op de vraag welke facetten het begrip kent en hoe deze kunnen worden gemeten. Paragraaf 3.2 belicht de gegevensbronnen die in Nederland voor dit type onderzoek beschikbaar zijn. In paragraaf 3.3 bespreken we keuzes die zijn gemaakt om tot een vaste aanpak voor het recidiveonderzoek te komen. De ontstaansgeschiedenis van de Recidivemonitor komt aan de orde en er zijn illustraties van de statistieken die standaard worden berekend. Paragraaf 3.4 bevat een opsomming van de factoren die bepalend zijn voor de omvang van de terugval. Op zichzelf is het meten van recidive een betrekkelijk theoriearm proces. Om de hoogte van de recidive te verklaren moet men de uitkomsten van het onderzoek in verband brengen met causale factoren die zijn ingebed in een criminologische theorie. De Recidivemonitor zelf verklaart niet, hij levert een beschrijving – een kwantitatief overzicht – van de terugval op. Om de ontwikkeling in het recidiveniveau beter te kunnen beschrijven worden de ruwe cijfers net als in Engeland wel gecorrigeerd voor verschuivingen in de samenstelling van de onderzoeksgroepen. Bovendien wordt rekening gehouden met registratie-effecten. Paragraaf 3.5 laat zien hoe dit in zijn werk gaat.

3.1 Anatomie van het recidivebegrip

In hoofdstuk 2 zagen we dat het onderzoek naar het verloop van criminele carrières vier dimensies kent: deelname, frequentie, ernst en duur (Blumstein et al., 1986). Een recidivemeting is een bijzondere vorm van onderzoek naar criminele carrières, een vorm van onderzoek waarbij het draait om de herhaling of voortzetting van het delinquente gedrag. Alle personen in het onderzoek zijn ‘deelnemers’. Er is in principe altijd een *uitgangsdelict*, het delict naar aanleiding waarvan de persoon in het onderzoek is opgenomen. Dit is niet noodzakelijk het eerste delict in de criminele carrière, een deel van de personen heeft mogelijk al een langere voorgeschiedenis. Maar het uitgangsdelict vormt wel altijd het startpunt voor het onderzoek. Vanaf dit punt bezien, wil men in essentie vier dingen van de onderzoekspersonen weten: óf ze hebben gerecidiveerd, hoe vaak ze hebben gerecidiveerd, hoe snel de recidive plaatsvond en hoe ernstig deze was. Het recidivebegrip kent met andere woorden vier eigen dimensies, te weten: prevalentie, frequentie, snelheid en ernst.

1 *Prevalentie*

Bij het vaststellen van de prevalentie van de recidive gaat het om de vraag of er sprake is van een nieuw delict. Deze vraag kan voor elk individu simpelweg met 'ja' of 'nee' worden beantwoord; op individueel niveau is prevalentie dus een dichotome variabele. Op groepsniveau wordt de prevalentie uitgedrukt als het percentage van de groep dat een nieuw delict heeft gepleegd. In veel studies is dit overigens de enige recidivemaat die wordt gebruikt.

Recidiveonderzoek kent een aantal vaste bouwstenen. Zo geldt in elke recidivestudie een *termijn* waarover wordt nagegaan of er nieuwe delicten zijn. Hoe langer deze periode duurt, des te hoger zal het percentage recidivisten zijn. Als er meer tijd verstreken is, is er immers meer gelegenheid geweest om in de fout te gaan. De personen waren langer 'at risk'. Het beginpunt van de observatieperiode kan de dag zijn waarop het uitgangsdelict is gepleegd, maar evengoed de datum waarop daders de straf werd opgelegd, of de dag waarop de betrokkene uit de gevangenis werd ontslagen. Meestal is de duur van de observatieperiode voor iedere persoon in het onderzoek gelijk, bijvoorbeeld een periode van twee of vier jaar. Maar soms doen zich verschillen voor in het begin- en/of eindpunt van de meetperiode. De duur van de follow-up hoeft dus niet voor iedereen hetzelfde te zijn.

Voordat de recidive kan worden vastgesteld dient de onderzoeker aan te geven wat hij daar precies onder verstaat. Het begrip moet worden geoperationaliseerd. Dit betekent dat het *bereik* van de recidive moet worden bepaald. Gaat het om herhaling van delinquent gedrag zoals omschreven in de wet, of is men ook geïnteresseerd in ander ongewenst of normafwijkend gedrag? De onderzoeker dient het begrip te specificeren. Indien het gaat om de bij wet verboden gedragingen, dan is de volgende vraag of alle misdrijven en overtredingen meetellen en op welke wetten men zich richt. Beperkt men zich tot het Wetboek van Strafrecht of beschouwt men een overtreding van de milieuwetgeving ook als een geval van recidive? De keuze van het bereik is van invloed op de resultaten van het onderzoek. Hoe meer gedragingen binnen de definitie vallen, des te hoger zullen de recidivepercentages zijn. Als men snelheidsovertredingen meeneemt, is bijna iedere volwassene die een motorrijtuig bestuurt een recidivist. Wanneer men zich beperkt tot verkrachting en moord, bijna niemand. Belangrijk is ook het *domein* waaruit men de recidivegegevens betreft. Het domein is het 'universum' van delicten dat men onderzoekt, het geeft globaal aan waar de grenzen liggen. De omvang van het onderzoeksdomein hangt af van de positie van de 'informatiehouder', dat is degene die weet (of aanneemt) dat de bedoelde gedragingen hebben plaatsgevonden. De dader of een slachtoffer is een informatiehouder. De politie ook. De politie krijgt niet alle daders in het vizier en weet ook niet alles van de verdachten die zij wel te spreken krijgt. De officier van justitie is ook een informatiehouder, hij vormt alleen een oordeel over de delicten die door

de politie aan hem worden doorgegeven. De rechter op zijn beurt doet alleen uitspraak over de delicten die door het Openbaar Ministerie (OM) onder zijn aandacht zijn gebracht. Het domein waarin politie en justitie actief zijn en informatie vastleggen, is onoverkomelijk beperkt. Indien men *alle* delicten die zijn gepleegd in kaart wil brengen, kan men zich niet verlaten op de gegevens van politie en justitie. Beperkt men zich echter tot die delicten die zijn ontdekt, gemeld en opgelost, of waarbij de verdachte is vervolgd en berecht, dan kan men gebruikmaken van politie- of justitiebronnen.

2 *Frequentie*

Naast de vraag of er recidive heeft plaatsgevonden is het van belang vast te stellen hoe vaak men opnieuw in de fout is gegaan. De recidivefrequentie geeft het aantal keren aan dat een persoon – volgens de informatiehouder en gedurende de observatieperiode – heeft gerecidiveerd. Het is een numerieke variabele, per persoon kan het aantal recidivegevallen eenvoudigweg worden geteld. Op groepsniveau wordt dan gebruikgemaakt van het gemiddelde of elke andere maat voor centrale tendentie (mediaan, modus enzovoort). Om de frequentie vast te stellen moet duidelijk zijn wat de *tel-eenheid* is, dit is het element waarvan het aantal wordt bepaald. In principe is dat uiteraard de strafbare gedraging (het delict: misdrijf of overtreding), maar het is niet altijd duidelijk hoe vaak een bepaalde gedraging heeft plaatsgevonden. Een proces-verbaal dat wordt opgesteld door de politie kan een ongespecificeerd aantal delicten bevatten. De dagvaarding van het OM ook. Deze beschrijft soms meerdere delicten en bevat vaak slechts een selectie van de delicten die in het proces-verbaal zijn terug te vinden. Indien een onderzoeker bij zijn recidivemeting gebruikmaakt van politie- of justitiegegevens, dan moet hij dus kiezen voor een grotere tel-eenheid dan het afzonderlijke strafbare feit. Het vaststellen van de frequentie komt dan neer op het tellen van het aantal nieuwe processen-verbaal of van het aantal nieuwe strafzaken.

3 *Snelheid*

Recidivesnelheid is het tempo waarin men recidiveert. Snelheid is een afgeleide van frequentie. Het is de verhouding tussen de recidivefrequentie en de tijd die is verstreken tot het laatste recidivegeval. Betreft men alleen het eerste geval van recidive in de analyse, dan krijgt het begrip echter een andere betekenis. In dat geval geeft de recidivesnelheid aan hoe lang het heeft geduurd voordat er werd gerecidiveerd. Deze maat kan van belang zijn, bijvoorbeeld om vast te stellen of een bepaalde interventie een recidive*vertragend* effect heeft gehad. Hoe langer terugvalt uitblijft, hoe beter het is. In deze betekenis is snelheid niets anders dan 'de tijd tot de eerste recidive'. Ze wordt uitgedrukt in dagen, maanden of jaren. Per groep kan de gemiddelde snelheid worden berekend, maar alleen voor de respondenten die hebben gerecidiveerd. De non-recidivisten tellen bij het

vaststellen van de gemiddelde snelheid niet mee.¹³ Voorwaarde voor het gebruik van deze maat is dat men met een zekere precisie het tijdstip van de terugval weet te bepalen. Dit kan aan de hand van een pleegdatum of -periode, of op basis van andere datums die in de gebruikte bron zijn vastgelegd.

4 *Ernst*

Naast het beantwoorden van de vragen of, hoe snel en hoe vaak de respondenten hebben gerecidiveerd, moet het onderzoek zicht bieden op de ernst van de nieuwe delicten. Om de ernst te kunnen bepalen moeten de delicten naar inhoud worden gerangschikt. Men kan daarvoor gebruikmaken van de standaardclassificatie van het CBS (Eggen & Van der Heide, 2004). Het CBS gebruikt zeven hoofdcategorieën om delicten in te delen: delicten beschreven in het Wetboek van Strafrecht (gewelddsmisdrijven, vermogensmisdrijven, vernielingen en openbare orde en overige misdrijven), de Wegenverkeerswet, de Wet op de economische delicten, de Opiumwet, de Wet wapens en munitie, het Wetboek van Militair Strafrecht en overige wetten. Maar men kan natuurlijk ook een fijnere indeling toepassen. Een en ander hangt af van de dadergroep die men volgt en de concrete probleemstelling van het onderzoek.

Als de recidivedelicten naar aard zijn ingedeeld, moeten ze naar zwaarte worden geordend. Er zijn minstens drie manieren om dit te doen. Op de eerste plaats kan men nagaan welke wetsartikelen zijn overtreden, om ze vervolgens te ordenen aan de hand van de strafdreiging. In de wet- en regelgeving staat beschreven welke (gevangenis)straffen bij overtreding van een bepaald artikel maximaal kunnen worden opgelegd. Met deze strafmaat drukt de wetgever uit hoe ernstig hij de delicten vindt. Hoe hoger de strafdreiging, hoe ernstiger het gepleegde delict. Deze methode om de ernst van de recidivedelicten te bepalen is aantrekkelijk om haar eenvoud, maar erg nauwkeurig is ze niet. Er wordt geen rekening gehouden met belangrijke details, zoals de feitelijke gedragingen van de dader en de omstandigheden waarin de delicten zijn gepleegd. Diefstal van een laag geldbedrag lijkt minder erg dan diefstal van een groot bedrag. Toch kan voor beide delicten maximaal vier jaar gevangenisstraf worden opgelegd en vallen zij, indien de maximale strafdreiging de maatstaf is, in dezelfde ernstcategorie.

Om rekening te houden met verschillen in de ernst van de delicten zou men als alternatief kunnen afgaan op de rechtspraktijk. De inschatting van de ernst wordt dan gebaseerd op de feitelijke afdoening. In tegenstelling tot de theoretische strafdreiging wordt in de sanctie die feitelijk wordt opgelegd wel rekening gehouden met de concrete omstandigheden waarin

13 Wel kan van de gehele groep de mediane snelheid worden berekend. Dit leidt tot uitspraken als: 'na 3,5 jaar heeft de helft van de groep gerecidiveerd'. Als niet iedereen even lang wordt gevolgd, is de mediaan een betere centrale maat dan de gemiddelde snelheid. Ze is minder gevoelig voor verschillen in observatietermijn.

het delict plaatsvond. Een zuivere maat is het echter nog niet, want in de afdoening klinkt wel meer door, zoals het strafrechtelijk verleden en de proceshouding van de dader en al dan niet structurele verschillen in straf-toemettingsbeleid. Bovendien is niet altijd duidelijk of een bepaalde afdoening zwaarder is dan een andere. Is een boete bijvoorbeeld lichter dan een voorwaardelijke vrijheidsstraf? Is twee weken gevangenisstraf voor een veelpleger zwaarder dan een werkstraf voor een first offender? Een derde, meer uitgebreide methode om de ernst van recidivedelicten te bepalen is ze te beschrijven in termen van een aantal onderliggende componenten, zoals de strafdreiging van het delict, de schade die en het letsel dat is toegebracht en andere verzwarende of juist verzachtende omstandigheden. De componenten kunnen worden gebruikt om per delict een samengestelde 'ernstscore' te berekenen. Hoe hoger de score, des te zwaarder het delict. Een dergelijke methode is in enkele studies toegepast, zij het in een ander verband. Om de 'zwaarte' van groepen justitiabelen in te schatten zijn op basis van de delictgegevens uit hun strafdossiers ernstscores berekend (Cozijn et al., 2000). Deze werkwijze zou ook bij recidiveonderzoek kunnen worden gehanteerd. Probleem is echter dat niet van alle delicten dossiers beschikbaar zijn en dat toepassing van de methode veel tijd kost. Bovendien is ze niet gevalideerd. Het is niet bekend hoe goed de scores de ernst van de delicten weergeven.

Prevalentie, frequentie, snelheid en ernst zijn de vier facetten van het recidivebegrip. Om deze te kunnen meten moet een aantal keuzes worden gemaakt: de termijn, het domein, het bereik en de tel-eenheid van het onderzoek moeten worden bepaald. De termijn, het domein, het bereik en de tel-eenheid zijn de 'bouwstenen' van het recidiveonderzoek. Als men dezelfde bouwstenen gebruikt, zijn de uitkomsten van het onderzoek onderling vergelijkbaar. Lopen de keuzes met betrekking tot de operationalisering van het recidivebegrip uiteen, dan kunnen de uitkomsten van de studies niet goed tegen elkaar worden afgezet en ontstaat er een scheef beeld. Het hanteren van een observatieperiode van zes jaar leidt tot hogere recidivepercentages en -frequenties dan een follow-up van zes maanden. Een zelfrapportagestudie leidt tot andere prevalentiecijfers dan de analyse van strafdossiers. Als men alleen seksuele delicten tot het bereik van het onderzoek rekent, dan valt de recidive lager uit. Telt men losse delicten, dan is de recidivefrequentie hoger dan wanneer men rechtszaken telt. Met het oog op de gebruikswaarde van het onderzoek is er veel voor te zeggen recidivemetingen te standaardiseren. Maar dit betekent dat er ten aanzien van de termijn, het domein, het bereik en de tel-eenheid vaste keuzes moeten worden gemaakt. Dit zal gebeuren in paragraaf 3.3. Eerst gaan we nader in op de mogelijke bronnen voor recidiveonderzoek.

3.2 Bronnen van recidivegegevens: een vergelijkende analyse

Er zijn verschillende methoden om recidivegegevens te verzamelen. Men kan daders in interviews op de man af vragen of zij nieuwe delicten hebben gepleegd. Men kan een schriftelijke daderenquête houden, of men kan informanten inschakelen. In al deze gevallen is de bron een persoon, de gegevens worden speciaal voor het onderzoek gegenereerd. Het is ook mogelijk om van bestaande gegevens uit te gaan, bijvoorbeeld door de dossiers te analyseren van politie, justitie of elke andere organisatie die kennis heeft van het delictgedrag van de respondenten. Of men kan gebruikmaken van de informatie die standaard door de instanties in hun administratiesystemen wordt vastgelegd. Elk type bron brengt, als het gaat om verzamelen van de gegevens, eigen logistieke problemen met zich mee. Voor vier bronnen zijn de mogelijkheden en beperkingen met betrekking tot het meten van de diverse facetten van het recidivebegrip geïnventariseerd. Er volgt een korte karakteristiek.¹⁴

1 *Zelfrapportage door daders*

Bij zelfrapportage vormt de dader de bron. Elke methode waarbij potentiële recidivisten gevraagd wordt informatie te geven over hun eigen delictgedrag, kan worden aangemerkt als zelfrapportage. Binnen de criminologie wordt zelfrapportage gebruikt om zicht te krijgen op het *dark number*, de delicten die niet ter kennis komen van de politie.¹⁵ In Nederland deed zich jaren terug een discussie voor over de waarde van de methode. Sommigen vinden het een bruikbaar instrument (Van der Heijden, Sijtsma & 't Hart, 1995; Nijboer, 1995; Swanborn, 1996). Anderen twijfelen aan de betrouwbaarheid of validiteit ervan (Bruinsma, 1994; Hessing & Elffers, 1995).¹⁶ Het grootste probleem bij het toepassen van zelfrapportage is het gevaar voor over- en onderrapportage: respondenten melden delicten die ze niet hebben gepleegd, of verzwijgen delicten die ze wel hebben gepleegd. In hun overzichtsartikel voeren Junger-Tas en De Haen-Marshall (1999) aan dat de toepassing van de methode sinds het midden van de vorige eeuw sterk is verbeterd. Ondanks de tekortkomingen achten zij zelfrapportage een bruikbare methode om criminaliteit te meten en de achtergronden daarvan te leren kennen.

14 Voor een uitgebreide, meer algemene verhandeling over bronnen op het gebied van de kwantitatieve criminografie, zie Van Kerckvoorde (1995). Van Kerckvoorde bespreekt de beperkingen van officiële statistieken, zelfrapportagemethoden en slachtofferenquêtes. Zijn conclusie is dat geen van deze bronnen een betrouwbaar beeld geeft van de omvang en verdeling van de criminaliteit.

15 Andere methoden om niet-ontdekte of niet-aangegeven delicten boven water te krijgen zijn slachtofferenquêtes (o.m. Eggen & Van der Heide, 2005), schattingen op basis van vangst-hervangst quota, sneeuwbalsteekproeven en 'detection controlled estimation' (Sikkel, Van der Heijden & Van Gils, 2006)

16 Betrouwbaarheid impliceert dat het gebruik van een instrument bij herhaalde meting van een onveranderd object tot hetzelfde resultaat leidt. Met validiteit wordt over het algemeen bedoeld dat een instrument meet wat het behoort te meten (zie o.m. Drenth & Sijtsma, 1990; Van den Brink & Mellenbergh, 2003).

2 HKS

De politieregio's maken gebruik van het zogeheten herkenningssysteem (HKS) om aangiftes van misdrijven en gegevens van verdachten op te slaan (Eggen & Van der Heide, 2005). Wat het HKS in aanleg geschikt maakt voor recidiveonderzoek is dat het voor elke verdachte een overzicht geeft van zijn antecedenten, zijn eerdere politiecontacten. HKS geeft in feite per persoon de criminele carrière weer. De overkoepelende, landelijke versie van het HKS wordt beheerd door de Dienst Nationale Recherche Informatie (DNRI), een onderdeel van het Korps Landelijke Politiediensten (KLPD), en is beperkt beschikbaar voor wetenschappelijk onderzoek.¹⁷ Op nationaal niveau dienen de gegevens van verdachten die in meer dan één regio actief zijn, te worden ontdekt. Daarnaast moet men rekening houden met de effecten van verjaring. Als er geen nieuwe registraties zijn, worden de oude zaken van een persoon na verloop van tijd verwijderd en is het beeld van zijn criminele carrière dus niet langer compleet.

3 Strafdossiers

Wanneer het OM vervolging van een verdachte overweegt, wordt er een strafdossier opgemaakt. Naast personalia en achtergronden van de verdachte bevat dit dossier de processen-verbaal van de politie. Hierin wordt de toedracht van de delicten beschreven, zowel van de kant van de mogelijke dader als vanuit het perspectief van het slachtoffer en eventuele getuigen. Indien de strafzaak wordt voorgelegd aan de rechter, bevat het dossier een dagvaarding waaruit blijkt waarvan de verdachte formeel wordt beschuldigd. Ook zijn afdoeningsgegevens opgenomen; in de vorm van een vonnis als het tot een rechtszaak kwam, of anderszins, als het OM de zaak zelf heeft afgedaan met een sepot of een transactie. Een selectie van de gegevens uit de strafdossiers wordt opgenomen in een gedigitaliseerd bestand, COMPAS, het administratiesysteem van de arrondissementsparketten. Een nadeel van strafdossiers als bron voor recidiveonderzoek is dat voor elk geval van recidive een apart dossier moet worden opgespoord en geanalyseerd. Strafdossiers worden niet op naam, maar op parketnummer bewaard.

4 JDS

Evenals het HKS biedt het Justitiële Documentatiesysteem (JDS) een overzicht van de criminele voorgeschiedenis van een verdachte. Het gaat nu echter niet om politie- maar om *justitie*contacten. Per persoon wordt in het JDS bijgehouden van welke delicten men in de loop der jaren werd ver-

17 Sinds enkele jaren werken de korpsen ook met een landelijke afslag van een verzameling van decentrale informatiesystemen, GIDS. Dit staat voor: geïntegreerde interactieve databank voor strategische bedrijfsinformatie. Deze vormt de basis voor gestandaardiseerde beleidsrapportages. GIDS bevat historische gegevens, een deel daarvan op persoonsniveau, en is in principe dus geschikt voor recidiveonderzoek.

dacht, welke instanties de strafzaken hebben afgehandeld en met welke sancties de zaken werden afgedaan. Vroeger werden de justitiële gegevens van de verdachten per arrondissement op kaart bijgehouden, maar in de jaren '90 van de vorige eeuw werd de registratie gecentraliseerd en geautomatiseerd. Nieuwe zaken worden tegenwoordig direct vanuit COMPAS aan het JDS toegevoegd. Het JDS wordt beheerd door JustId, de Justitiële Informatiedienst, gevestigd in Almelo. Ook het JDS kent verjaringsregels. In de Wet justitiële en strafvorderlijke gegevens is vastgelegd welke informatie wordt bijgehouden, hoe lang deze wordt bewaard en wie de gegevens mag inzien. De registratie is primair bedoeld voor politie- en justitieambtenaren aan elk strafdossier wordt een uittreksel uit het JDS toegevoegd. Maar ook onderzoekers kunnen onder bepaalde voorwaarden inzage krijgen.

Het HKS, de strafdossiers en het JDS zijn bestaande registratiesystemen. Ze worden gebruikt in verschillende fasen van de justitiële keten. Het HKS bij de opsporing, de strafdossiers bij de vervolging, en het JDS als de strafzaken zijn afgedaan. Een deel van de delicten waarover via zelfrapportage gegevens worden verzameld komt niet in de justitiële keten terecht. Daarin wijkt deze bron van de andere af. Tabel 3.1 geeft een globale inschatting van de mogelijkheden en de beperkingen van de vier bronnen. In de eerste kolom vinden we de basiselementen van het recidiveonderzoek terug. In de eerste rij wordt het *domein* van de bronnen aangegeven. Zelfrapportage bestrijkt het grootste universum van mogelijke delicten. Het is denkbaar dat via deze methode een beeld wordt verkregen van alle delicten die door de respondenten zijn gepleegd. Bij gebruik van de andere bronnen is dat de facto niet het geval.¹⁸ Het HKS biedt alleen zicht op de delicten waarvan een aangifte bestaat en waarbij de politie een verdachte heeft weten op te sporen. Het domein van de strafdossiers en het JDS is weer een deelverzameling van de delicten die zijn vastgelegd in het HKS. Strafdossiers en het JDS bevatten alleen gegevens over de delicten die door de politie zijn opgehelderd en die voor eventuele vervolging zijn doorverwezen naar het OM.

De tweede rij van de tabel bevat de *tel-eenheden* van de vier bronnen. Bij zelfrapportage is het meest elementaire onderdeel het delict, bijvoorbeeld een concreet geval van zwartrijden of een inbraak. In het HKS is de kleinste eenheid die wordt geregistreerd het antecedent. Dit is in de meeste gevallen één delict, maar niet altijd. Het kan ook gaan om meer strafbare feiten in één, bijvoorbeeld bij een serie van inbraken. Bij het gebruik van strafdossiers kan men verschillende tel-eenheden kiezen, maar niet het afzonderlijke delict. De dossiers omvatten de processen-verbaal van de

18 Alleen als iedereen 24 uur per dag in de gaten zou worden gehouden of iedere dader al zijn misstappen uit eigen beweging zou aangeven, zouden de autoriteiten van elk gepleegd delict op de hoogte zijn. De vraag is in hoeverre er onder zulke onwaarschijnlijke condities nog sprake zou zijn van crimineel gedrag.

politie waarin de actuele antecedenten worden beschreven. Het OM maakt van één of meer processen-verbaal een strafzaak en geeft deze een parketnummer. Voor het OM vormt de strafzaak de eenheid voor verdere vervolging. De officier beslist of hij de strafzaak zelf afdoet of voor de rechter brengt. Het komt voor dat er op enig moment tegen een persoon meer strafzaken tegelijk lopen. In dat geval kunnen de zaken worden gevoegd tot één rechtszaak waarin de rechter uitspraak doet over alle aangevoerde feiten. In het JDS zijn geen processen-verbaal opgenomen. Bij deze bron kan men als tel-eenheid dus alleen kiezen tussen de strafzaak en de rechtszaak.

Hoe groter het domein dat de bron bestrijkt en hoe kleiner de tel-eenheid is, des te uitgebreider is het zicht dat de gegevensbron biedt op de recidive van de betrokkenen. Op grond van deze twee criteria is zelfrapportage de beste bron. De methode van zelfrapportage levert meer waarnemingen op dan het sterk gefilterde HKS, de strafdossiers of het JDS. Maar een uitgebreider zicht levert niet altijd een beter beeld op, dat hangt ook af van de kwaliteit van de waarnemingen.¹⁹ De onderste helft van tabel 3.1 geeft daarom aan hoe goed de overige aspecten van het recidivebegrip gemeten kunnen worden, gegeven het domein en de tel-eenheid van de betreffende bron. Positieve scores wijzen op veel precisie en flexibiliteit. Lage scores betekenen dat de mogelijkheden om op de onderscheiden onderdelen valide en betrouwbare recidivemetingen uit te voeren, beperkt zijn. Omdat het gaat om een globale inschatting, zijn er slechts drie grove kwalificaties mogelijk: 'laag', 'middel' en 'hoog'.

Tabel 3.1 Domein, tel-eenheid en de mogelijkheden van vier bronnen om in verschillende opzichten betrouwbare en valide recidivemetingen te doen

	Zelfrapportage	HKS	Strafdossier	JDS
Domein	alle gepleegde delicten	delicten die door de politie zijn opgehelderd	delicten waarbij het OM vervolging overweegt	delicten waarbij het OM vervolging overweegt
Tel-eenheid	delict	antecedent	antecedent, strafzaak of rechtszaak	strafzaak of rechtszaak
Bereik	laag	middel	hoog	hoog
Termijn	laag	hoog	hoog	middel
Prevalentie	middel	middel	hoog	hoog
Frequentie	middel	middel	hoog	hoog
Snelheid	middel	hoog	hoog	middel
Ernst	laag	laag	hoog	middel

19 Zoals bleek in de parabel van de gebroeders Eénoog en Oogklep, die elkaar betwisten om de heerschappij over een verder verlaten eiland Epistemica. Beiden vonden dat hen die eer toekwam; Eénoog omdat hij alles kon zien en Oogklep omdat hij alles wát hij kon zien, goed zag.

Afgaand op de waarden in de tabel vormt niet zelfrapportage maar het strafdossier de beste bron van recidivegegevens. Op elk van de onderscheiden aspecten wordt een hoge score behaald. Strafdossiers bevatten zeer veel, goed bruikbare gegevens. De informatie over het delictgedrag wordt door professionals vastgelegd. Zowel in het proces-verbaal, op de dagvaarding als in het vonnis komen delictomschrijvingen voor. Een onderzoeker kan de operationele definitie van recidive (het 'bereik') dus naar believen aanpassen. De vele datums die in de dossiers te vinden zijn, maken het gemakkelijk om vast te stellen wanneer de recidive precies heeft plaatsgevonden (hoge score op termijn en snelheid), en voor het bepalen van de recidivefrequentie kan men kiezen uit verschillende tellenheden. De ernst van de recidive ten slotte, kan op elk van de eerder besproken wijzen worden vastgesteld: aan de hand van de wettelijke strafdreiging, aan de hand van de feitelijke afdoening of aan de hand van een combinatie van factoren. Een dossier bevat bijvoorbeeld ook gegevens over de schade die met het delict is veroorzaakt of over het letsel dat daarbij is ontstaan.

Binnen het domein dat met deze bron wordt bestreken, bieden strafdossiers dus de beste mogelijkheden om betrouwbare en valide metingen te doen. Alle delicten uit het domein vindt men in de strafdossiers terug. Bij zelfrapportage is dit niet het geval. Uit onderzoek van Huizinga en Elliot (1986) naar de correspondentie tussen zelfgerapporteerde en officieel geregistreerde arrestatiegegevens bleek dat op zijn minst een derde van de ondervraagden een deel van de feitelijk plaatsgevonden arrestaties verborg, of vergat te rapporteren. Een vijfde van de respondenten verzweeg zelfs alle arrestaties. Eerder ontdekten Elliot en Voss (1974) al dat de kans op onderrapportage groter is naarmate de gepleegde delicten ernstiger zijn. Hindelang, Hirschi en Weis (1979) stelden vast dat Afro-Amerikaanse respondenten minder snel delicten toegaven dan blanke respondenten. Aanwijzingen als deze hebben geleid tot de veel gedeelde opvatting dat zelfrapportage vooral geschikt zou zijn om lichtere vormen van criminaliteit in kaart te brengen (Kleck, 1982; Angenent, 1984; Van Kerckvoorde, 1995; Farrington et al., 1996; Messner & Rosenfeld, 1997; Junger-Tas & De Haen-Marshall, 1999; De Haan et al., 1999). Bij zwaardere of bijzondere dadergroepen leidt de methode van zelfrapportage tot vertekende resultaten.

Omdat de validiteit en betrouwbaarheid van de langs deze weg verzamelde gegevens op zijn best een wisselend beeld laten zien, vormen zelfrapportages geen goede basis voor doorlopend, algemeen recidiveonderzoek. Strafdossiers zijn hiervoor echter evenmin geschikt. Het is te bewerkelijk om voor elk recidivedelict een complex strafdossier te analyseren, zeker als de onderzoeksgroep omvangrijk is. Uit praktische overwegingen zal men daarom vaak voor bronnen als het HKS of het JDS kiezen. Dit zijn geautomatiseerde informatiesystemen waarin de onderzoeksgegevens per persoon al in de juiste volgorde zijn gezet. Gelet op de scores in de tabel is

de kwaliteit van beide bronnen redelijk tot goed te noemen. In het voordeel van het HKS spreekt dat het een groter domein van delicten beslaat en het een kleinere tel-eenheid kent. Een belangrijk nadeel is dat het alleen voorlopige delictgegevens bevat en geen informatie over de afdoening heeft. Op het moment dat het HKS wordt gevuld, staat nog niet vast dat de verdachte het delict heeft gepleegd. De politie heeft de zaak nog in onderzoek. Hangende het onderzoek wordt aan het delictgedrag veelal de zwaarst denkbare kwalificatie toegekend (De Haan et al., 1999: 20). Dit wordt dan later door de officier of de rechter bijgesteld. Het wetsartikel op het proces-verbaal blijkt niet zelden af te wijken van het wetsartikel dat uiteindelijk op het vonnis staat vermeld (Projectteam SRM, 1997). Op basis van HKS-gegevens is het met andere woorden niet geheel zeker dat de verdachte heeft gerecidiveerd en zo ja, met welk delict. De prevalentie, ernst en frequentie van de recidive worden daardoor minder nauwkeurig gemeten.

Het JDS kent deze problemen niet, of althans veel minder. De delicten die in het JDS worden vastgelegd, zijn juridisch getoetst. In elk geval door de officier van justitie, maar vaak ook door de rechter en bij hoger beroep of cassatie zelfs door meer instanties. Op grond van de in het JDS vastgelegde gegevens staat vast dat een persoon heeft gerecidiveerd. Duidelijk is ook waaruit de recidive bestond. Omdat het JDS afdoeningen bevat, kan de ernst van de recidive bovendien op meer manieren worden ingeschat dan op basis van het HKS. Een nadeel is dat het enige tijd kan duren voordat een nieuw delict in het systeem is opgenomen. Voor recidiveonderzoek met een zeer korte observatietermijn is deze bron dus minder geschikt. Op alle andere punten moet het JDS echter hoger worden aangeslagen dan het HKS. *Binnen* zijn domein kan de recidive uiterst efficiënt worden gemeten. De delicten in het JDS zijn gewaarmerkt, de validiteit van het JDS is gegarandeerd. Ze wordt alleen bedreigd door administratiefouten en gerechtelijke dwalingen.²⁰

Zelfrapportage is van grote waarde voor de criminologie. Het is een methode om delicten die normaal gesproken verborgen blijven, boven water te krijgen. Indien men van grote, uiteenlopende groepen daders de recidive wil bepalen, kan men echter beter kiezen voor het JDS. Het probleem dat deze bron betrekking heeft op slechts een deel van de criminaliteit, kan worden omzeild door de definitie van 'recidive' aan te passen. Een meting op basis van het JDS brengt niet het recidivegedrag, maar de *strafrechtelijke* recidive in kaart. Dit is het deel van de feitelijke recidive dat aanleiding gaf tot hernieuwd ingrijpen door politie en justitie. Hoe groot het aandeel is en hoe het zich verhoudt met het dark number zal nooit helemaal duidelijk worden. Crimineel gedrag is, of men het nu leuk vindt

20 Een klein percentage personen komt onder verschillende namen in het JDS voor. Zodra bekend, worden deze 'ontdubbeld'. In 2006 werd de Commissie Evaluatie Afgesloten Strafzaken (CEAS) ingesteld. Indien er bij zware strafzaken twijfel blijft bestaan over de schuldigverklaring, onderzoekt deze commissie of de bewijsvoering 'ernstige manco's' vertoont (De Ridder, Klein Haarhuis & De Jongste, 2008).

of niet, gehuld in een hardnekkige, epistemologische nevel. De mist trekt hoogstens plaatselijk even op, wanneer een onderzoeker via soms ingenuïe vraagtechnieken aan een specifieke groep mogelijke recidivisten tracht te ontfutselen wat men eigenlijk niet kwijt wil. De essentie van het probleem is dat men ook via zelfrapportage nooit zal weten in hoeverre men daarin is geslaagd. Wil men recidiveonderzoek doen op grote schaal en bij voortduring, dan zal men moeten beschikken over een constante aanvoer van betrouwbare en valide gegevens. Het JDS meet inderdaad slechts een deel van de criminaliteit, maar dat deel wordt wel volledig en correct gemeten. Behoudens mogelijke registratie-effecten – waarover in paragraaf 3.4.3 meer – meet het JDS de strafrechtelijke recidive op een bestendige manier.

3.3 De WODC-Recidivemonitor

Het idee om op grote schaal en met vaste regelmaat recidivemetingen uit te voeren, is ontstaan in het begin van de jaren '90 van de vorige eeuw. Onderzoekers die werkten aan de evaluatie van strafrechtelijke interventies hadden behoefte aan vergelijkingsmateriaal. Het recidiveonderzoek dat zij verrichtten stond veelal op zichzelf, de verzamelde cijfers misten een ijkpunt. Door de ontwikkeling van een vaste standaard zou de gebruikswaarde van het onderzoek kunnen worden verhoogd. Het duurde enkele jaren voordat aan de totstandkoming van de Recidivemonitor kon worden gewerkt. De belangstelling voor uitstroomcijfers van justitiële inrichtingen was inmiddels toegenomen. Het traject 'van beleidsbegroting naar beleidsverantwoording' speelde daarbij een rol (Ministerie van Financiën, 2000). Veranderingen in het systeem van financiering van overheidsdiensten noopte het ministerie van Justitie op zoek te gaan naar geldige prestatie-indicatoren. Op het terrein van de 'effectieve tenuitvoerlegging' ging de aandacht al snel uit naar recidive. De beleidsdirecties vroegen het WODC een monitor te ontwikkelen. De eerste stap bestond uit het bundelen van het recidiveonderzoek dat eerder op dit instituut was verricht. Het merendeel van de studies was gebaseerd op gegevens uit het JDS. Uit de aanpak van de verschillende studies werd een voorlopige, gemeenschappelijke werkwijze gededuceerd.

De tweede stap in de ontwikkeling van de Recidivemonitor was de automatisering van de dataverzameling. Het JDS bevat de justitiële documentatie van miljoenen (rechts)personen, maar is op zichzelf niet geschikt als onderzoeksbestand. Het systeem is bedoeld om politie- en justitiemedewerkers te ondersteunen bij de opsporing en vervolging van individuele personen. Het is ontworpen om per individu een uittreksel te geven van zijn justitiële gegevens en niet in staat om tabellen te genereren waarin voor groepen personen gegevens op een rij worden gezet. Om onderzoek te doen moeten de gegevens uit het JDS eerst worden ingevoerd in een

bestand; een tijdrovend proces en een van de redenen waarom het recidiveonderzoek in Nederland zo lang kleinschalig is gebleven. Voordat de standaard van de Recidivemonitor breed kon worden toegepast, moest dus eerst een aantal processen worden geautomatiseerd.

3.3.1 *De OBJD*

In 1999 kreeg een projectgroep van het ministerie van Justitie opdracht de informatie uit het JDS toegankelijk te maken voor beleidsanalyses en statistisch onderzoek (Directie Informatisering, 1999). Dit leidde uiteindelijk tot de komst van de onderzoek- en beleidsdatabase justitiële documentatie (OBJD). De OBJD is speciaal met het oog op het onderzoek van de Recidivemonitor ontwikkeld. Zij is een kopie van het JDS, waarbij de onderliggende structuur van de database zo is aangepast dat de gegevens wel groepsgewijs kunnen worden weggeschreven.

De OBJD is geanonimiseerd. Ze bevat dezelfde gegevens als het JDS, althans voor zover die voor wetenschappelijk onderzoek van belang kunnen zijn, maar heeft geen namen of andere naar individuele personen te herleiden gegevens. De paraktnummers en identificatienummers waaronder de personen in het JDS voorkomen, zijn versleuteld. De OBJD voldoet daarmee aan de privacywetgeving. Het systeem is alleen geschikt voor wetenschappelijk onderzoek en niet bedoeld om individuele personen in op te sporen. Die functie is voorbehouden aan het JDS en de Justitiële Informatiedienst.

De OBJD wordt om de drie maanden ververst. Elk kwartaal worden de zaken die nieuw in het JDS zijn opgenomen aan de OBJD toegevoegd. Een belangrijk verschil met het JDS is dat de gegevens in de OBJD niet worden verwijderd als zij zijn verjaard. De Wet justitiële en strafvorderlijke gegevens schrijft voor dat indien een persoon geen nieuwe justitiecontacten opbouwt, zijn oude zaken na verloop van tijd uit het JDS worden verwijderd. In de OBJD blijven ze beschikbaar. Ook de zaken van de personen die zijn overleden of ouder zijn dan tachtig jaar, blijven behouden. Op basis van de OBJD kan van alle personen die in Nederland verdachte waren in een strafzaak, de volledige strafrechtelijke carrière in kaart worden gebracht. De enige beperking is dat de persoon na 1996 (nog) een justitiecontact moet hebben gehad.²¹

Er zijn twee manieren om de justitiële gegevens van een onderzoeksgroep uit de OBJD te halen: direct of door tussenkomst van de Justitiële Informatiedienst. Indien de onderzoeksgroep kan worden beschreven in termen van de velden die in de OBJD zijn opgenomen, kunnen de gegevens van

21 De oude zaken van een persoon worden in het JDS ingevoerd als er zich een nieuwe melding voordoet. Indien men onderzoek doet naar de strafrechtelijke carrière van personen die al enige tijd geen strafzaak meer hebben gehad, dan moeten hun zaken eerst vanaf de kaarten uit het handmatige systeem aan het JDS worden toegevoegd. Bij de volgende driemaandelijkse verversing zijn deze personen ook in de OBJD te vinden.

de respondenten rechtstreeks uit de database worden geëxtraheerd. Een voorbeeld van dergelijk onderzoek is een meting onder volwassenen die in 1997 een onvoorwaardelijke vrijheidsstraf kregen opgelegd. Aan de hand van de geboortedatum, het jaar van beslissing en informatie over de afdoening kunnen de personen in het systeem worden opgespoord. Met de versleutelde persoonsnummers kunnen hun strafzaken uit de OBJD worden geselecteerd.

Gaat het echter om een meer specifieke dadergroep waarvan niet alle kenmerken in de OBJD zijn opgenomen, dan is er een omweg nodig via de Justitiële Informatiedienst. Een groep daders die hun detentie in een bepaalde gevangenis hebben uitgezeten, kan bijvoorbeeld niet direct worden opgespoord, doordat de OBJD niet bijhoudt in welke inrichtingen vrijheidsstraffen worden geëxecuteerd. In zo'n geval dient men te beschikken over de personalia van de mensen in het onderzoek (i.c. namen en parketnummers). De Justitiële Informatiedienst zoekt de nummers van de respondenten in het JDS op. Deze nummers worden versleuteld en teruggestuurd naar het WODC, waar de justitiële gegevens uit de OBJD worden geselecteerd. Omdat met deze procedure de anonimiteit van de respondenten wordt opgeheven, is voor dit type onderzoek toestemming nodig van de Minister van Justitie.

Box 3.1 Corporate crime en organisatiecriminaliteit

Het JDS bevat niet alleen informatie over natuurlijke personen, maar houdt ook de strafzaken van bedrijven, organisaties en instellingen bij. Naar de criminaliteit van 'corporate actors' (Coleman, 1974) is niet veel onderzoek gedaan. Meest recent publiceerden Loschnig-Gspandl (2000), Simpson et al. (2002) en Nelken (2002) over dit onderwerp. In Nederland verscheen een studie van Van den Berg, Aidala en Beenackers (2002) naar 'organisatiecriminaliteit', een term geïntroduceerd door Van de Bunt (1992). Organisatiecriminaliteit is een breder begrip dan 'corporate crime'. Ook natuurlijke personen kunnen bij de uitoefening van hun functie organisatiecriminaliteit plegen. Bij 'corporate crime' worden rechtspersonen aansprakelijk gesteld. Hoewel deze gemakkelijk van gedaante kunnen wisselen en het dus lastiger zou zijn om herhaalde overtredingen te meten, blijkt dat organisaties, bedrijven en instellingen betrekkelijk vaak opnieuw de fout in gaan.

Tabel 3.2 Totaal aantal strafzaken in de carrière van natuurlijke personen en rechtspersonen vervolgd in 1997 (in % van het kolomtotaal)

	Soort persoon		Totaal
	Natuurlijk	Rechtspersoon	
1	29,59	37,17	29,97
2	14,07	15,13	14,12
3 tot 4	16,42	14,60	16,32
5 tot 10	20,08	15,20	19,84
11 of meer	19,84	17,91	19,74
Totaal aantal personen	293.586	15.654	309.240
Totaal aantal zaken	2.462.874	171.419	2.634.293
Gemiddeld	8,39	10,95	8,52
Mediaan	3	2	3
Modus	1	1	1
Maximum	936	6.667	6.667

Bron: OBJD

Tabel 3.2 laat zien dat in 1997 15.654 rechtspersonen verdachte waren in een strafzaak. Veel minder dan het aantal natuurlijke personen. Duidelijk is echter ook dat bedrijven wel degelijk recidiveren. Voor meer dan 60% van de rechtspersonen was de strafzaak in 1997 niet het enige justitiecontact. Aan de strafzaak in 1997 gingen één of meer strafzaken vooraf, of er volgden nieuwe zaken in de jaren daarop.* Bij natuurlijke personen heeft 70% meer dan één justitiecontact. Gemiddeld ligt de totale frequentie voor rechtspersonen zelfs hoger dan voor natuurlijke personen. Voor een deel ligt dit aan een beperkt aantal 'uitbijters' in de staart van de verdeling. Er zijn elf bedrijven die tot juli 2005 al meer dan duizend keer werden vervolgd voor een overtreding of een misdrijf. Als deze elf bedrijven buiten beschouwing worden gelaten, daalt het gemiddelde aantal justitiecontacten voor rechtspersonen naar 9,4. Vermoedelijk gaat het hier om grote autoverhuurbedrijven die de verkeersovertredingen van hun klanten op hun naam krijgen. Maar zeker is dit niet, want de identiteit van de rechtspersonen in de OBJD is niet bekend.

* De observatieperiode eindigde op 1 juli 2005.

Bewerking van de ruwe gegevens

De informatie die wordt opgeslagen in de OBJD is niet direct klaar voor analyse, de ruwe gegevens dienen eerst te worden bewerkt. Het ontwerpen van vaste procedures voor het verwerken van de OBJD-gegevens vormde de derde stap in de ontwikkeling van de Recidivemonitor. De meeste velden in de OBJD bevatten losse tekststrings, wisselende stukken tekst die

eerst moeten worden gecodeerd en ingedeeld voordat er verder mee kan worden gewerkt. De delictinformatie in de OBJD is vastgelegd in termen van overtreden wetsartikelen. Aan deze artikelen worden de wettelijke strafmaxima gekoppeld, zodat per zaak de strafbare feiten naar ernst kunnen worden gerangschikt en het zwaarst bedreigde feit kan worden geselecteerd.²² Daarnaast wordt aan elk feit of elke artikelcombinatie de cijfercode toegekend die het Centraal Bureau voor de Statistiek gebruikt om delicten te classificeren. Op basis van deze code kunnen de strafbare feiten op verschillende manieren naar soort worden ingedeeld.

Ook de afdoeningsinformatie wordt samengevat. Dit gebeurt op twee manieren. Per zaak wordt nagegaan of er sprake is van een bepaald type afdoening. Daarnaast wordt, afgaand op de zwaarste component van de straf, de afdoening van elke zaak ingedeeld in één van zeven hoofdcategorieën: vrijheidsbeperkende maatregelen, onvoorwaardelijke vrijheidsstraffen, taakstraffen, voorwaardelijke vrijheidsstraffen, boetes, transacties en beleidssepts.²³ Met deze aanpak houden we zicht op combinaties van straffen, maar is ook duidelijk wat per zaak de belangrijkste sanctie was. Naast het soort afdoening wordt natuurlijk ook de hoogte of de duur van de straf bijgehouden.

3.3.2 *De operationele definitie van recidive*

Met de keuze van de OBJD als bronsysteem is het *domein* van de Recidivemonitor bepaald. Alleen de criminaliteit die het Openbaar Ministerie te verwerken krijgt, wordt met het onderzoek in kaart gebracht. De delicten die niet worden opgespoord en de delicten die wel worden opgespoord maar niet bij de officier van justitie terechtkomen, blijven buiten beschouwing. De recidive die wordt gemeten, is de 'strafrechtelijke recidive', dat deel van het delictgedrag dat onder aandacht komt van het OM. Strafrechtelijke recidive hangt positief samen met feitelijke recidive. In het algemeen zal het zo zijn dat een persoon die veel nieuwe delicten pleegt, een grotere kans heeft opnieuw te worden vervolgd. De verhouding tussen het aantal nieuwe justitiecontacten en het feitelijke aantal nieuw gepleegde delicten zal echter niet voor iedereen hetzelfde zijn. Sommigen lopen tegen de lamp bij hun eerste nieuwe delict, anderen plegen delicten maar ontspringen de dans. De strafrechtelijke recidive is daarom hoogstens een 'proxy', een representatie van de werkelijke recidive. Maar het begrip staat ook op zichzelf. De recidive zoals gemeten in het onderzoek van de Recidivemonitor staat voor 'terugkeer in het strafrechtelijke systeem'. Het geeft aan of justitie zich opnieuw over een persoon heeft moeten ontfermen. Binnen het gegeven domein van delicten kan de reci-

22 Indien sprake is van een 'strafbare poging tot (uitlokking van) een misdrijf' dan wel 'strafbare voorbereiding' of 'medeplichtigheid', wordt de strafdreiging conform het wetboek van strafrecht gecorrigeerd.

23 Zie bijlage 2 voor een opsomming van de vonnisonderdelen die in de OBJD voorkomen.

dive uiterst betrouwbaar en valide worden vastgesteld. De informatie over de strafzaken komt via COMPAS en het JDS automatisch in de OBJD terecht. De gegevens zijn bovendien, zoals eerder aangegeven, op verschillende niveaus juridisch op juistheid getoetst.

Het *bereik* of de operationele definitie van het recidivebegrip geeft aan welke delicten uit het domein concreet worden gerekend tot de gevallen van recidive. Het JDS bestaat eigenlijk uit twee parallelle systemen, één voor de registratie van overtredingen (de zgn. kantonfeiten) en één voor de registratie van misdrijven (rechtbankfeiten). Overtredingen worden normaal gesproken afgedaan door het kantongerecht, een lagere rechtbank. Misdrijven worden in eerste aanleg behandeld door arrondissementsrechtbanken. Het JDS neemt niet alle overtredingen op, alleen de overtredingen waarvan in de wet staat beschreven dat bij herhaling een zwaardere sanctie volgt, worden bijgehouden. 90% van de overtredingen opgenomen in het JDS zijn verkeersdelicten, waarvan ruim 60% overtredingen van de maximumsnelheid. Onderzoekers van het WODC namen in het verleden vaak alleen de rechtbankfeiten mee. De (verkeers)overtredingen bleven bij recidivemetingen doorgaans buiten beschouwing (Van Emmerik, 1985, 1989; Van der Werff, 1986; Van der Laan, 1991; Spaans, 1994; Leuw, 1995, 1999). De Recidivemonitor sluit bij deze praktijk aan. Alleen misdrijven vallen dus onder het bereik van het onderzoek.²⁴

Bij de metingen van de Recidivemonitor geldt de strafzaak als *tel-eenheid*. Het JDS registreert overtreden wetsartikelen en niet het aantal keren dat deze wetsartikelen zijn overtreden. Als er binnen een strafzaak meer wetsartikelen worden genoemd, is niet duidelijk hoeveel strafbare feiten de verdachte heeft gepleegd. Achter één artikel kan een serie misdrijven schuilgaan, terwijl een combinatie van wetsartikelen voor één delict kan staan. Men kan in het JDS dus geen losse misdrijven tellen, de 'strafzaak' is het kleinste, vaste element waarover informatie wordt bijgehouden. Zaken die worden gevoegd tot één rechtszaak, worden als afzonderlijke strafzaken geteld. Zij staan onder hun eigen parketnummer in het JDS. Een persoon die werd vervolgd voor een nieuwe zaak waarbij drie strafzaken werden gevoegd, heeft volgens de richtlijnen van de Recidivemonitor dus vier keer gerecidiveerd.

Niet alle strafzaken worden meegenomen. Een klein gedeelte – in de periode 1997-2003 9% – eindigt in een technisch sepot of een vrijspraak, of wordt op andere wijze technisch afgedaan. Hier is niet vast komen te staan dat de verdachte de strafbare feiten heeft gepleegd en daarom blijven deze zaken buiten beschouwing. Alleen strafzaken waarin de rechter een schuldigverklaring uitspreekt of die door de officier van justitie worden afgedaan met een beleidssepot of een transactie, tellen mee. In

24 In studies naar verkeerscriminaliteit worden overtredingen wel meegenomen. Zie Blom & Wartna (2004).

dergelijke gevallen is sprake van een ‘geldige afdoening’.²⁵ Strafzaken die nog niet zijn afgedaan of zaken die in hoger beroep dienen, worden ook meegeteld. De kans dat zij eindigen in een technische uitspraak is gering.²⁶

Vaste criteria voor het meten van de recidive

Met het vaststellen van het domein, het bereik en de tel-eenheid is de operationele definitie van het recidivebegrip bepaald. Er is sprake van recidive als zich tijdens de observatieperiode van het onderzoek een nieuwe strafzaak voordoet naar aanleiding van een misdrijf dat niet met een technische beslissing werd afgedaan. Om onderscheid te kunnen maken naar de ernst en de aard van de strafrechtelijke recidive worden naast deze algemene omschrijving additionele criteria gehanteerd. Tabel 3.3 geeft de vaste criteria van de Recidivemonitor weer. De ‘ernstige’ recidive is een deelverzameling van de ‘algemene’ recidive. Bij ernstige recidive gaat het om misdrijven waarvoor minstens vier jaar gevangenisstraf kan worden opgelegd. De ‘zeer ernstige recidive’, met delicten met een strafdreiging van acht jaar of meer, is een subselectie van de ernstige recidive.²⁷

De eerste drie criteria hebben altijd dezelfde betekenis. Dit geldt niet voor ‘speciale’ en ‘specifieke’ recidive. De termen werden eerder gebruikt door Van der Werff (1979, 1986). De invulling van deze criteria hangt af van de aard van de uitgangszaak, de zaak naar aanleiding waarvan de personen in het onderzoek zijn opgenomen. Onder verkeersdelinquenten bijvoorbeeld, is speciale recidive een nieuw justitiecontact vanwege een verkeersdelict. Bij zedendelinquenten is er sprake van speciale recidive wanneer zij opnieuw voor een zedendelict een strafzaak hebben. Specifieke recidive is herhaling van hetzelfde delict. Een dronken rijder die opnieuw voor rijden onder invloed wordt veroordeeld, en een exhibitionist die opnieuw voor schennis van de eerbaarheid is vervolgd, hebben zich schuldig gemaakt aan specifieke recidive.

Aan de hand van de criteria uit tabel 3.3 kan de recidive in een onderzoeksgroep geschakeerd worden beschreven, terwijl tegelijkertijd vergelijkbaarheid is gegarandeerd. De *termijn* waarover de recidive wordt bepaald, staat binnen de Recidivemonitor niet vast, maar bedraagt minimaal twee jaar. Omdat niet alle strafbare feiten even snel bij het OM worden aangemeld, is tijd nodig om het beeld compleet te krijgen. Een kortere observatieperiode zou leiden tot onderschatting van de recidive. Na het plegen en de ontdekking van een delict duurt het enige tijd voordat een verdachte is opgespoord en de zaak kan worden overgedragen aan de officier van justitie. Na twee jaar is echter voldoende tijd verstreken.

25 Formeel staat bij een beleidssepot of een transactie niet vast dat de verdachte de dader is. De verdachte is niet schuldig verklaard of ‘veroordeeld’. Niettemin worden deze gevallen in het onderzoek meegenomen.

26 Bijlage 2 bevat een overzicht van de geldige afdoeningen en de technische beslissingen van strafzaken.

27 Bijlage 3 geeft een indeling van de meest voorkomende misdrijven naar strafdreiging.

Na die tijd zal in zo goed als alle vroege gevallen van recidive het politieonderzoek zijn afgerond, is het delict doorgegeven en zijn de gegevens verwerkt in het JDS. Om het exacte tijdstip van de recidive te bepalen, gaan we uit van de pleegdatum van het nieuwe delict. De start van de observatieperiode valt, als er geen aanvullende informatie beschikbaar is, voor elke persoon samen met de inschrijfdatum van zijn uitgangszaak.²⁸

Tabel 3.3 Vaste recidivecriteria van de WODC-Recidivemonitor

criterium	omschrijving
Algemene recidive	Alle nieuwe strafzaken naar aanleiding van een misdrijf, ongeacht de aard en ernst van de gepleegde delicten, behalve de zaken die eindigen in vrijspraak, een technisch sepot of een andere technische uitspraak (= een 'geldig justitiecontact')
Ernstige recidive	Nieuwe, geldige justitiecontacten n.a.v. een misdrijf met een wettelijke strafdreiging van 4 jaar of meer*
Zeer ernstige recidive	Nieuwe, geldige justitiecontacten n.a.v. een misdrijf met een wettelijke strafdreiging van 8 jaar of meer
Speciale recidive	Nieuwe, geldige justitiecontacten n.a.v. hetzelfde type misdrijf als in de uitgangszaak
Specifieke recidive	Nieuwe, geldige justitiecontacten n.a.v. overtreding van het hetzelfde wetsartikel als in de uitgangszaak

* Delicten met een lagere strafdreiging maar waarvoor wel voorlopige hechtenis kan worden opgelegd, vallen ook in deze categorie.

De observatieperiode kent een minimum duur van twee jaar. Een maximum is er niet. Er bestaat niet één ideale duur van de follow-up. De ideale duur verschilt per onderzoeksgroep en hangt af van de specifieke vragen waarop men een antwoord zoekt. Binnen de Recidivemonitor wordt daarom niet één, maar worden meerdere observatieperioden aangehouden. De prevalentie, frequentie en ernst worden voortschrijdend in de tijd in kaart gebracht. Alle statistieken worden berekend over elk vol jaar dat sinds de uitgangszaak verstreken is. Het kleinste aantal observatiepunten is dus twee: één en twee jaar na de uitgangszaak. Of er meer meetmomenten zijn, hangt af van de 'selectieperiode' en het moment waarop de gegevensverzameling werd stopgezet. Het onderzoek dat plaatsvindt onder ex-terbeschikkinggestelden bijvoorbeeld, kent voor een deel van de groep een observatietermijn van meer dan dertig jaar. Dit onderzoek heeft betrekking op personen die vanaf 1974 uit de tbs werden ontslagen.

3.3.3 Globale opzet van het vaste onderzoek

De Recidivemonitor kent vijf groepen waarnaar doorlopend onderzoek wordt gedaan:

28 Behalve als de zaak is afgedaan met een onvoorwaardelijke vrijheidsstraf. In dat geval wordt de recidive (ook) berekend vanaf de geschatte datum van invrijheidsstelling.

- 1 Volwassenen met een strafzaak wegens een misdrijf, afgedaan met een transactie of een beleidssepot, of een straf of maatregel opgelegd door de rechter.
- 2 Jeugdigen met een strafzaak wegens een misdrijf, afgedaan met een transactie of een beleidssepot, of een straf of maatregel opgelegd door de rechter.²⁹
- 3 Volwassenen die werden ontslagen uit een penitentiaire inrichting (ex-gedetineerden).
- 4 Jongeren uit een justitiële opvang- of behandelinrichting (ex-JJI-pupillen).
- 5 Personen die een intramurale tbs-maatregel kregen opgelegd (ex-terbeschikkinggestelden).

Van deze vijf onderzoekspopulaties wordt jaarlijks de strafrechtelijke recidive bepaald. Centrale vraag is steeds of deze is toe- of afgenomen. In het volgende hoofdstuk zullen de groepen uitvoeriger worden beschreven. De eerste twee bestaan uit bijna alle natuurlijke personen die sinds 1997 als verdachte van een misdrijf in aanraking kwamen met de Nederlandse justitie. De drie overige monitoren richten zich op de personen die zijn uitgestroomd uit een justitiële inrichting. Naast het vaste onderzoek in de vijf daderpopulaties worden overigens ook eenmalige metingen uitgevoerd bij meer specifieke dadergroepen. Ook dan worden de facetten van de recidive op dezelfde wijze gemeten.

Vaste statistieken

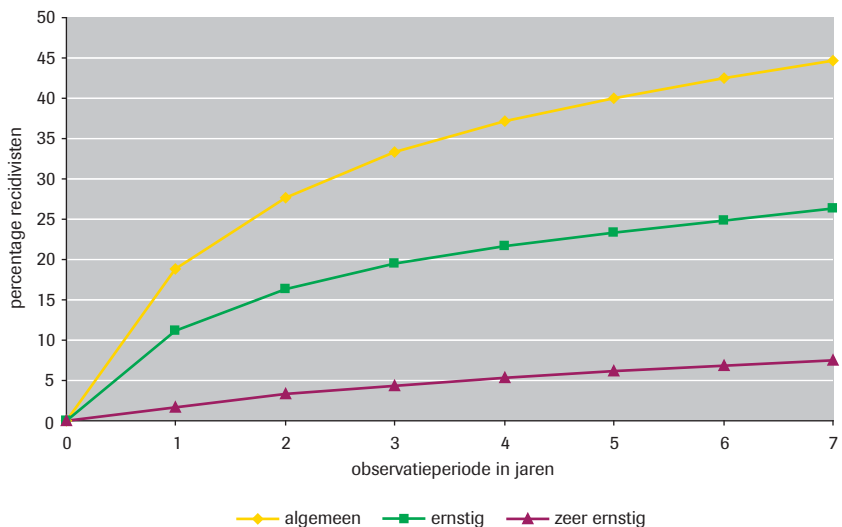
De *prevalentie* van de recidive wordt vastgesteld door het percentage daders te berekenen dat na de uitgangszaak één of meer nieuwe strafzaken had. Dit gebeurt met behulp van survivalanalyse, een techniek die in paragraaf 2.3 werd geïntroduceerd. Survivalanalyse houdt rekening met verschillen in de duur van de periode dat de personen in het onderzoek konden worden gevolgd. De kans op recidive in een groep wordt niet eenmalig geschat, maar op elk moment van de observatieperiode, op basis van de waarnemingen die op dat moment beschikbaar zijn (zie o.m. Klein & Moeschberger, 1997). Het voordeel van survivalanalyse is dat niet elke respondent de volle lengte van de observatietermijn hoeft te worden gevolgd. Als er voldoende waarnemingen zijn, kan het percentage aan het slot van de termijn op grond van een deel van de groep worden vastgesteld.

Figuur 3.1 geeft de voortschrijdende recidivepercentages te zien van de groep volwassen daders met een strafzaak afgedaan in 1997 (Wartna, Tollenaar & Blom, 2005). Bij deze onderzoeksgroep bedroeg de observatieperiode minimaal zes en maximaal zeven jaar. Omdat niet iedereen de

²⁹ Omdat personen die van de officier van justitie een transactie kregen of een beleidssepot, formeel niet zijn vervolgd, zullen de eerste twee groepen worden aangeduid als volwassenen en jeugdigen 'met een strafzaak'.

volle periode werd geobserveerd, zijn de percentages voor het laatste jaar geschat op basis van een deel van de groep. Tot zes jaar waren de waarnemingen compleet. In de figuur zien we hoe de onderscheiden recidivecriteria zich empirisch tot elkaar verhouden. De algemene recidive bedraagt na zeven jaar 45%. Van de daders met een strafzaak afgedaan in 1997 heeft 45% binnen zeven jaar na de uitgangszaak een nieuw justitiecontact naar aanleiding van enig misdrijf. De ernstige recidive, nieuwe justitiecontacten naar aanleiding van misdrijven met een strafdreiging van vier jaar of meer, is na zeven jaar 26%. De verhouding tussen beide vormen van recidive is stabiel. De ernstige recidive maakt – in deze groep – stevast bijna 60% uit van de algemene recidive. Zeer ernstige recidive, met delicten met een strafdreiging van acht jaar of meer, komt veel minder vaak voor. Na zeven jaar is het percentage daders vervolgd voor een zeer ernstig misdrijf opgelopen tot 7,5%. Figuur 3.1 toont naast de prevalentie ook de recidive-snelheid. Hoe eerder de percentages een bepaald niveau bereiken, des te sneller treedt de recidive op.

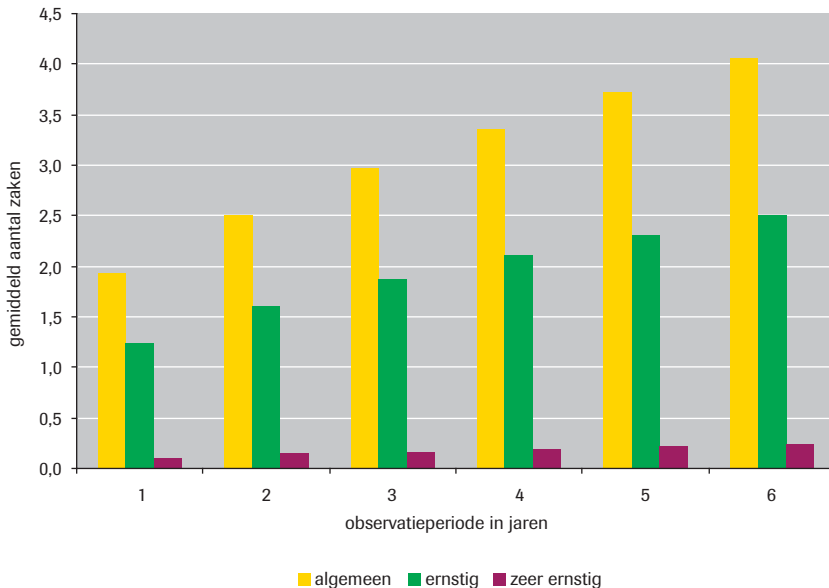
Figuur 3.1 Prevalentie van algemene, ernstige en zeer ernstige recidive onder volwassenen met een strafzaak afgedaan in 1997 (N=133.331)



De *frequentie* van de recidive wordt uitgedrukt in het gemiddelde aantal nieuwe justitiecontacten. Deze statistiek wordt in de Recidivemonitor alleen berekend over degenen die hebben gerecidiveerd, de non-recidivisten tellen niet mee. Ook de gemiddelde frequentie kan cumulatief en voortschrijdend in de tijd worden afgebeeld. In tegenstelling tot de

prevalentie moet bij het berekenen van de frequentie de observatieperiode voor alle respondenten even lang zijn, anders is de vergelijkbaarheid niet optimaal. Figuur 3.2 geeft de cumulatieve, gemiddelde recidivefrequenties voor de recidivisten onder de volwassen daders met een strafzaak afgedaan in 1997, uitgesplitst naar de ernst van de gepleegde delicten. Na zes jaar hebben degenen die opnieuw met justitie in aanraking kwamen gemiddeld vier nieuwe strafzaken. Daarvan waren er gemiddeld 2,5 naar aanleiding van ernstige delicten. Zeer ernstige recidive komt veel minder vaak voor.

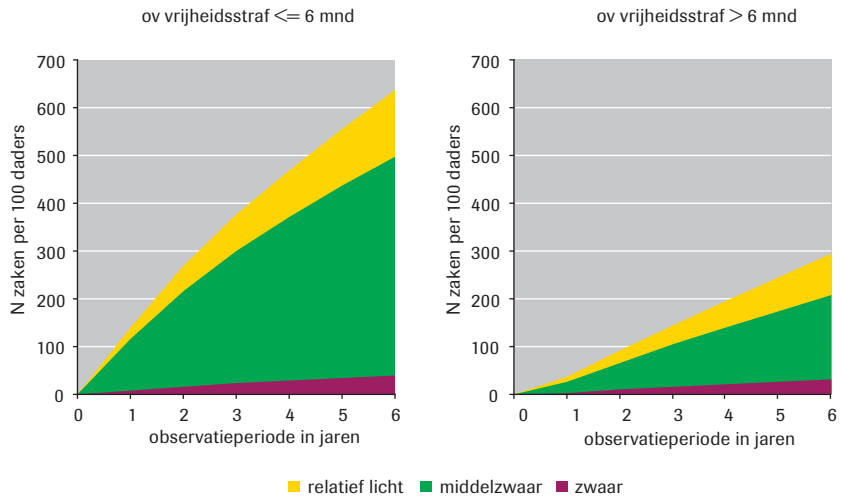
Figuur 3.2 Gemiddeld aantal nieuwe justitiecontacten voor de recidivisten onder volwassen daders met een strafzaak afgedaan in 1997; naar ernst van de recidive



De *ernst* van de recidive wordt binnen de monitor op verschillende manieren in beeld gebracht. Op de eerste plaats wordt er bij het meten van de prevalentie, snelheid en frequentie onderscheid gemaakt tussen algemene, ernstige en zeer ernstige recidive. Verder kan van elke onderzoeksgroep een tabel worden getoond van de verdeling van de recidivedelicten naar type delict en soort afdoening. De aard van de recidivedelicten wordt ingedeeld in zeven klassen. Hetzelfde geldt voor de afdoeningen van de recidivezaken. Deze zijn grofweg naar zwaarte gerangschikt. Een vrijheidsbenemende maatregel (een terbeschikkingstelling of een plaatsing in een inrichting voor jeugdigen) wordt opgevat als de zwaarste sanctie, een

transactie in termen anders dan een geldbedrag als de lichtste. Bij combinatievonnissen is het zwaarste onderdeel bepalend voor de indeling.

Figuur 3.3 Omvang van de recidive onder volwassen daders met een korte of lange vrijheidsstraf opgelegd in 1997; naar ernst van de vervolgte delicten



Ook in de weergave van ‘omvang’ van de recidive wordt de ernst van de delicten verdisconteerd. De omvang van de recidive is een overkoepelende maat, een samenvatting van de prevalentie, snelheid, frequentie en ernst. De omvang geeft, voortschrijdend in de tijd, het totale aantal recidivezaken in een onderzoeksgroep, onderscheiden naar de ernst van de vervolgte delicten. Bij het berekenen van de omvang tellen de non-recidivisten mee. Hoe meer non-recidivisten er zijn, des te geringer de omvang van de recidive. Tenzij natuurlijk de recidivisten die er zijn vaker dan gemiddeld hebben gerecidiveerd. Dan wordt het totaalbeeld weer anders. Figuur 3.3 geeft de recidiveomvang voor de kort- en langgestraften uit 1997. Om een goede vergelijking mogelijk te maken worden de aantallen berekend per groep van honderd daders. Zo is in één oogopslag te zien dat de recidive onder kortgestraften, met straffen tot zes maanden, veel omvangrijker is dan onder personen met een vrijheidsstraf langer dan zes maanden. Honderd kortgestraften hebben binnen zes jaar na de oplegging van hun straf in totaal 638 nieuwe justitiecontacten op hun conto bijgeschreven. Het merendeel daarvan (circa 80%) betreft een ‘middelzwaar’ delict met een strafdreiging van vier tot acht jaar. De omvang van de reci-

dive onder langgestraften is met een totaal van 295 nieuwe strafzaken per honderd daders in zes jaar tijd, ruim twee keer zo klein.³⁰

3.4 Analyse van de recidivecijfers

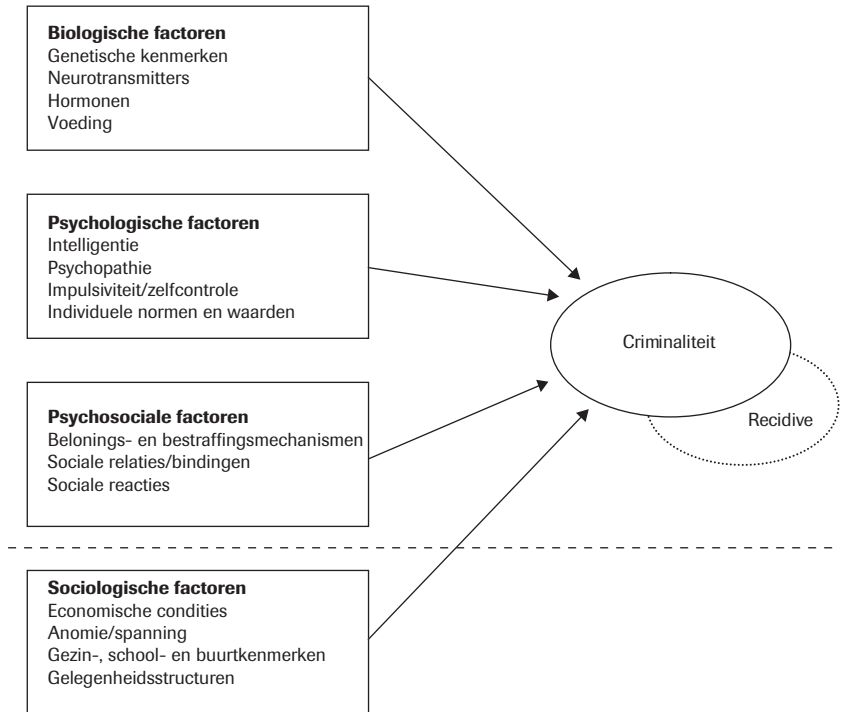
De procedures die tot dusver zijn besproken, leveren een beschrijving op van de recidive in de onderzoeksgroepen. De cijfers laten zien hoe omvangrijk de terugval is en, indien de groep over een langere periode wordt gevolgd, welke ontwikkelingen zich hierin hebben voorgedaan. Een antwoord op de vraag *waarom* de recidivecijfers zijn gestegen of gedaald, bieden de statistieken van de Recidivemonitor niet. Om de cijfers te verklaren zijn aanvullende gegevens en analyses nodig. We bespreken drie soorten factoren die gerelateerd kunnen worden aan de cijfers en die bepalend zijn voor de hoogte van de recidive: causale factoren, beschrijvende kenmerken en registratie-effecten.

3.4.1 Causale factoren

Sommige daders vallen terug, andere niet. Wat onderscheidt de een van de ander? Zoeken naar de oorzaken van recidive is als het zoeken naar de oorzaken van criminaliteit. Recidive kan net als delictgedrag in het algemeen worden gezien als de uitkomst van een proces waarop verschillende soorten van factoren van invloed zijn. In hun overzicht van de criminologie bespreken Lissenberg, Van Ruller en Van Swaaningen (2001) hedendaagse criminologische theorieën. Zij leggen daarbij een grof onderscheid aan tussen biologische, psychologische, psychosociale en sociologische benaderingen. Op elk van deze niveaus zijn causale factoren aan te wijzen die het vóórkomen van criminaliteit, of meer in het bijzonder het optreden van recidive, kunnen verklaren. Figuur 3.4 geeft hiervan een schematische weergave. De niveaus sluiten elkaar niet uit, maar meestal ligt het accent van een criminologische theorie binnen één van de vier categorieën.

30 Onder de kortgestraften bevinden zich relatief veel 'veelplegers', vandaar het grote verschil in de omvang van de recidive. Deze cijfers zijn berekend vanaf de inschrijfdatum van de uitgangszaak. Zij omvatten dus de periode dat men vastzat in de inrichting. Er zijn ook prevalentiecijfers beschikbaar die zijn berekend vanaf de datum van invrijheidsstelling, dus exclusief het 'incapacitatie-effect' van de oorspronkelijke detentie.

Figuur 3.4 Voorbeelden van causale factoren die een rol spelen bij de verklaring van (herhalings)criminaliteit



Natuurlijk is ook een andere indeling mogelijk. In hun themaboek over geweld onder jongeren gebruiken De Haan et al. (1999) drie categorieën om de risicofactoren die zij relevant achten, te rubriceren. Zij onderscheiden persoonlijke kenmerken (zoals genetische, fysiologische en psychologische factoren), sociaal-culturele factoren (gezin, school, groeps- en buurtkenmerken) en algemeen-maatschappelijke factoren, waaronder sociale samenhang en economische factoren. Het theoretisch kader dat de onderzoeker hanteert, bepaalt welke factoren of welke soorten van factoren bij het zoeken naar de oorzaken van (herhalings)criminaliteit van belang worden geacht. Het theoretisch kader biedt een leidraad voor de interpretatie van de uitkomsten van het onderzoek. De benadering die men kiest is vaak afhankelijk van de dadergroep die wordt bestudeerd. In onderzoek onder serieverkrachters zullen de oorzaken van de recidive doorgaans in een andere hoek worden gezocht dan bij een studie onder adolescenten die zich schuldig hebben gemaakt aan vandalisme. Niet alle criminologen denken daar overigens hetzelfde over. Sommigen gaan ervan uit dat de oorzaken van criminaliteit universeel van aard zijn. Zij

onderschrijven algemene criminologische theorieën en proberen uiteenlopende vormen van normoverschrijdend gedrag met behulp van enkele centrale concepten te verklaren. Anderen stellen dat er bij uiteenlopende dadergroepen verschillende ideeën nodig zijn om criminaliteit te doorgronden. Zij vinden het acceptabel dat er verschillende theorieën naast elkaar bestaan, met elk hun eigen toepassingsgebied.

Meta-analyses van oorzaken van criminaliteit

Bij breed opgezet onderzoek naar de oorzaken van crimineel gedrag, onder grote groepen volwassenen of jeugdige daders, kunnen algemene theorieën allicht behulpzaam zijn. In een meta-analyse van 214 meest Amerikaanse studies test Pratt (2001) de empirische evidentie van negen bekende, criminologische theorieën: social disorganisation, anomie/strain, absolute deprivation/conflict, relative deprivation, routine activities, deterrence/rational choice, support/altruïsme en subculturele theorieën.³¹ Pratt inventariseerde welke indicatoren in de studies werden gebruikt om de geldigheid van de theorieën te testen en ging vervolgens na welke typen indicatoren consistent samenhangen met de in de studies gemeten omvang van de criminaliteit. De criminaliteit werd steeds op een andere manier gemeten, maar door het gebruik van een overkoepelende maat waarin de sterkte van de samenhang tussen de indicatoren en de gemeten criminaliteit wordt uitgedrukt, konden de uitkomsten van de individuele studies onderling worden vergeleken.

Zaken die wijzen op 'geconcentreerde achterstand', zoals het percentage niet-blanke personen, het aandeel gebroken gezinnen en de hoeveelheid armoede, bleken het criminaliteitsniveau het beste te voorspellen. Volgens de indeling van Pratt zijn dit indicatoren van de social disorganisation-theorie en de absolute deprivation/conflict-theorie. De social disorganisation-theorie stelt dat criminaliteit ontstaat in buurten waarin instituten als school, kerkgemeenschappen en vrijwillersorganisaties in verval raken en niet langer de sociale controle kunnen uitoefenen die nodig is om de bewoners in het gareel te houden. De absolute deprivation/conflict-theorie gaat ervan uit dat 'criminaliteit' een sociaal construct is dat voortkomt uit de animositeit tussen de *haves* en de *havenots*. Voor deze twee theorieën werd in de 214 studies dus de meeste steun gevonden. Andere theorieën, zoals de 'rational choice/deterrence'-benadering, waarin de calculerende dader centraal staat, en subculturele theorieën van criminaliteit die een centrale rol toekennen aan het bestaan van afwijkende normen en waarden in grootstedelijke jeugdbendes en -groepen, kwamen

31 Het voert te ver om deze theorieën in detail toe te lichten. We laten hun namen ook onvertaald. Voor een Nederlandstalige bespreking, zie Lissenberg, Van Ruller en Van Swaaningen (2001). Ook Vold, Thomas en Snipes (2002) geven een breed overzicht van gangbare criminologische theorieën, evenals Ellis en Walsh (2000) en Lilly et al. (1997). Tierney (1996) en Downes en Rock (2003) beperken zich tot een uiteenzetting van sociologische benaderingen. Akers en Jensen (2003) redigeerden een themaboek over sociale leertheorieën.

er minder goed van af. De indicatoren van deze theorieën, zoals het aantal politieagenten en de urbanisatiegraad, bleken een geringe gemiddelde effectgrootte te hebben. Zij vertoonden dus geen sterke, stabiele samenhang met de omvang van de criminaliteit.

Pratts meta-analyse levert een opsomming op van factoren die voor de interpretatie van de uitkomsten van de recidivemetingen van belang kunnen zijn. De vraag is of zijn resultaten ook opgaan in Nederland, waar de maatschappelijke tegenstellingen wellicht wat minder scherp zijn dan in Amerika. Bovendien zijn er meer grootheden die een rol spelen bij het al dan niet vóórkomen van recidive. Een belangrijke beperking van de studies die Pratt selecteerde is dat de metingen zich afspeelden op macro-niveau. Algemene statistieken over de relevant geachte causale factoren worden in verband gebracht met het niveau van de criminaliteit in een buurt, een stad of een regio. De onderzoekseenheden waren dus groepen en geen personen. Het overzicht van Pratt mist hierdoor de factoren die op individueel niveau een verklaring kunnen bieden voor de prevalentie van criminaliteit.

In de meta-analyse van Cottle, Lee en Heilbrun (2001) krijgen individuele kenmerken wel een kans. Deze auteurs overzagen 23 empirische studies, gericht op de verklaring van recidivisme bij jeugdige daders.³² Dertig factoren werden onderscheiden, ingedeeld in acht categorieën: demografische variabelen, delictgeschiedenis, sociale factoren en gezinsrelaties, schoolvariabelen, intellectuele kenmerken, het gebruik van genotsmiddelen, klinische factoren (psychopathologie, gedragsstoornissen) en het gebruik van gecombineerde risicoscores. Op basis van de geselecteerde studies berekenen Cottle, Lee en Heilbrun de gewogen gemiddelde effectgrootte van de voorspellers van recidive van jeugdigen.³³

32 Cottle, Lee en Heilbrun merken terecht op dat etiologisch onderzoek naar recidivisme andere uitkomsten kan opleveren dan studies naar de oorzaken van criminaliteit in het algemeen. Bij recidiveonderzoek heeft men immers te maken met minder heterogeniteit; alle respondenten vertoonden ooit delinquent gedrag.

33 Zie paragraaf 2.5 voor uitleg van de technische termen.

Tabel 3.4 Voorspellers van recidive bij jeugdigen (N_{TOT}=15.265)

Factor	Z _r	p	n	k	n _x
Leeftijd bij eerste opname	-0,346	**	720	3	2.273
Leeftijd bij eerste justitiecontact	-0,341	**	1.225	8	3.298
Lichte vorm van psychopathologie	0,305	**	953	7	2.244
Gezinsproblemen	0,277	**	1.054	5	2.165
Gedragsproblemen	0,255	**	1.667	7	6.949
Vrijtijdsbesteding	-0,233	**	588	2	1.343
Delinquente vrienden	0,204	**	1.525	7	2.842
Duur van eerste detentie	0,187	**	641	3	1.022
Aantal uithuisplaatsingen	0,184	**	424	2	617
Aantal eerdere opnames	0,174	**	585	3	699
Type delict	0,159	**	10.267	7	81.345
Score vaardigheidstest	-0,153	**	506	3	599
Overmatig drugsgebruik	0,149	**	1.111	6	1.273
Totaalscore IQ-test	-0,142	**	1.756	5	5.014
Deelname aan speciaal onderwijs	0,130	*	432	2	473
Gebruik risicotaxatie-instrument	0,118	**	10.353	6	60.117
Slachtofferschap misbruik	0,112	**	9.949	5	59.436
Gender (mannelijk)	0,111	**	9.671	3	58.698
Verbale score IQ-test	-0,111	*	716	4	522
Eenoudergezin	0,070	**	10.501	5	3.793
Ernstige psychopathologie	0,069		346	2	ns
Etniciteit (niet-blanke minderheden)	0,067	**	10.121	6	30.018
Socio-economische status	0,065	**	10.363	3	36.703
Aantal eerdere arrestaties	0,058	**	10.155	7	26.145
Aanwezigheid op school	-0,048		299	2	ns
Pathologie bij ouders	0,047		529	3	ns
Vaardigheidsscore IQ-test	-0,031		491	2	ns
Schooldiploma	-0,028		10.025	6	ns
Hulpverleningsgeschiedenis	0,019		9.366	2	ns
Drugsgebruik	0,014		9.366	2	ns

Z_r = gewogen gemiddelde effect size, k = aantal unieke steekproeven, n_x = aantal respondenten bij wie het verband zou moeten ontbreken om het gevonden resultaat non-significant te maken.

* p < 0,01; ** p < 0,001.

Bron: Cottle, Lee & Heilbrun, 2001 (eigen vertaling)

De resultaten staan weergegeven in tabel 3.4. Hoe hoger de waarde Z_r, hoe belangrijker de factor is in het voorspellen van latere recidive. Naast de effectgrootte en het significantieniveau van de voorspellers geeft de tabel het aantal jeugdigen en het aantal onderzoeksgroepen waarbij het verband met recidive kon worden vastgesteld. De waarde in de

laatste kolom van de tabel geeft aan hoe robuust het gevonden verband is. Dit heeft te maken met het *file drawer*-probleem (zie 2.5). Studies met non-significante resultaten worden minder vaak gepubliceerd in internationale tijdschriften en hebben dus een kleinere kans om in een meta-analyse te worden opgenomen. N_x geeft de omvang van een fictieve steekproef waarbij het verband tussen de voorspeller en recidive *niet* zou moeten worden gevonden, om het verband dat wel werd aangetroffen te ontcrachten. Hoe hoger n_x , des te overtuigender is de gevonden samenhang tussen de voorspeller en de recidive met de geselecteerde studies.³⁴ De sterkste voorspellers van recidive bij jeugdigen staan boven in de lijst. Bij de top vijftien zijn vijf van de zes factoren uit het domein 'delictgeschiedenis' opgenomen en vier van de zeven sociale en gezinsfactoren. Deze twee domeinen zijn volgens Cottle, Lee en Heilbrun dus het belangrijkste. Andere belangrijke determinanten van recidive bij jeugdigen zijn de aanwezigheid van lichte vormen van psychopathologie, zoals stress en angstaanvallen, en gedragstoornissen.

Bonta, Law en Hanson (1998) voerden een meta-analyse uit van de recidive bij volwassen daders met een psychische stoornis. Zij gebruikten dezelfde indeling van factoren als Cottle, Lee en Heilbrun. Men zou verwachten dat bij deze groep klinische variabelen zoals intelligentie, de aan- of afwezigheid van stemmingswisselingen of psychoses een grotere rol zouden spelen, maar dat bleek niet het geval. De enige klinische factor die een verklaring vormde voor de recidive onder psychisch gestoorde daders was het bestaan van een antisociale persoonlijkheid. Voor de rest bleken dezelfde factoren van belang als bij 'normale' daders. De beste voorspellers bevonden zich wederom in het domein 'delictgeschiedenis'. Bijna alle indicatoren uit dit domein bleken significant. Ook demografische factoren speelden een rol. Leeftijd, sekse en burgerlijke staat bleken in de 58 studies die Bonta, Law en Hanson bundelden consistent gerelateerd aan de kans op recidive. Gebrekkige woonomstandigheden, problematische gezinsrelaties en drugsgebruik droegen er ook aan bij. De conclusie van Bonta, Law en Hanson luidt dat recidive bij psychisch gestoorde beter kan worden verklaard vanuit een algemene sociaalpsychologische benadering dan vanuit het perspectief van de psychopathologie. De behandeling van hun psychische stoornis mag goed zijn voor het welbevinden van de patiënten, maar leidt volgens de auteurs niet per se tot reductie van de kans op criminele recidive.

Zijn de oorzaken van criminaliteit dan toch universeel? Bonta, Law en Hanson vonden geen verschillen tussen psychisch gestoorde en normale daders. Bij seksuele delinquenten liggen de zaken wellicht anders. Hanson en Bussière (1998) overzagen 61 studies waarin werd gezocht naar de beste voorspellers van (seksuele) recidive bij dit type daders. Voor algemene

34 Zie Rothstein (2007) voor een bespreking van deze en andere methoden om een inschatting te maken van *publication bias* in de uitkomsten van een meta-analyse.

recidive kwamen dezelfde voorspellers uit de bus als onder niet-seksuele delinquenten: jonge, mannelijke en ongetrouwdeaders met eerdere justitie- of hulpverleningscontacten lopen het grootste risico. Bij de voorspelling van seksuele recidive zijn deze factoren ook van belang, maar minder dan factoren die wijzen op seksuele deviantie, gemeten via fallometrie of gestandaardiseerde vragenlijsten. Volgens Hanson en Bussière is het van belang bij het gebruik van risicotaxatie-instrumenten onder seksuele delinquenten onderscheid te maken tussen de predictie van algemene en de predictie van speciale recidive. In feite stellen zij hiermee vast dat seksuele recidive andere oorzaken kent dan recidive in het algemeen. Gendreau, Little en Goggin (1996) bespreken de resultaten van 131 studies gericht op de predictie van algemene recidive bij volwassenen. Deze meta-analyse is van belang omdat er voor het eerst gebruik wordt gemaakt van het containerbegrip 'criminogene behoeften'. De auteurs volgen Andrews en Bonta (1994), die een onderscheid maken tussen 'statische' en 'dynamische' criminogene factoren. Statische factoren zijn kenmerken die gerelateerd zijn aan de kans op recidive, maar die niet kunnen worden beïnvloed. Leeftijd, sekse, het criminele verleden of een voorgeschiedenis van ander probleemgedrag zijn voorbeelden van statische criminogene factoren. Dynamische criminogene factoren zijn wel veranderbaar. Het kan gaan om attitudes, cognities en gedrag, maar ook om leefomstandigheden en vaardigheden. Volgens Andrews en Bonta vormen dynamische factoren aangrijpingspunten voor behandelprogramma's gericht op de reductie van recidive.

Gendreau, Little en Goggin laten zien dat de dynamische factoren de belangrijkste voorspellers zijn van recidive onder volwassenen, naast het criminele verleden, de aanwezigheid van een antisociale persoonlijkheid en de omgang met delinquente vrienden. Zonder verdere toelichting stellen zij dat deze indicatoren passen bij sociale-leertheorieën (Burgess & Akers, 1966; Akers, 1998). Deze theorieën benadrukken dat crimineel gedrag – net als elk ander gedrag – het gevolg is van leerprocessen die plaatsvinden in de sociale omgeving. Crimineel gedrag is aangeleerd en onderhevig aan subtiele, meest informele bekrachtigingsmechanismen als waardering en afkeuring, beloning en straf. De genoemde factoren zijn belangrijker dan sociale klasse, intelligentie en klinische factoren zoals stress en angstgevoelens. Helaas geven de auteurs niet aan welke factoren uit de 131 studies zij precies tot het domein van de 'criminogene behoeften' hebben gerekend. Uit eerdere en latere publicaties wordt echter duidelijk dat het gaat om antisociale attitudes, drugsgebruik, gebrek aan inlevingsvermogen, impulsiviteit, gebrekkige sociale vaardigheden en problemen op het vlak van werk, opleiding en het omgaan met geld.

De term ‘criminogene behoeften’ is een verzamelterm voor factoren waarvan via onderzoek is gebleken dat zij verband houden met het vóórkomen van criminaliteit. Hij is verbonden met de *what works*-benadering, de onderzoekstraditie waarin wordt nagegaan welke interventies succesvol zijn in het terugdringen van de recidive. Welk pakket van criminogene factoren men concreet op het oog heeft en welke ‘evidence-based’ interventies men als gevolg daarvan voorstaat, varieert. Afhankelijk van de dadergroep wisselen de samenstelling en het relatieve gewicht van de kritieke factoren.

De leefgebieden van de RISC

In Nederland werd voor de reclassering een algemeen diagnose-instrument ontwikkeld op basis van de bevindingen uit de *what works*-literatuur (Adviesbureau Van Montfoort/Reclassering Nederland, 2004). Dit instrument, de RISC, wordt sinds 2004 ingezet bij het maken van voorlichtingsrapportages. Afname van de RISC leidt tot een onderbouwde schatting van het recidiverisico, maar ook van de mate waarin dit risico nog te beïnvloeden valt. In navolging van onderzoekers die in Engeland werkten aan OASys, een soortgelijk instrument, delen de makers van RISC de dynamische criminogene factoren van de (volwassen) verdachten in tien ‘leefgebieden’ in (zie box 3.2). Op deze gebieden zoekt men dus naar oorzaken van het (aanhoudende) criminele gedrag. Daarnaast verzamelen de reclasseringswerkers ook informatie over het delictverleden en het uitgangsdelict. Van der Knaap, Leenarts en Nijssen (2007) onderzochten de interbeoordelaarsbetrouwbaarheid, interne consistentie en congruente validiteit van de RISC.³⁵ Het instrument scoorde een voldoende op elk van deze psychometrische kwaliteiten.

Box 3.2 De 10 leefgebieden van dynamische, criminogene factoren opgenomen in de RISC

1	huisvesting en wonen	6	drugsgebruik
2	opleiding, werk en leren	7	alcoholgebruik
3	inkomen en omgaan met geld	8	emotioneel welzijn
4	relaties met partner, gezin en familieleden	9	denkpatronen, gedrag en vaardigheden
5	relaties met vrienden en kennissen	10	(procriminele) houding

³⁵ De congruente of ‘soortgenoot’-validiteit werd ingeschat met behulp van StatRec, een actuairieel predictie-instrument dat op basis van het populatieonderzoek van de Recidivemonitor kon worden ontwikkeld (zie paragraaf 3.4.2).

De bespreking van de meta-analyses en de introductie van de RISc leveren een bloemlezing op van factoren die, als zij worden geplaatst in een theorie over crimineel gedrag, een verklaring zouden kunnen bieden voor de hoogte van de recidive binnen een onderzoeksgroep. Compleet is het overzicht niet. Zo ontbreken fysiologische kenmerken aan de lijst. Het onderzoek naar de biologische fundamenteën van antisociaal gedrag is sterk in opkomst (zie o.m. Raine, 1993; Raine et al., 1997; Bock & Goode, 1996; Rose, 2000; Rowe, 2002; Walsh, 2002). Popma (2006) onderzocht de rol van het stresshormoon cortisol bij jongeren met een Halt-afdoening. Fishbein et al. (2006) ontdekten een verband tussen de regulatie van dit hormoon en de kans op behandel succes bij Amerikaanse gedetineerden. In Nederland deed Zaalberg (2007) in navolging van Gesch et al. (2002) onderzoek naar de effecten van voedingssupplementen op agressief en regelschendend gedrag bij jongvolwassen gedetineerden. Het lijkt slechts een kwestie van tijd voordat de eerste meta-analyses verschijnen waarin de effectgroottes van factoren uit het biologische, sociale en psychologische domein tegen elkaar worden afgezet. Tegelijk zal ook de theorievorming over de interactie tussen de verschillende domeinen toenemen. Want constateren dat biologische en sociale factoren op elkaar inwerken is één ding, begrijpen hoe dat precies in zijn werk gaat is nog iets anders.³⁶

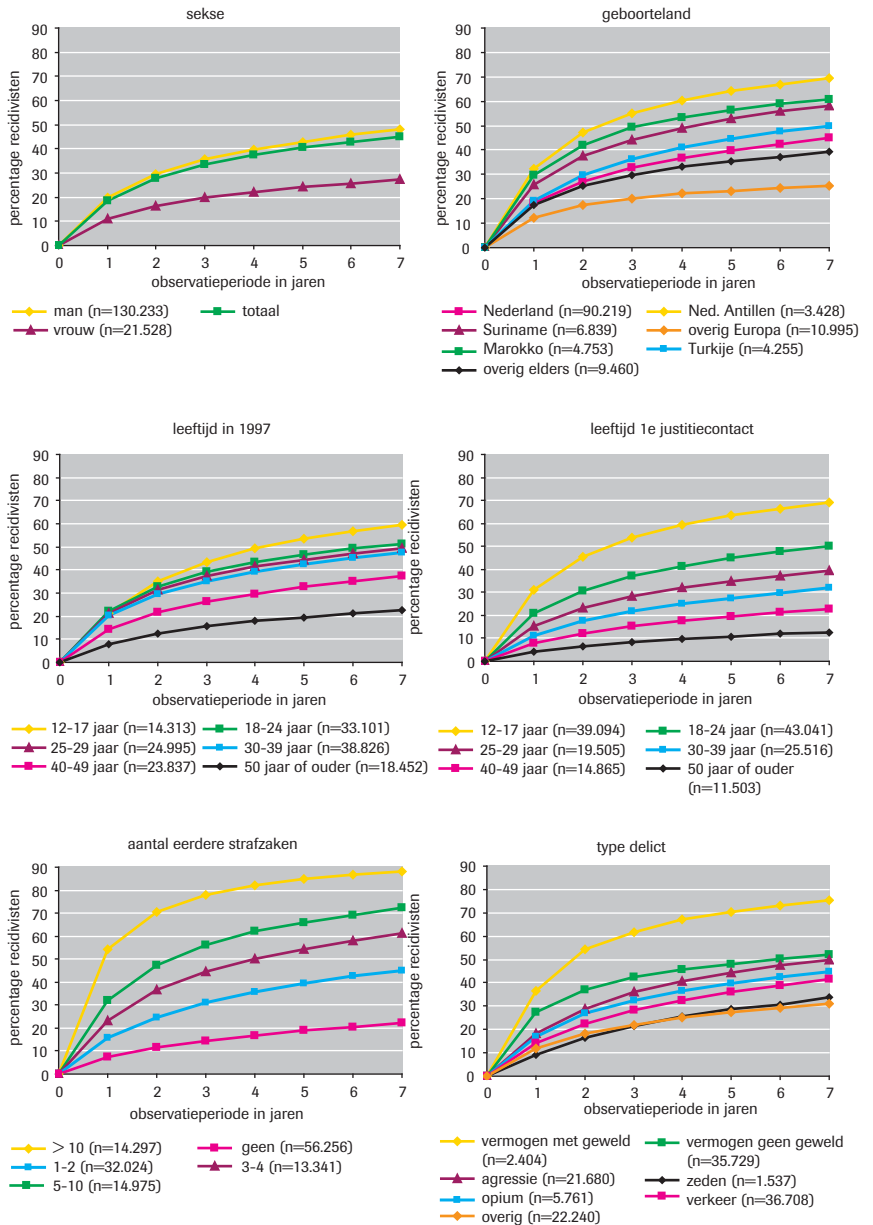
3.4.2 *Moderatoren: uitsplitsingen naar achtergrondkenmerken*

De OBJD bevat geen fysiologische kenmerken van daders en geeft ook geen informatie over de psychosociale omstandigheden waarin zij verkeerden in de periode voor en na de uitgangszaak. De OBJD bevat slechts enkele eenvoudige achtergrondkenmerken: de sekse, het geboorteland en de geboortedatum. Daarnaast houdt het systeem justitiële gegevens bij: delictinformatie, de namen van de instanties die zich over de strafzaken hebben gebogen, en gegevens over de afdoening van de strafzaken. Op basis van de laatste gegevens kan van elke groep de strafrechtelijke recidive worden berekend. Aan de hand van de persoons- en zaakkenmerken kan de recidive nader worden omschreven. Figuur 3.5 laat zien hoe de recidivepercentages onder personen die in Nederland werden vervolgd, uitvallen als ze worden uitgesplitst naar sekse, geboorteland, leeftijd, het aantal eerdere justitiecontacten en type delict. De gegevens zijn afkomstig van een meting onder alle volwassen en jeugdige daders met een strafzaak afgedaan in 1997 (Wartna, Tollenaar & Blom, 2005). Het gaat in het totaal om ruim 130.000 strafrechtelijke meerderjarigen en 15.000 minderjarigen.³⁷

36 Voor een bespreking van de ontwikkelingen op het raakvlak van biologie en criminologie, zie onder meer Blokland, Thienpont & Donker (2005) en De Kogel (2008).

37 Een klein percentage (0,8%) van de minderjarigen werd berecht volgens het volwassenenstrafrecht.

Figuur 3.5 Prevalentie van algemene recidive onder daders vervolgd in 1997; uitgesplitst naar verschillende persoonskenmerken



Figuur 3.5 maakt duidelijk dat de kans op nieuwe justitiecontacten samenhangt met de sekse, de leeftijd en het geboorteland van de dader, het aantal eerdere justitiecontacten dat men heeft opgebouwd en het type delict bij de uitgangszaak. Grofweg kunnen we stellen dat:

- mannen en jongens een hogere kans op recidive hebben dan vrouwen en meisjes;
- de kans op recidive onder daders geboren in sommige niet-westerse landen groter is;
- bij jonge daders het recidiverisico hoger is dan bij oudere;
- het risico hoger is naarmate men jonger was ten tijde van de eerste strafzaak;
- de recidivekans stijgt met het aantal eerdere strafzaken; en
- het risico het grootst is na een vermogensdelict met gebruikmaking van geweld.

Op zich zijn dit geen verrassende resultaten. De rol van het criminele verleden kwam bij de eerder besproken meta-analyses al aan het licht. Bij grootschalige recidivestudies in het buitenland zijn soortgelijke verbanden aangetroffen. In Engeland (Lloyd, Mair & Hough, 1994), Schotland (Scottish Executive, 2003), Frankrijk (Kensey & Tournier, 2005), Noorwegen (SSB, 2002) en de Verenigde Staten (Langan & Levin, 2002) kwamen sekse, leeftijd en het aantal eerdere veroordelingen steeds als de beste voorspellers van de kans op recidive uit de bus. De samenhang met deze achtergrondkenmerken is een bekend criminologisch feit, dat al werd beschreven in Engelse en Duitse rapporten uit de eerste helft van de vorige eeuw.³⁸ In Nederland was Van der Werff (1978) de eerste die het verband tussen sekse, leeftijd, delictgeschiedenis en het vóórkomen van recidive op deze schaal met empirisch materiaal beschreef.

Uit de meta-analyse van Pratt (2001) kwam naar voren dat het percentage niet-blanken in een gemeenschap indicatief is voor de omvang van de criminaliteit. Uit Nederlandse gegevens blijkt dat Antilliaanse, Marokkaanse, Surinaamse en Turkse daders vaker en sneller recidiveren dan daders die zijn geboren in Nederland. De vraag welke oorzaken hieraan ten grondslag liggen en met welke theoretische noties het verschil in recidive tussen allochtone en autochtone Nederlanders het best kan worden verklaard, is hier niet aan de orde. De rol van de kenmerken die in de OBJD zijn opgenomen, is niet *explicatief* van aard. Ze worden louter *descriptief* gebruikt. Het zijn moderatoren, beschrijvende kenmerken die worden aangewend om de strafrechtelijke recidive onder justitiebelen in meer detail in kaart te brengen. Op grond van hun correlatie met de uitkomstmaat worden zij binnen de Recidivemonitor ingezet als controle- of matchingvariabelen die ervoor zorgen dat de opeenvolgende

38 Mayr (1917), geciteerd door Kerner (1996); Command Paper (1935, 1936), geciteerd door Radzinowicz (1945).

dadergroepen onderling beter vergelijkbaar zijn (zie verder 3.5). Ze worden niet, althans niet in dit proefschrift, opgevat als causale factoren die een verklaring voor de recidivecijfers kunnen geven.

Multivariate analyse

Een vraag die binnen de Recidivemonitor wel aan de orde komt, is hoe de kenmerken uit figuur 3.5 gezamenlijk samenhangen met de kans op recidive. Afgaand op de spreiding in de percentages lijken sekse, startleeftijd (dit is de leeftijd ten tijde van het eerste justitiecontact) en het aantal eerdere justitiecontacten de sterkste voorspellers te zijn. Maar er is sprake van ‘multicollineariteit’: een deel van de samenhang tussen een predictor en de criteriumvariabele – in dit geval de recidive – moet worden toegeschreven aan samenhang tussen de predictoren. Om na te gaan welke unieke rol achtergrondkenmerken spelen in het voorspellen van de recidivekans, wordt binnen de Recidivemonitor gebruikgemaakt van Cox regressie. Cox regressie is een speciale vorm van regressieanalyse waarbij in feite gewerkt wordt met twee afhankelijke variabelen: een statusvariabele die aangeeft of een bepaalde gebeurtenis heeft plaatsgevonden en een numerieke variabele die uitdrukt hoe lang het duurde voordat deze gebeurtenis plaatsvond. Cox regressie heeft veel weg van logistische regressie. Ook daar wordt op basis van de waarden van onafhankelijke variabelen een vergelijking opgesteld die het al dan niet voorkomen van – in dit geval – recidive zo goed mogelijk voorspelt. Het verschil is dat in Cox regressie bij het schatten van de rol van de verschillende covariaten rekening wordt gehouden met het tijdstip waarop de recidive plaatsvond (Cox, 1972; Klein & Moeschberger, 1997; Royston, 2001).

Tabel 3.5 geeft een voorbeeld van de resultaten van een Cox regressie. Ze hebben betrekking op de volwassen daders met een strafzaak afgedaan in 1997. De tabel geeft de omvang, het significantieniveau en het 95%-betrouwbaarheidsinterval van de *hazard ratio's* van de predictoren die in het model zijn opgenomen. De hazard ratio geeft aan hoezeer de aanwezigheid van het kenmerk van invloed is op de kans op het vóórkomen van recidive. De hazard ratio is de verhouding tussen het recidiverisico in de groep met het kenmerk en het ‘normale’ recidiverisico in een referentiegroep. Hoe meer de hazard ratio afwijkt van 1, hoe sterker het kenmerk op zichzelf beschouwd, dus los van de andere voorspellers in het model, van invloed is op de kans dat de daders op enig moment opnieuw de fout in gaan. Is de hazard ratio groter dan 1, dan is het recidiverisico hoger. Is de hazard ratio kleiner dan 1, dan is er ten opzichte van de referentiegroep en bij gelijke scores op de andere voorspellers sprake van een verlaagd risico. De maat in de laatste kolom van de tabel drukt het relatieve belang van de voorspellers uit. Hoe hoger de waarde van δ_a , des te groter de rol die de factor speelt.

Tabel 3.5 Invloed van enkele dader- en delictkenmerken op de kans op strafrechtelijke recidive van volwassen daders met een strafzaak afgedaan in 1997 (N=133.331)

Factor	hazard ratio	sign.	95%-betrouwbaarheidsinterval	δ_a
<i>Sekse</i>				
Man	1			
Vrouw	0,628	0,000	0,611 - 0,647	1,18
<i>Leeftijd bij inschrijving van de 1e strafzaak</i>	0,944	0,000	0,942 - 0,945	2,08
<i>Leeftijd bij inschrijving van de uitgangszaak</i>	1,016	0,000	1,015 - 1,018	1,22
<i>Type delict uitgangszaak</i>				
Vermogen geen geweld	1			
Vermogen met geweld	1,113	0,000	1,076 - 1,190	1,02
Agressie	0,840	0,000	0,820 - 0,860	1,07
Zeden	0,612	0,000	0,561 - 0,668	1,05
Opiumwet	0,747	0,000	0,717 - 0,779	1,07
Overig	0,700	0,000	0,681 - 0,719	1,14
Verkeer	0,812	0,000	0,795 - 0,830	1,10
<i>Aantal eerdere justitiecontacten</i>	1,018	0,000	1,018 - 1,019	1,24

δ_a is de absolute toename van de hazard ratio bij een verschuiving van de waarde van de covariaat ter grootte van één standaarddeviatie. Bij dichotome variabelen bedraagt de standaarddeviatie $\sqrt{p(1-p)}$, waarbij p de kans aangeeft dat het kenmerk voorkomt.

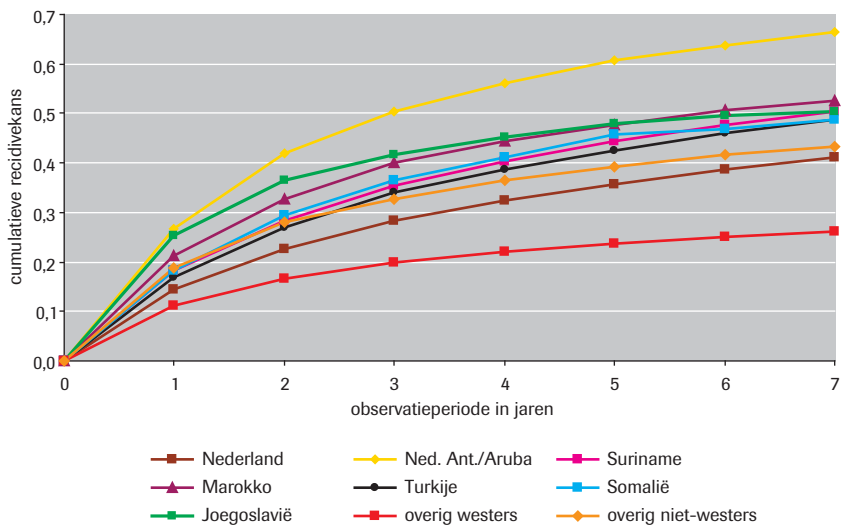
Te zien is dat ook bij constanthouding van de andere factoren in het model mannelijke daders een grotere kans hebben om te recidiveren. De kans dat een vrouwelijke dader opnieuw in de fout gaat is – bij gelijke scores op de andere kenmerken – steeds ongeveer 40% kleiner dan bij mannelijke daders. Men zou ook kunnen zeggen dat de kans dat mannen recidiveren over het algemeen ruim anderhalf keer zo groot is als die van vrouwen ($1 : 0,628 = 1,59$).

Verder blijkt dat de recidivekans kleiner wordt naarmate men ouder was toen men voor het eerst in aanraking kwam met justitie. Met elk jaar dat de strafrechtelijke carrière later begon, neemt het risico met 5,6% af. Is er een verschil van twee jaar, dan is de kans 31% kleiner ($5,6^2$). Op zichzelf hangt de leeftijd bij de uitgangszaak negatief samen met de recidivekans (zie figuur 3.5), maar gevoegd bij de rol die de startleeftijd speelt, vertoont deze factor een positieve samenhang. De kans op strafrechtelijke recidive is na een zedendelict het kleinst en na een vermogensdelict met geweld het grootst. Het aantal vroegere justitiecontacten draagt ook bij aan de voorspelling. Met elke eerdere strafzaak neemt de kans dat de strafrechtelijke carrière zich na de uitgangszaak verder voortzet, met 1,8% toe.

Afgaand op de waarden in de laatste kolom van de tabel is de startleeftijd

de belangrijkste voorspeller, gevolgd door het aantal eerdere contacten, de leeftijd bij de uitgangzaak en het geslacht van de dader.

Figuur 3.6 Baselines van cumulatieve recidivekansen voor volwassen daders met een strafzaak afgedaan in 1997 (N=133.331); naar geboorteland en bij gemiddelde waarden op de overige covariaten



Het geboorteland ontbreekt in de tabel. Het Cox model gaat uit van 'proportionaliteit', dat wil zeggen dat het effect van een factor op de hazard ratio constant is in de tijd. Dat ging echter niet op voor 'geboorteland'. Daarom is deze factor als stratumvariabele in het voorspellingsmodel opgenomen. Dit betekent dat voor elk geboorteland een eigen baseline van de recidivekansen wordt geschat. In figuur 3.6 staan deze baselines afgebeeld, bij de gemiddelde waarden op de andere factoren in het model. Volwassen daders geboren op de Nederlandse Antillen of Aruba geven over het algemeen het hoogste recidiverisico te zien. De baselines voor personen geboren buiten Nederland liggen hoger dan de baseline voor personen die binnen de landsgrenzen geboren zijn, met uitzondering van de groep daders afkomstig uit overige westerse landen. De laatste groep bestaat voor een groot deel uit toeristen en die ziet de Nederlandse justitie minder vaak terug. Uit het feit dat de baselines van sommige geboortelanden elkaar kruisen, blijkt de noodzaak van stratificatie.³⁹

³⁹ In het model zijn geen interacties tussen factoren meegenomen, omdat zij ofwel niet significant waren ofwel de voorspelkracht van het model verminderden. In het laatste geval is er sprake van 'overfitting': de p-waarden van de hazard ratio's in de schattingssteekproef zijn significant, maar het opnemen van de factoren in het model heeft een negatief effect op de kwaliteit van de voorspelling in andere steekproeven (Harrell, Lee & Mark, 1996; zie ook Hosmer & Lemeshow, 2000).

Cox regressie en logistische regressie zijn geschikte middelen om de complexe samenhang tussen achtergrondkenmerken en de prevalentie van recidive te analyseren. Met behulp van deze technieken zijn scherpere uitspraken mogelijk over de recidivekansen in de verschillende subgroepen van het onderzoek, scherper althans dan op basis van univariate analyses. Regressiemodellen houden rekening met onderlinge verbanden tussen covariaten en leggen zo de eigen bijdrage van de in het onderzoek gemeten kenmerken bloot. Binnen de Recidivemonitor worden de modellen voor verschillende doeleinden gebruikt:

- *Risicotaxatie*. Op basis van resultaten als in tabel 3.5 en figuur 3.6 is een predictie-instrument ontwikkeld, de StatRec. Met deze schaal kan per individu een grove, voorlopige inschatting worden gedaan van de kans op latere recidive (Wartna, Tollenaar & Bogaerts, 2009, in druk).
- *Benchmarking*. De uitstroomresultaten van justitiële inrichtingen kunnen beter met elkaar worden vergeleken als men rekening houdt met verschillen tussen de groepen die zij huisvestten. Zie verder hoofdstuk 6.
- *Effectstudies*. De kennis van de samenhang tussen de achtergrondkenmerken en het vóórkomen van recidive maakt het mogelijk om bij vergelijkend onderzoek betere controlegroepen samen te stellen.

De belangrijkste toepassing is echter het corrigeren van de ruwe recidivepercentages. Verschuivingen in de samenstelling van de dadergroepen die in contact komen met justitie, beïnvloeden de hoogte van de recidivecijfers en dat maakt het lastig de percentages van de opeenvolgende jaren met elkaar te vergelijken. Met Cox regressie kan de invloed van de verschuivingen worden ingeschat en geneutraliseerd. In paragraaf 3.5 komen we hier op terug.

3.4.3 Registratie-effecten

Een andere factor die het zicht op de feitelijke ontwikkeling van de recidive bemoeilijkt, is het mogelijke bestaan van registratie-effecten. Politie- en justitiecijfers vormen zoals bekend geen goede afspiegeling van de omvang van de criminaliteit. Er is sprake van een 'dark number'. Veel delicten blijven onopgemerkt, worden niet aangegeven of niet opgelost. Omdat het onderzoek van de Recidivemonitor betrekking heeft op de criminaliteit die onder de aandacht komt van het Openbaar Ministerie, zouden schommelingen in het recidiveniveau het gevolg kunnen zijn van een wijziging in het aandeel van de criminaliteit dat het OM te verwerken krijgt. Als het aantal geregistreerde delicten toeneemt terwijl de feitelijke omvang van de criminaliteit constant blijft, stijgen de recidivecijfers toch. Er zou dan sprake van een *registratie-effect*.

Om een inschatting te kunnen doen van het 'dark number' worden in Nederland slachtofferenquêtes gehouden (Eggen & Van der Heide, 2005).

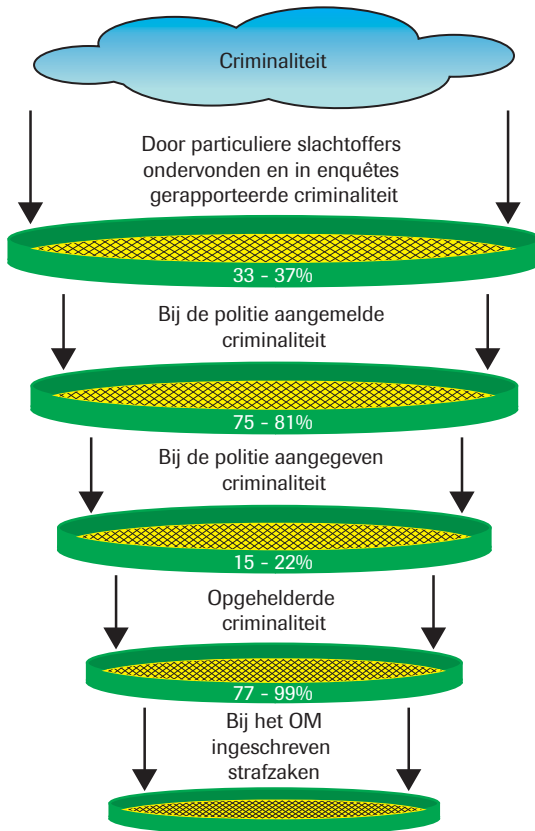
Van Kerckvoorde (1995) bespreekt de methode. Zijn conclusie is somber: ook slachtofferenquêtes geven geen compleet beeld van de omvang en verdeling van de criminaliteit. Buiten beschouwing blijven slachtofferloze delicten (zoals de meeste drugs- en verkeersdelicten), slachtofferdiffuse delicten (zoals de meeste milieudelicten) en delicten waarvan organisaties het slachtoffer zijn. Van de criminaliteit waarnaar in de enquêtes wel wordt gevraagd, is onduidelijk welk deel door de respondenten wordt gerapporteerd. Het kan zijn dat respondenten voorvallen zijn vergeten, dat men zaken bewust verzwijgt, of indeelt in verkeerde categorieën of perioden. Omgekeerd kan er sprake zijn van dubbelstellingen en andere vormen van overrapportage. Slachtoffers brengen hun ervaringen in meer dan één delictcategorie onder, dikken het verhaal aan of verzinnen voorvallen die niet hebben plaatsgevonden.

De term 'geregistreerde criminaliteit' verwijst doorgaans naar de criminaliteit zoals vastgelegd in de cijfers van politie en justitie. Maar ook slachtoffers en enquêteurs registreren criminaliteit. Of het nu gaat om een slachtoffer, een dader, een getuige, een onderzoeker of een politieagent, elke 'informatiehouder' brengt de gebeurtenissen rond het delictgedrag onder in zijn eigen referentiekader met eigen definities, conventies en betekenissen. In elke schakel van de 'kennisketen' vindt filterwerking plaats. Figuur 3.7 geeft de effecten daarvan weer. De eerste schakel van de keten is meteen de missing link. De criminaliteit laat zich niet kennen, ook niet door het afnemen van slachtofferenquêtes. Het zou mooi zijn als er sprake zou zijn van een vaste verhouding tussen de omvang van de criminaliteit en wat daarvan via slachtofferenquêtes aan het licht komt, maar erg aannemelijk is dat niet. Sterker, het zou goed kunnen dat zich vroeg in de keten grotere schommelingen in de registratie voordoen dan verderop. Onderzoekers, politieagenten en juristen hebben vanuit hun professie belang bij het hanteren van vaste definities en het voeren van een gestandaardiseerde administratie. Daders en slachtoffers niet. Hun opvattingen van crimineel gedrag en overlast zijn vermoedelijk veranderlijker dan die van politie en justitie. De motivatie om de informatie waarover zij beschikken op een consistente manier door te geven, evenzeer. Het is dus lastig om op grond van enquêteresultaten te concluderen dat de criminaliteit is toe- of afgenomen, want ook op dit laagste niveau van de kennisketen doen zich registratie-effecten voor.

Over de verhoudingen tussen de overige schakels in figuur 3.7 is iets meer bekend. Uit antwoorden van respondenten van slachtofferenquêtes blijkt dat 33 tot 37% van de delicten die via de enquêtes boven water komen, wordt aangemeld bij de politie. 75 tot 81% daarvan leidt volgens opgave van de slachtoffers zelf tot ondertekening van een proces-verbaal of een ander officieel document. De percentages zijn berekend op basis van tijdreeksen uit *Criminaliteit en Rechtshandhaving* (Eggen & Van der Heide, 2005: 329 e.v.). Ze hebben betrekking op de periode 1992-2004. Van alle processen-verbaal die in de periode 1990-2004 naar aanleiding van een

misdrijf werden opgemaakt, werd 15 tot 22% opgehelderd (ibid.: 332). Na 1985 daalde het landelijke ophelderingspercentage, maar de laatste jaren stijgt het weer. Een deel van de processen-verbaal wordt afgewikkeld door de politie zelf. Tussen 1996 en 2004 werd naar schatting 77 tot 99% van de opgehelderde misdrijven voor vervolging doorgegeven aan het OM.

Figuur 3.7 Stroomdiagram geregistreeerde criminaliteit



De lagen van de trechter sluiten niet bepaald naadloos op elkaar aan. In elke schakel van de keten doen zich verschuivingen voor in de definitie van wat wordt geregistreerd. Het domein van de delicten verandert, hun betekenis verglijdt en zo hier en daar verspringt de tel-eenheid. We mogen dus geen al te harde conclusies verbinden aan de cijfers van figuur 3.7. Niettemin is duidelijk dat slechts een beperkt deel van de – door particuliere slachtoffers ondervonden en in enquêtes gerapporteerde – delicten het einde van de keten haalt. Voor delicten die buiten de enquêtes vallen, ligt dit anders. Slachtofferloze ‘haaldelicten’, zoals rijden onder invloed en

de meeste overtredingen van de Opiumwet, kennen – als zij zijn ontdekt – ophelderingspercentages van bijna 100% (Smit et al., 2003). Voor de via slachtofferenquêtes gemeten criminaliteit geldt echter dat, zeer grof geschat, niet meer dan 10% uiteindelijk bij justitie terechtkomt. Welke factoren zijn verantwoordelijk voor de sterke filterwerking en geven zo aanleiding tot het bestaan van registratie-effecten? In eerste instantie lijken drie zaken van belang:⁴⁰

- 1 *Aangiftebereidheid*. De aangiftebereidheid geeft aan welk deel van de feitelijk gepleegde delicten bij de politie terechtkomt. Hoe meer delicten door burgers of instellingen bij de politie worden aangegeven, des te omvangrijker zijn de geregistreerde criminaliteit en de recidive die op basis daarvan wordt vastgesteld. De aangiftebereidheid kan worden bepaald aan de hand van uitkomsten van de landelijke slachtofferenquêtes van het CBS (2006). Voor elk uitgangsjaar wordt berekend welk deel van de ‘ondervonden’ criminaliteit werd opgenomen in een proces-verbaal van aangifte.⁴¹
- 2 *Ophelderingspercentage*. Wanneer er aangifte is gedaan, moet het misdrijf worden opgelost. Een misdrijf wordt doorgaans als opgehelderd beschouwd indien er bij de politie ten minste één verdachte bekend is (Smit et al., 2003). Het ophelderingspercentage kan worden berekend door het totaal aantal gehoorde verdachten te delen door het totale aantal opgemaakte processen-verbaal. Hoe meer delicten de politie ‘rond’ krijgt, des te groter is de instroom van strafzaken op de arrondissementsparketten. Met andere woorden, hoe harder de politie haar best doet, hoe meer criminaliteit er ogenschijnlijk is. Dit is een voorbeeld van een ‘prestatieparadox’ (zie o.m. Wiebrens, 1999; Wittebrood, 2002; Van Thiel & Leeuw, 2003). Omdat een belangrijk deel van de strafzaken betrekking heeft op daders die al eerder werden vervolgd, leidt betere opsporing ook tot hogere recidivecijfers.
- 3 *Politieafdoeningen*. Niet alle opgeloste delicten worden doorgegeven naar het OM. Een deel wordt afgehandeld door de politie zelf. Haltafdoeningen komen bijvoorbeeld niet bij justitie terecht en er zijn politietransacties en -sepots, zaken die wel als opgelost worden beschouwd maar die om uiteenlopende redenen niet worden doorgezet. In principe hebben politieafdoeningen een dempend effect op de cijfers van de Recidivemonitor. Als de politie meer zelf afdoet, zal het aantal strafzaken ingeschreven bij het OM dalen en wordt er minder strafrechtelijke recidive gemeten. In de praktijk is dit effect echter gering. Het aantal politieafdoeningen is de laatste jaren gestegen, het verschil

40 De wetgever bepaalt welke gedragingen van mensen en organisaties in formeel-juridische zin strafbaar zijn. Aan deze factor wordt hier verder geen aandacht besteed. Aangenomen wordt dat zich gedurende de periode van het onderzoek geen grote verschuivingen hebben voorgedaan in het scala van gedragingen die wettelijk zijn gedefinieerd als misdrijven en dus onder ons begrip van recidive vallen.

41 Het is wellicht beter te spreken van ‘meldingsbereidheid’. Niet alle meldingen van delicten leiden tot een aangifte. Zie figuur 3.7.

tussen het aantal gehoorde verdachten en het aantal ingeschreven strafzaken is in de periode 1996-2004 tot 23% toegenomen (Van Eggen & Van de Heide, 2005: 332, 370). Dit komt door de toepassing van Halt, maar vooral door de groei van het aantal 'stille' sepots, zaken die terzijde worden geschoven door de politie en dus niet aan het OM worden doorgegeven (Van Tulder, 2002). Beide ontwikkelingen drukken echter niet op de recidivecijfers. Halt-zaken draaien om delicten die voorheen ook al niet door het OM werden behandeld. De zaken die tegenwoordig op politieniveau niet-vervolgbaar worden beschouwd, werden vroeger op OM-niveau geseponneerd en bij de meting van de strafrechtelijke recidive worden technische sepots niet meegeteld.

Naast deze drie factoren is er nog een verschijnsel dat in directe zin van invloed zou kunnen zijn op de hoogte van de cijfers:

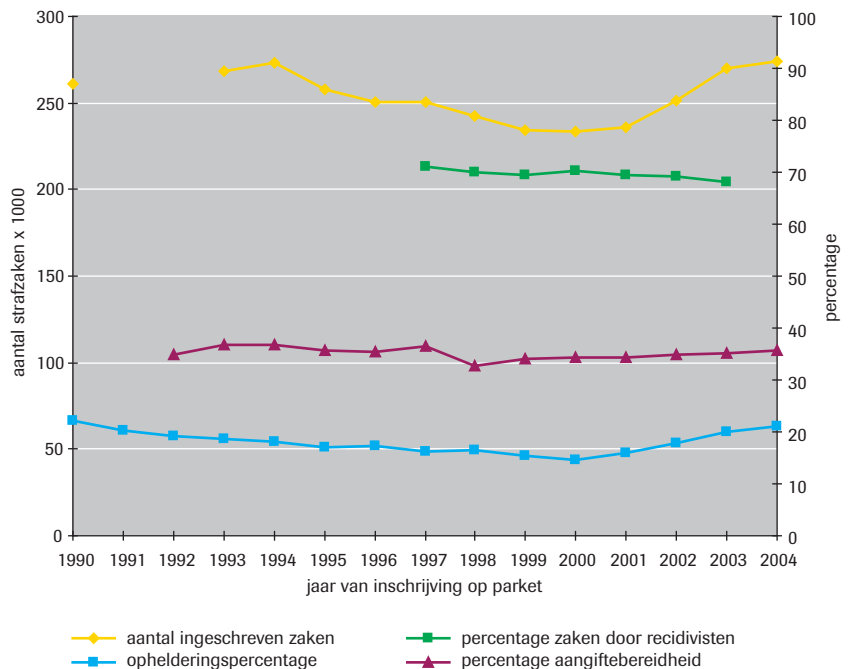
- 4 *Gerichtheid op bekende daders*. De gemeten recidive zou ook toenemen wanneer politie en justitie zich bij de opsporing en vervolging van verdachten meer dan anders zouden concentreren op de personen die al eerder werden vervolgd. Dit effect wordt ook wel het 'round up the usual suspects'-verschijnsel genoemd (Garland, 2000) en lijkt op het verschijnsel 'creaming' zoals omschreven door Lipsky (1980). Om de werkdruk het hoofd te kunnen bieden richten functionarissen zich doorgaans op de relatief eenvoudige gevallen. Het is gemakkelijker daders op te sporen als er al gegevens over hen bestaan. Bij bekende daders is de pakkans normaal gesproken groter dan bij first offenders, bijvoorbeeld omdat men over een DNA-profiel van hen beschikt, men vertrouwd is met hun modus operandi of omdat men bekende daders simpelweg meer in de gaten houdt. Worden de kansen dat een recidivist wordt ingerekend gelijkmatig verhoogd en is er dus sprake van een vaste verhouding tussen de pakkansen van recidivisten en first offenders, dan is deze factor van geen belang. Maar als de verhouding tussen die kansen verandert, bijvoorbeeld omdat de prioriteiten bij de opsporing en vervolging worden verlegd,⁴² kan dit ertoe leiden dat de recidivecijfers uit het ene jaar minder goed vergelijkbaar zijn met die uit een ander jaar.

De uitkomsten van de Recidivemonitor winnen aan waarde als duidelijk kan worden gemaakt hoezeer zij 'vervuild' zijn door registratie-effecten. Het aantal politieafdoeningen speelde zoals we zagen geen rol, maar wijzigingen in de aangiftebereidheid, het ophelderingspercentage en de gerichtheid op bekende daders zijn bronnen van mogelijke vertekening. Figuur 3.8 geeft weer welke ontwikkelingen zich op deze drie factoren hebben voorgedaan sinds de jaren '90 van de vorige eeuw. In de figuur

42 De invoering van prestatiecontracten waarin onder meer is vastgelegd hoeveel processen-verbaal jaarlijks aan het OM worden verstuurd, zou hier debet aan kunnen zijn (zie o.m. Hoogenboom, 2006). Een belangrijk deel van het Veiligheidsprogramma is gericht op de aanpak van 'veelplegers' (zie hoofdstuk 1).

wordt ook het aantal zaken weergegeven dat bij het OM werd ingeschreven. Met uitzondering van de tijdreeks van de gerichtheid op bekende daders zijn de cijfers afkomstig uit het rapport *Criminaliteit en Rechts-handhaving* (Eggen & Van der Heide, 2005). De gerichtheid op bekende daders wordt uitgedrukt in het percentage strafzaken waarbij een recidivist de verdachte was. Dit percentage is berekend op basis van gegevens uit de OBJD. Het ophelderingspercentage is de verhouding tussen het aantal delicten dat de politie als opgelost beschouwt en het aantal aangegeven delicten. De aangiftebereidheid wordt vastgesteld door het aantal aangegeven delicten te delen door het totale aantal ondervonden delicten. Deze aantallen zijn geschat op basis van slachtofferenquêtes. De cijferreeksen in figuur 3.8 hebben betrekking op heel Nederland. De invloed van deze factoren kan alleen op landelijk niveau worden vastgesteld.

Figuur 3.8 Het aantal misdrijfzaken en drie exogene determinanten van strafrechtelijke recidive



Bron: Eggen & Van der Heide (2005) / WODC-Recidivemonitor

Het aandeel strafzaken dat op rekening komt van recidivisten, is tussen 1997 en 2003 stabiel. Het percentage ligt tussen 69 en 71%. Na 2000 is misschien een lichte daling te zien, maar substantieel is deze niet. Mocht de recidive in de onderzoeksperiode zijn gestegen (waarover in

het volgende hoofdstuk meer), dan komt dit dus *niet* omdat de politie de opsporing meer is gaan richten op bekende daders, personen die al eerder in aanraking kwamen met justitie. Hetzelfde geldt in feite voor de aangiftebereidheid. Sinds de jaren '90 van de vorige eeuw schommelt de aangiftebereidheid tussen 33 en 37%. Ruim een derde van alle door particulieren ondervonden delicten wordt bij de politie aangemeld. Dit percentage is sinds 1998 nauwelijks veranderd, dus veel invloed op de geregistreerde recidive kan de aangiftebereidheid de laatste jaren niet hebben gehad. We zien wel dat het ophelderingspercentage na de eeuwwisseling vrij sterk is toegenomen. Na een daling in de jaren '90 van de vorige eeuw is het aandeel opgeloste delicten na het jaar 2000 gestegen van 15 naar 21%. Er lijkt bovendien een duidelijke relatie te bestaan tussen het percentage opgeloste delicten en het aantal strafzaken dat uiteindelijk bij het Openbaar Ministerie werd ingeschreven. De twee curven kennen grofweg hetzelfde verloop. Het ligt voor de hand te veronderstellen dat een deel van de recente stijging in het aantal ingeschreven strafzaken veroorzaakt is door de verhoging van het ophelderingspercentage. Een betere opsporing leidt dus inderdaad tot meer criminaliteit en zal dan ook tot hogere recidivecijfers hebben geleid. Maar in feite is dit – in elk geval voor een deel – een schijneffect: er wordt niet per se meer gerecidiveerd, er wordt meer recidive *geregistreerd*.

3.5 Correctie van ruwe recidivepercentages

Registratie-effecten en verschuivingen in de samenstelling van de onderzoeksgroepen vertroebelen het zicht op de ontwikkeling van de recidive. In de Recidivemonitor worden daarom naast ruwe ook gecorrigeerde percentages berekend. De gecorrigeerde cijfers laten zien hoe de recidive zich heeft ontwikkeld los van de verschuivingen in de dadergroepen en de eventuele veranderingen in aangiftebereidheid en het ophelderingspercentage. De correctie van de ruwe cijfers vindt plaats met behulp van een variant van Cox regressie. Het model schat de invloed van de gemeten kenmerken op de recidivekans en berekent of er significante registratie-effecten zijn. Op basis van de schattingen van het model worden de ruwe recidivecijfers bijgesteld.

Voor elk van de vaste daderpopulaties werd een model gefit: voor de volwassen daders met een strafzaak, de jeugdige daders met een strafzaak, de ex-gedetineerden, ex-JJI-pupillen en ex-terbeschikkinggestelden. We volgden daarbij dezelfde procedure als Schmidt en Witte (1988) bij hun analyse van de recidiveprevalentie in twee groepen ex-gevangenen. De modellen werden gefit op één helft van de populatie en gevalideerd op de andere helft. Deze procedure levert een regressievergelijking op waarmee voor elke subgroep van daders de gecorrigeerde recidive kan worden

bepaald. Bijlage 4 geeft de parameters van de vijf modellen.⁴³ In de modellen konden alleen kenmerken worden opgenomen uit de OBJD. Het gaat dan in eerste instantie om de voorspellers uit figuur 3.5: sekse, leeftijd en geboorteland van de dader, het type delict, zijn startleeftijd en het aantal eerdere justitiecontacten. Daarnaast werden jaarreeksen van ophelderingspercentages en percentages aangiftebereidheid als zaakkenmerken aan de gegevens van de respondenten toegevoegd. De aangiftebereidheid bleek in *geen* van de vijf populaties een significante bijdrage te doen aan de voorspelling van de recidive. Deze factor speelt kennelijk, op landelijk niveau althans, geen rol van betekenis. Het opwaartse effect van het ophelderingspercentage was alleen bij de volwassenen zichtbaar. Sinds 2000 is een stijging te zien in het aandeel opgehelderde delicten. Alleen in het model voor de volwassen justitiabelen was de bijdrage van deze factor significant en positief en werden de recidivecijfers dus ook op dit punt (licht) bijgesteld. In paragraaf 4.2.1 zullen we daarvan de resultaten zien. Bij het zoeken naar de juiste mix van variabelen stond de voorspelkracht voorop. Bij het fitten van de voorspellingsmodellen ging het dus niet om het verklaren van de recidive of het toetsen van hypothesen afgeleid uit een criminologische theorie, noch om het vinden van een zo spaarzaam mogelijk model. Het belangrijkste was dat de recidive die met het model werd voorspeld de werkelijke recidive in de groep zo dicht mogelijk benadert.⁴⁴ Bijlage 5 bevat een weergave van de overeenkomst tussen de voorspelde en geobserveerde recidive in de vijf onderzoekspopulaties. De fit is over het algemeen goed te noemen. Bij terbeschikkinggestelden is vanaf de middellange termijn sprake van een systematische onderschatting van de geobserveerde recidive, tot maximaal 4,6%. In het begin van de observatieperiode is het verschil echter ook in deze groep miniem. Twee jaar na het opleggen van de sancties of het vertrek uit de inrichting is de voorspelfout in geen van de vijf onderzoekspopulaties groter dan 0,9%. Dit betekent dat de modellen op dat tijdstip in de follow-up-periode een nauwkeurige schatting geven van de invloed van de factoren in het model. Er vindt met andere woorden op deze punten een juiste correctie van de ruwe cijfers plaats.

43 Het jaartal van afdoening/uitstroom fungeerde steeds als stratum. De gecorrigeerde percentages hebben een gemeenschappelijk deel, de voorspelde recidive in een referentiegroep, en een eigen deel, de baseline schatting van het betreffende jaar. In elk deelonderzoek gold de eerste cohort als referentiegroep, behalve bij de tbs. Daar vormde de laatste cohort het uitgangspunt voor de schattingen.

44 Zie onder meer Pedhazur (1982) voor het verschil tussen 'predictive' en 'explanatory' onderzoek. Via verklarend onderzoek wil men begrijpen hoe verschijnselen zijn gerelateerd. Bij voorspellend onderzoek is de aard van de relaties tussen afhankelijke en onafhankelijke variabelen van secundair belang. Alles draait om het optimaliseren van de correlatie tussen de set van predictoren en de criteriumvariabele.

3.6 Slot: meten is weten?

Omdat ze zijn ontdaan van mogelijke registratie-effecten, en de invloed van bepaalde verschuivingen binnen de dadergroepen is geneutraliseerd, bieden de gecorrigeerde cijfers beter zicht op de *netto-ontwikkeling* van de recidive. De gecorrigeerde cijfers laten zien hoe de recidive zich in de onderzoeksperiode zou hebben ontwikkeld indien de betreffende daderpopulatie op een aantal belangrijke punten niet van samenstelling was veranderd en in elk onderzoeksjaar bovendien hetzelfde ophelderingspercentage gold. De hoogte van de recidivepercentages is daarmee echter nog niet verklaard. Oorzaken en achtergronden worden met het standaardonderzoek van de Recidivemonitor niet blootgelegd, er wordt hooguit een aantal zaken uitgesloten. Als men uitspraken wil doen over oorzaken van de recidive, dan zal men de cijfers in verband moeten brengen met causale factoren zoals opgesomd in paragraaf 3.4.1. De OBJD, het bronsysteem van de WODC-Recidivemonitor, bevat slechts enkele statische kenmerken. Zij geven een goede voorspelling van de prevalentie van recidive, maar bieden er geen verklaring voor. De rol van deze moderatoren is zuiver *actuarieel*. Het zijn controlevariabelen die worden ingezet om eventuele instroomverschillen te corrigeren en de dadergroepen over de opeenvolgende jaren beter vergelijkbaar te maken.

4 Hoe hoog is de recidive in Nederland?

Uitkomsten van landelijk onderzoek

Nu de werkwijze van de Recidivemonitor is toegelicht, kunnen we ingaan op de uitkomsten van de metingen. Na een korte inleiding volgt in paragraaf 4.2 een beschrijving van de ontwikkelingen die zich in de recidive van de vaste daderpopulaties van de monitor hebben voorgedaan. In paragraaf 4.3 wordt een vergelijking gemaakt met cijfers uit het buitenland. Aan het slot van het hoofdstuk wordt een poging gedaan een verklaring te vinden voor de gesignaleerde trends.

4.1 De derde ronde van het landelijk onderzoek

De start van de periodieke metingen van de Recidivemonitor vond plaats in 2005, maar de resultaten die hier worden besproken zijn afkomstig van de derde ronde van het landelijk onderzoek. In dat jaar werd de selectieperiode van de vijf deelonderzoeken verlengd tot 2004. Dit betekent dat de metingen betrekking hebben op personen die tot 1 januari 2005 te maken kregen met een justitiële interventie. Tabel 4.1 geeft een overzicht van de onderzoeksgroepen. Het gaat om zo goed als alle personen die tussen 1997 en 2004 één of meer keren als verdachte van een misdrijf te maken kregen met de Nederlandse justitie, plus alle personen die in die jaren een justitiële inrichting verlieten. Voor de ex-terbeschikkinggestelden, die vanwege de geringere aantallen zijn verdeeld in cohorten van vijf jaar, zijn cijfers beschikbaar over de periode 1974-2003.

De eerste twee deelonderzoeken hebben betrekking op jeugdigen en volwassenen die een straf of maatregel kregen opgelegd voor het plegen van een misdrijf. De sancties zijn verdeeld in zes hoofdgroepen: onvoorwaardelijke vrijheidsstraffen, voorwaardelijke vrijheidsstraffen, werk- en leerstraffen, geldstraffen, beleidssepots en een categorie 'overige straffen en maatregelen'. Doel van het onderzoek is het beschrijven van de ontwikkeling in de strafrechtelijke recidive die volgt op het opleggen van deze typen sancties. De drie laatste groepen omvatten de cliënten van de Dienst Justitiële Inrichtingen (DJI). Deze organisatie is verantwoordelijk voor de tenuitvoerlegging van intramurale sancties, de voorlopige hechtenis en de opvang van vreemdelingen en onder toezicht gestelde jongeren die uit huis werden geplaatst. Een deel van de personen uit de eerste twee onderzoeksgroepen werd veroordeeld tot een onvoorwaardelijke vrijheidsstraf, kreeg tbs opgelegd of zat in voorlopige hechtenis. Zij komen dus ook voor in een van de DJI-groepen, er is sprake van een zekere overlap. De meeste daders kregen evenwel een extramurale sanctie opgelegd, zoals een boete, een taakstraf of een beleidssepot. In totaal waren er bijna 865.000 unieke personen bij het onderzoek betrokken.

Tabel 4.1 De vaste onderzoeksgroepen van de WODC-Recidivemonitor – meting 2007

Groep	Omschrijving	Selectie- periode	Lengte cohort	Observatie- periode	Aantal observaties
Volwassen daders	Meerderjarigen met een strafzaak wegens een misdrijf, afgedaan met een transactie, een beleidssepot of een straf of maatregel door de rechter	1997-2004	1 jaar	2-9 jaar	1.177.887
Jeugdige daders	Minderjarigen met een strafzaak wegens een misdrijf, afgedaan met een transactie, een beleidssepot of een straf of maatregel door de rechter	1997-2004	1 jaar	2-9 jaar	132.930
Ex-gedetineerden	Personen die werden ontslagen uit een penitentiaire inrichting voor volwassenen	1997-2004	1 jaar	2-9 jaar	158.588
Ex-JJI-pupillen	Jongeren uitgestroomd uit een justitiële opvang- of behandelinrichting	1997-2004	1 jaar	2-9 jaar	23.325
Ex-terbeschikkinggestelden	Personen van wie de (deels) intramurale uitgevoerde tbs-maatregel formeel werd beëindigd	1974-2003	5 jaar	2-31 jaar	2.203

Hoewel er meer gegevens beschikbaar zijn, zal in dit hoofdstuk vooral aandacht worden besteed aan de *prevalentie* van algemene recidive, binnen twee jaar na oplegging van de straf of het ontslag uit de inrichting. De korte observatieperiode maakt het mogelijk ook het uitgangsjaar 2004 bij de analyses te betrekken, zodat hier een betrekkelijk actueel beeld kan worden geschetst van de strafrechtelijke recidive in Nederland. Centrale vraag is steeds of het percentage recidivisten gedurende de onderzoeksperiode is toe- of afgenomen. In de slotparagraaf zal ook aandacht worden besteed aan de *omvang* van de recidive in de diverse sectoren. Bij de meting onder ex-terbeschikkinggestelden richten we ons op de ernstige en de speciale recidive.⁴⁵ Tevens zal worden ingegaan op de recidive op langere termijn. In de tbs-sector zijn gegevens voorhanden over een periode van meer dan dertig jaar.

Om beter zicht te krijgen op de netto-ontwikkeling van de strafrechtelijke recidive zullen naast de ruwe ook *gecorrigeerde* percentages worden gepresenteerd. In het vorige hoofdstuk is uiteengezet hoe de bijstelling van de cijfers in zijn werk gaat. Voor elke populatie werd op basis van de beschikbare gegevens een model ontwikkeld waarmee de 'instroomverschillen' tussen de cohorten kunnen worden geneutraliseerd. De gecorrigeerde percentages geven aan hoe hoog de strafrechtelijke recidive in een bepaald jaar zou zijn geweest indien er zich in de onderzoekspopulaties geen verschuivingen hadden voorgedaan op de kenmerken sekse, leeftijd, geboorteland, type delict, startleeftijd en de omvang van het strafrecht-

45 Zie paragraaf 3.3.2 voor de definitie van deze criteria.

lijke verleden. Bovendien wordt rekening gehouden met registratie-effecten.

4.2 Recidive in Nederland

We beginnen met de resultaten van het onderzoek onder volwassenen met een strafzaak. In paragraaf 4.2.2 volgt het recidivebeeld van minderjarige daders. Daarna komen de uitstroomresultaten van de drie sectoren van de justitiële inrichtingen aan de orde: het gevangeniswezen, de justitiële jeugdinrichtingen en de tbs-sector.

4.2.1 Volwassen daders met een strafzaak

Tabel 4.2 geeft de verdeling van de sancties te zien die werden opgelegd aan de volwassen daders die in de periode 1997-2004 voor het plegen van een misdrijf in aanraking kwamen met justitie. Het gaat niet om unieke personen. Een persoon kan in meer dan één cohort voorkomen, maar slechts één keer in één jaar.⁴⁶ De laatste drie jaar van de onderzoeksperiode is het aantal daders sterk toegenomen. In 2004 bedroeg het aantal volwassenen met minstens één strafzaak bijna 170.000, in 1997 lag dit aantal nog op 140.000; een stijging van 22%. Het totale aantal strafzaken dat bij het OM werd ingeschreven laat een soortgelijke toename zien. Eggen en Van der Heide (2005: 370) wijzen op de sterke uitbreiding van de capaciteit van de rechterlijke macht. In Nederland komen steeds meer mensen in aanraking met het strafrecht.

Tabel 4.2 Volwassen daders met een strafzaak; naar jaar van afdoening en type straf (in kolompercentages)

	1997	1998	1999	2000	2001	2002	2003	2004
Straf	n=139.476	n=138.207	n=139.590	n=136.615	n=138.742	n=147.273	n=168.253	n=169.731
Onvw. vrij ≤ 6 mnd	6,8	6,7	7,5	7,6	8,2	8,4	8,1	7,1
Onvw. vrij > 6 mnd	4,6	4,5	4,2	4,4	5,0	5,6	5,4	5,1
Werkstraf	9,6	10,1	11,3	11,1	12,1	14,1	15,1	16,3
Leerstraf	0,1	0,1	0,1	0,1	0,2	0,5	0,6	0,8
Vw. vrij	8,1	7,8	8,0	7,3	5,9	4,4	3,8	3,2
Geldstraf	62,1	62,1	61,3	61,1	61,8	59,9	60,6	60,9
Beleidssepot	8,7	8,6	7,4	8,2	6,6	6,7	6,2	6,4
Overig	0,2	0,2	0,2	0,2	0,2	0,3	0,2	0,3

⁴⁶ Dit om te vermijden dat er statistische afhankelijkheid ontstaat tussen de observaties binnen één cohort. Bij trendstudies dienen de cohorten dezelfde lengte te hebben. Hier is dat dus één jaar.

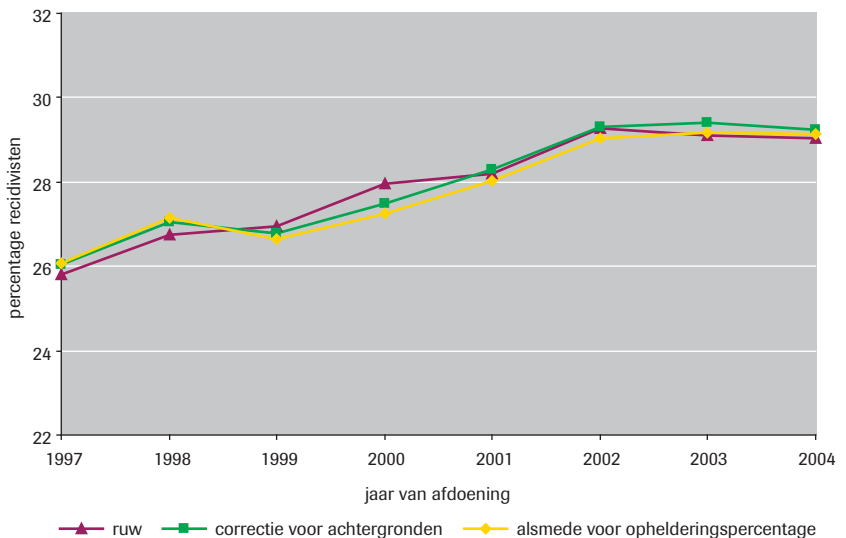
De geldstraf werd in de onderzoeksperiode verreweg het meest gebruikt. Rond 60% van de volwassen daders kreeg met deze sanctie te maken. Naast boetes worden onder geldstraffen overigens ook de transacties gerekend die bestaan uit het betalen van een geldsom. Bij onvoorwaardelijke gevangenisstraffen wordt bij de volwassenen onderscheid gemaakt tussen korte en lange vrijheidsstraffen. Korte vrijheidsstraffen duren niet langer dan zes maanden. In de loop van de onderzoeksperiode is het totale aandeel onvoorwaardelijke vrijheidsstraffen gestegen, maar in de laatste twee jaar weer iets gedaald. Het aandeel werkstraffen nam vanaf 2000 toe. Het beleidssepot en de geheel voorwaardelijke vrijheidsstraf kwamen de laatste jaren van de onderzoeksperiode minder vaak voor. Leerstraffen worden bij volwassenen niet vaak toegepast. Zij zullen in het vervolg van het onderzoek samen met de werkstraffen worden ondergebracht in de categorie 'taakstraffen'. De categorie 'overige straffen en maatregelen' omvat de overige transactievormen en de vrijheidsbenemende maatregelen, zoals de tbs en de plaatsing in een inrichting voor jeugdigen (PIJ). Deze sancties worden hier verder buiten beschouwing gelaten. De recidive na de tbs komt in paragraaf 4.2.5 aan de orde, aan de uitstroomresultaten van de PIJ-maatregel wordt in paragraaf 4.2.4 aandacht besteed. In bijlage 6 is een overzicht opgenomen van de achtergrondkenmerken van de volwassen daders. De verschuivingen op de 'OBJD-kenmerken' zijn niet erg groot. De grootste veranderingen deden zich voor in de verdeling van de leeftijd en het type delict. Het aandeel jongvolwassenen (met een leeftijd van 18 tot 25 jaar) is tussen 1997 en 2004 gestegen van 23 naar 26%. Het aandeel geweldsdelicten nam toe van 8 naar 13%. De verschuivingen op de achtergrondkenmerken zijn van belang omdat ze van invloed zijn op het recidiveniveau. Als er meer veelplegers worden vervolgd, gaat het recidivepercentage omhoog. Als er meer vrouwen worden bestraft, dalen de recidivecijfers.

Ruwe en gecorrigeerde recidivepercentages

Figuur 4.1 laat zien welke gevolgen de correcties voor verschuivingen in de populatie en voor registratie-effecten te zien geven. De ruwe recidive (de paarse lijn) stijgt tussen 1997 en 2004 van 26 naar 29%. Aan het eind van de onderzoeksperiode hadden dus bijna drie op de tien volwassen daders binnen twee jaar een nieuwe strafzaak. De correctie voor verschuivingen op achtergrondkenmerken (de groene lijn) doet daar weinig aan af. Kennelijk waren de instroomverschillen tussen de cohorten niet erg groot. Per saldo worden de recidivepercentages slechts licht bijgesteld, in sommige jaren iets naar boven, in andere iets naar beneden. Het effect van de correctie voor registratie-effecten (de gele lijn) is eenduidiger. Bij de volwassenen bleek alleen het ophelderingspercentage een significante bijdrage te doen aan de voorspelling van de recidive; de ontwikkelingen in de aangiftebereidheid veranderden de cijfers niet. Het percentage opgehelderde delicten steeg na de eeuwwisseling van 15 naar 21% (zie figuur 3.8).

In figuur 4.1 zien we dan ook dat in de tweede stap van de correctie de recidivepercentages aan het eind van de onderzoeksperiode naar beneden worden bijgesteld. Een klein deel van de toename van het recidivepercentage tegen het eind van de observatieperiode is dus een registratie-effect en komt voor rekening van de stijging van het ophelderingspercentage.⁴⁷

Figuur 4.1 Ruwe en gecorrigeerde percentages 2-jarige algemene recidive bij volwassen daders met een strafzaak; naar jaar van afdoening



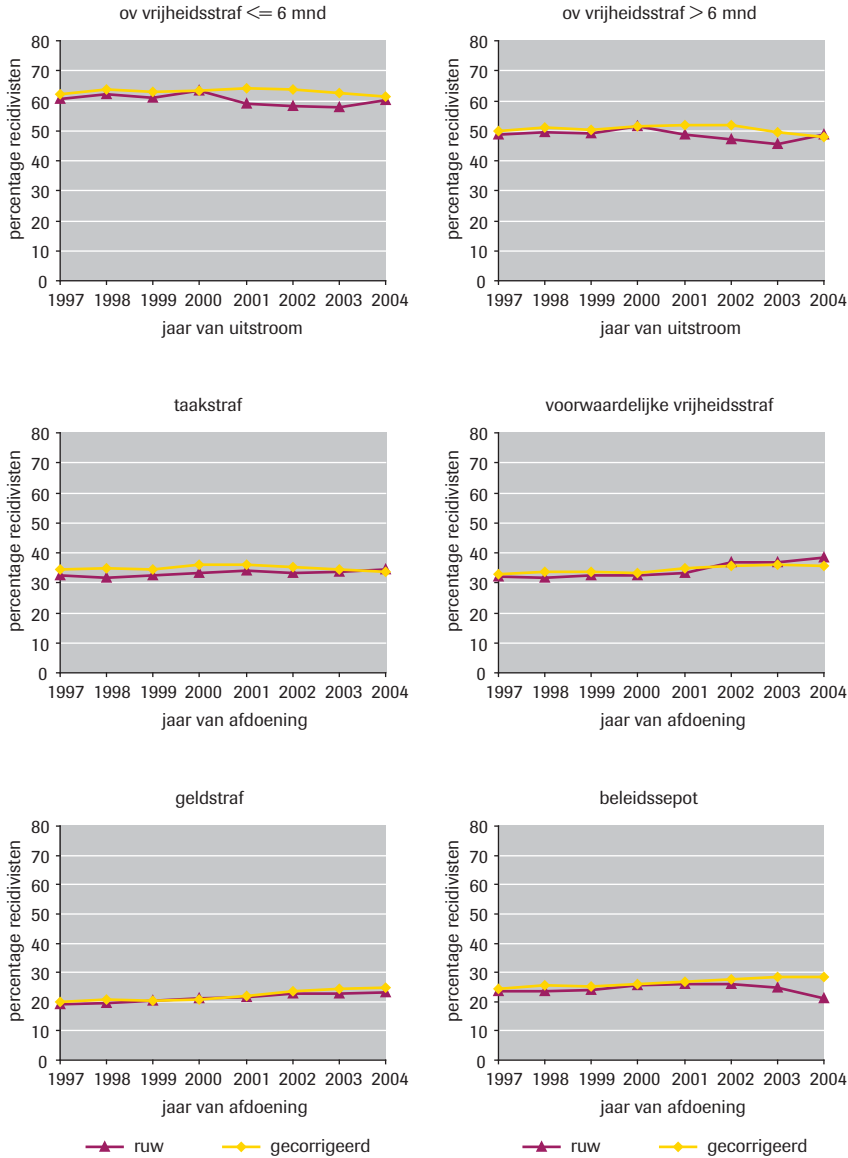
Recidivebeeld per sanctie

Figuur 4.2 geeft de ruwe en gecorrigeerde recidivepercentages uitgesplitst naar sanctie. De terugval na een gevangenisstraf is duidelijk het hoogst. Maar dat mocht ook worden verwacht, want deze straf wordt opgelegd aan personen met een hoger recidiverisico. Volwassen daders die een gevangenisstraf krijgen, hebben andere achtergronden dan de personen die een boete of een beleidssepot krijgen. Ze hebben andere delicten gepleegd en hebben veel vaker een strafrechtelijk verleden.

De correcties van de ruwe percentages – beide correcties worden nu in één keer getoond – zijn over het algemeen niet groot. Dit wijst erop dat de sanctiegroepen door de jaren heen niet veel zijn veranderd, althans niet wat betreft het samengestelde recidiverisico berekend op basis van de in

⁴⁷ Bij de correctie van de ruwe percentages gold 1997 als 'referentiejaar'. In dat jaar zouden de gecorrigeerde percentages samen moeten vallen met het ruwe percentage. Dat doen ze niet, er is een klein verschil van maximaal 0,3%. Dit is de voorspelfout.

Figuur 4.2 Ruwe en gecorrigeerde percentages 2-jarige algemene recidive in acht opeenvolgende populaties van volwassen daders; naar soort afdoening



dit onderzoek gemeten achtergronden. Bij de onvoorwaardelijke vrijheidsstraffen en de beleidssepts zijn de verschillen nog het grootst. Na 2000 worden de ruwe percentages met enkele procenten naar boven bijgesteld. Dit betekent dat deze sancties in die jaren steeds vaker werden opgelegd aan personen met een lager recidiverisico. De ruwe recidivepercentages geven een daling te zien, maar die kan zo goed als geheel worden toegeschreven aan veranderingen in de samenstelling van de groep die met een gevangenisstraf of een beleidssepot te maken kreeg. Bij de beleidssepts is na correctie zelfs sprake van een lichte toename van het recidivepercentage, net zoals bij de veelvoorkomende geldstraf en de voorwaardelijke vrijheidsstraf.

4.2.2 Jeugdige daders met een strafzaak

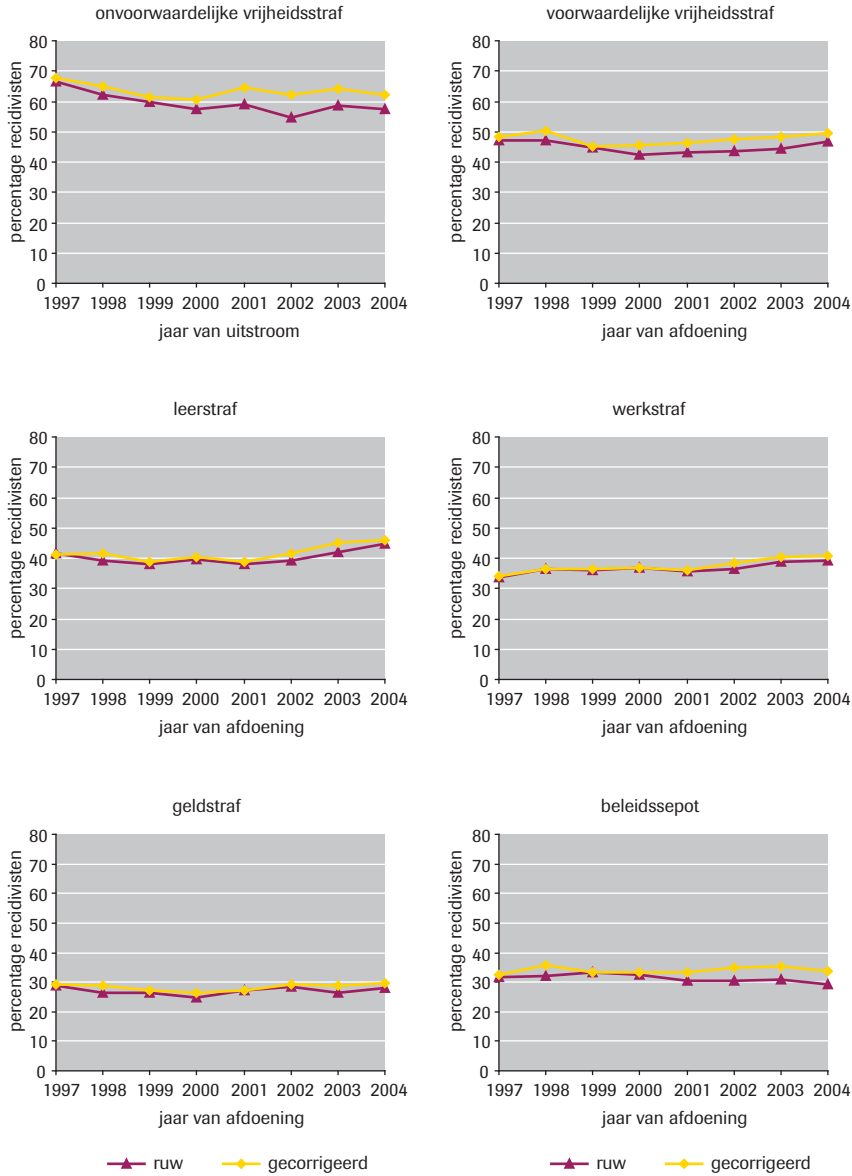
Tabel 4.3 geeft een overzicht van het recidiveonderzoek onder jeugdige daders. Het aantal minderjarigen met een strafzaak is in de periode 1997-2004 spectaculair gestegen, van 14.000 naar ruim 21.000; een toename van meer dan 50%. Onvoorwaardelijke vrijheidsstraffen komen onder jeugdigen niet vaak voor, maar het aandeel is in de loop van de onderzoeksperiode wel groter geworden. Ook de werkstraf en de leerstraf werden steeds vaker toegepast. In 2004 vormden zij samen ruim 60% van alle opgelegde straffen.⁴⁸ De geheel voorwaardelijke vrijheidsstraf, de geldstraf en het beleidssepot werden tussen 1997 en 2003 steeds minder vaak gebruikt. Het aandeel van de geldstraf en het beleidssepot is in acht jaar tijd gehalveerd.

Tabel 4.3 Jeugdige daders met een strafzaak; naar jaar van afdoening en type straf (in kolompercentages)

	1997 n=13.942	1998 n=14.007	1999 n=14.658	2000 n=15.113	2001 n=15.926	2002 n=18.402	2003 n=19.610	2004 n=21.272
Straf								
Onvw. vrij ≤ 6 mnd	3,3	2,6	3,3	3,3	4,9	5,6	5,7	5,2
Onvw. vrij > 6 mnd	2,2	2,4	2,8	2,8	3,2	3,2	2,3	2,3
Werkstraf	42,6	42,5	44,0	44,9	46,4	47,0	51,7	57,2
Leerstraf	9,1	11,9	13,7	15,2	15,7	19,0	14,8	13,5
Vw. vrij	6,5	6,5	6,0	5,8	4,5	3,9	4,2	3,8
Geldstraf	17,8	16,8	13,8	12,2	10,9	9,7	10,9	8,8
Beleidssepot	17,5	16,1	15,1	14,6	13,4	10,6	9,5	8,5
Overig	1,0	1,3	1,3	1,2	1,0	0,9	0,8	0,9

48 De vervolgstaffen die binnen hetzelfde jaar aan recidivisten werden opgelegd, zijn hierbij niet meegeteld. Voor het totaal zie Eggen & Van der Heide, 2005: 441.

Figuur 4.3 Ruwe en gecorrigeerde percentages 2-jarige algemene recidive in acht opeenvolgende populaties van jeugdige daders; naar soort afdoening



Bijlage 6 toont de achtergrondkenmerken van de jeugdigen in de opeenvolgende cohorten. Meest in het oog springt de sterke toename van het aantal meisjes. Het percentage meisjes stijgt van 12% in 1997 naar 17% in 2004. Verder valt op dat het aandeel gewelds- en openbare ordedelicten is toegenomen. In 2004 waren deze delicttypen samen goed voor 45% van de misdrijfzaken van jeugdigen, in 1997 was dit nog 36%. Voor iets minder dan drie kwart van alle minderjarige daders was de uitgangzaak het eerste justitiecontact. Daarin is gedurende de onderzoeksperiode weinig veranderd. Het merendeel van de jongeren was first offender; de uitgangzaak vormde voor hen het eerste justitiecontact.

Recidivebeeld per sanctie

Figuur 4.3 geeft de ruwe en gecorrigeerde recidivepercentages uitgesplitst naar sanctie. De cijfers zijn bijgesteld voor de verschuivingen op de achtergrondkenmerken. Er was geen sprake van een significant en positief effect voor de stijging in het landelijke ophelderingspercentage. Vanwege de lage aantallen zijn de onvoorwaardelijke vrijheidsstraffen in één categorie samengevoegd.

De verschillen tussen de ruwe en gecorrigeerde percentages zijn het grootst bij de onvoorwaardelijke vrijheidsstraf, de voorwaardelijke vrijheidsstraf en de beleidssepots, maar voor alle sancties geldt dat de ruwe recidivepercentages naar boven worden bijgesteld. We mogen dus concluderen dat de populatie van jeugdige daders in het algemeen 'lichter' is geworden. Toch is de recidive in de loop van de onderzoeksperiode na de meeste sancties met enkele procenten toegenomen. Alleen bij de zwaarste sanctie, de onvoorwaardelijke vrijheidsstraf, is er na correctie voor instroomverschillen een daling te zien, van ruim 5%.

4.2.3 Ex-gedetineerden

Een deel van de volwassenen over wie in paragraaf 4.2.1 werd gerapporteerd, verbleef in een gevangenis of een huis van bewaring. Zij kregen een onvoorwaardelijke vrijheidsstraf opgelegd of zaten vast in voorlopige hechtenis. Omdat van deze personen de exacte uitstroomdatum bekend is, kan de terugval van deze 'ex-gedetineerden' nauwkeuriger worden bepaald. De Dienst Justitiële Inrichtingen (DJI) levert het WODC jaarlijks de persoonsgegevens aan die nodig zijn om de justitiële documentatie van deze groep uit de OBJD te kunnen extraheren.

Tabel 4.4 geeft een overzicht van de onderzoeksantallen. Nederland gold jarenlang als een van de landen met het laagste aantal gedetineerden per hoofd van de bevolking. Sinds de jaren '80 van de vorige eeuw is de capaciteit van het gevangeniswezen echter enorm gegroeid. Tot 1985 kwam het aantal plaatsen in huizen van bewaring en gevangenissen niet boven de 5.000 uit. In 2004 was het aantal cellen toegenomen tot meer dan 18.000.

De totale (meerderjarige) bevolking nam in dezelfde periode met een factor 1,4 toe.

In de tabel zien we de gevolgen van de capaciteitsuitbreidingen terug. Het aantal uitgestroomde gedetineerden is in de onderzoeksperiode met 42% gestegen. Alleen in 1999 en 2002 vielen de uitstroomaantallen lager uit dan in het jaar daarvoor.⁴⁹ Overigens zijn niet alle ex-gedetineerden opgenomen in het onderzoek. Degenen die uitstroonden uit vreemdelingenbewaring, strafrechtelijke vreemdelingen die aansluitend werden uitgezet en personen die de inrichting vanuit voorlopige hechtenis verlieten maar nog niet waren afgestraft, werden uitgesloten. In totaal zijn ruim 101.000 verschillende personen bij het onderzoek betrokken.⁵⁰

Tabel 4.4 geeft ook de gemeten achtergrondkenmerken. Het aandeel vrouwelijke ex-gedetineerden is in de periode 1997-2004 meer dan verdubbeld. In 2003 kwam het percentage vrouwen voor het eerst uit boven 10%. De stijging heeft te maken met de ingebruikname van nieuwe detentiecentra bestemd voor drugkoeriers. Deze centra huisvesten relatief veel vrouwelijke gedetineerden. Net als bij de volwassen en jeugdige daders in het algemeen is de proportie gedetineerden die vast kwamen te zitten voor een geweldsdelict, toegenomen. Ook het aandeel overtreders van de Opiumwet steeg fors, vooral in 2003. Van de ex-gedetineerden uit dat jaar werd één op de vijf veroordeeld voor een Opiumwetdelict. Ook dit heeft te maken met de opening van detentiecentra voor 'bolletjesslikkers'. Een andere opvallende verschuiving heeft betrekking op het strafrechtelijke verleden van de ex-gedetineerden. Het percentage first offenders, personen die nooit eerder strafrechtelijk werden vervolgd, is tot 2003 substantieel toegenomen en valt in het laatste jaar van de onderzoeksperiode weer terug. Het aandeel 'veelplegers', personen die voorafgaand aan de detentie frequent met justitie in aanraking kwamen, is nog steeds erg groot, maar nam in de laatste jaren af. De gemiddelde leeftijd van de ex-gedetineerden nam intussen toe. De groep 40-plussers is groter geworden en dat ging ten koste van het aandeel 25- tot 30-jarigen.

49 In 1999 ging de Dienst Justitiële Inrichtingen (DJI) over op een nieuw beleidsinformatiesysteem en werden nog niet alle verblijven in dat systeem geregistreerd. Het relatief lage aantal in 2002 heeft te maken met het tekort aan celruimte. Om 'heenzendingen' tegen te gaan werd de executie van lopende vonnissen en van subsidiaire hechtenis steeds vaker uitgesteld. Het aantal korte verblijven nam hierdoor sterk af, met als gevolg dat er minder mensen uitstroonden.

50 Een deel van de ex-gedetineerden stroomde in meer dan één jaar uit. Een persoon kan per jaar slechts één keer voorkomen. Voor iemand die binnen een jaar meermalen uitstroonde, telt het eerste verblijf als de uitgangzaak. Een volgend verblijf wordt, indien er sprake is van een nieuwe strafzaak, als een geval van recidive aangemerkt.

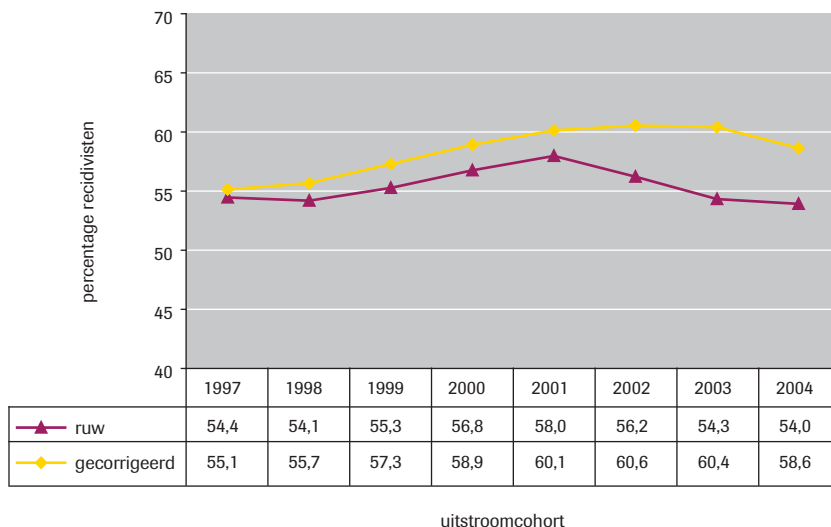
Tabel 4.4 Onderzoeksaantallen en achtergronden van ex-gedetineerden; naar jaar van uitstroom (in kolompercentages)

	1997 n=18.281	1998 n=18.631	1999 n=17.977	2000 n=19.862	2001 n=19.991	2002 n=17.015	2003 n=20.887	2004 n=25.936
<i>Sekse</i>								
Man	95,9	95,2	92,4	92,0	94,7	93,3	88,6	90,2
Vrouw	4,1	4,8	7,6	8,0	5,3	6,7	11,4	9,8
<i>Leeftijd bij uitstroom</i>								
Tot 20 jaar	3,6	4,0	4,5	5,2	4,4	4,6	4,6	4,0
20-24 jaar	18,3	17,6	17,1	16,9	17,5	18,9	18,3	17,9
25-29 jaar	23,0	21,6	20,9	19,5	18,2	18,7	17,7	16,5
30-39 jaar	34,1	34,6	34,5	34,9	35,8	34,5	34,2	33,5
40-49 jaar	14,6	15,2	16,5	17,0	17,3	16,6	18,0	19,6
50+	5,2	5,6	5,2	5,3	5,6	5,2	5,9	7,2
Onbekend	1,2	1,3	1,3	1,1	1,3	1,4	1,2	1,3
<i>Geboorteland</i>								
Nederland	54,9	53,9	52,9	52,6	49,8	47,6	47,5	51,7
Marokko	8,4	8,5	8,3	7,8	7,9	8,0	7,3	7,2
Nederlandse Antillen	5,9	6,3	6,3	7,1	7,8	8,7	11,1	9,4
Suriname	10,8	10,3	10,1	9,4	9,4	8,7	8,4	8,7
Turkije	3,6	3,8	3,8	3,3	3,2	3,4	3,2	3,2
Overige westers	7,9	8,4	8,4	9,3	10,0	10,5	9,8	8,7
Overige niet-westers	8,0	8,5	9,8	10,1	11,5	12,5	12,1	10,6
<i>Type delict</i>								
Openbare orde*	6,0	5,9	6,8	6,7	6,7	6,3	5,3	6,4
Vermogen zonder geweld	36,6	36,6	37,5	36,8	37,6	36,9	35,3	35,4
Vermogen met geweld	9,1	8,6	7,9	8,1	10,2	12,0	11,5	10,0
Geweld	8,3	8,1	8,5	9,4	10,3	11,6	11,5	11,7
Zeden	2,6	2,3	2,1	2,1	2,1	2,6	2,3	2,2
Drugs	12,4	12,6	11,6	11,1	13,2	17,0	20,6	16,1
Verkeer	7,2	7,4	7,1	6,4	5,4	3,4	3,8	5,9
Overig misdrijf	2,9	2,7	2,6	3,0	2,6	2,7	2,8	2,9
Kantonfeit (overtreding)	0,2	0,2	0,3	0,3	0,2	0,0	0,0	0,1
<i>Strafrechtelijk verleden</i>								
0 eerdere contacten	12,5	12,8	12,7	13,4	15,0	18,8	20,0	14,2
1-2 eerdere contacten	14,0	14,1	14,5	14,2	14,0	14,0	14,7	17,1
3-4 eerdere contacten	10,6	10,9	10,7	10,7	10,7	9,8	10,3	11,8
5-10 eerdere contacten	21,8	21,5	21,6	21,3	20,5	19,6	19,3	20,9
11-19 eerdere contacten	17,7	17,5	16,8	16,6	16,3	15,1	14,6	15,6
20 of meer eerdere contacten	23,5	23,2	23,7	23,8	23,4	22,7	21,1	20,4

Tabel 4.4 Onderzoeksaantallen en achtergronden van ex-gedetineerden; naar jaar van uitstroom (in kolompercentages) (vervolg)

	1997 n=18.281	1998 n=18.631	1999 n=17.977	2000 n=19.862	2001 n=19.991	2002 n=17.015	2003 n=20.887	2004 n=25.936
<i>Leeftijd 1e strafzaak</i>								
12-17 jaar	39,1	38,8	37,7	38,1	37,8	37,6	35,8	37,5
18-24 jaar	32,5	32,1	32,8	32,4	31,5	31,5	31,8	32,0
25-29 jaar	11,8	11,7	12,4	12,4	12,4	12,3	12,4	11,5
30-39 jaar	10,6	11,0	11,1	11,3	12,0	12,0	12,7	11,9
40-49 jaar	3,1	3,4	3,3	3,2	3,5	3,6	4,3	4,0
50 jaar of ouder	1,1	1,3	1,1	1,1	1,2	1,2	1,4	1,4
<i>Verblijfsduur</i>								
Tot 1 maand	28,9	32,3	35,4	35,5	23,5	18,6	18,8	24,3
1-3 maanden	23,9	22,6	23,8	23,9	20,0	23,7	24,5	24,9
3 tot 6 maanden	15,8	15,2	15,0	14,5	14,7	20,2	20,6	19,5
6 maanden tot 1 jaar	18,7	18,4	15,8	15,8	14,8	21,3	21,4	18,7
1 jaar of meer	11,9	10,6	9,0	9,3	9,7	12,2	12,0	10,9

* Vernieling, lichte agressie en openbare ordedelicten.

Figuur 4.4 Ruwe en gecorrigeerde percentages 2-jarige algemene recidive in acht opeenvolgende cohorten van ex-gedetineerden

Ruwe en gecorrigeerde percentages

Figuur 4.4 brengt de gevolgen van deze verschuivingen in beeld. Het ruwe recidivepercentage schommelt tussen 54 en 58%. Meer dan de helft van de ex-gedetineerden heeft binnen twee jaar een nieuw justitiecontact wegens het plegen van een misdrijf. Na 2001 daalt het percentage, omdat er meer vrouwen, first offenders en ouderen zijn uitgestroomd; subgroepen die over het algemeen minder hoge recidivepercentages laten zien. Na correctie voor de verschuivingen op de achtergrondkenmerken is de afname van de recidive zo goed als verdwenen. Na 2002 daalt het niveau nog iets, maar gerekend over de gehele onderzoeksperiode is er onder ex-gedetineerden sprake van een stijging van de strafrechtelijke recidive. De toename bedraagt 3,5%.

4.2.4 Jongeren uit justitiële jeugdinrichtingen

Nederland kent twee typen justitiële jeugdinrichtingen (JJI's): opvanginrichtingen en behandelinrichtingen. De opvanginrichtingen zijn bedoeld voor de uitvoering van de voorlopige hechtenis en de jeugddetentie.⁵¹ De behandelinrichtingen zijn bestemd voor de verzorging, opvoeding en behandeling van jongeren met ernstige gedragsproblemen. De meeste van hen worden op civielrechtelijke titel opgenomen, in het kader van een ondertoezichtstelling (ots).⁵² In de behandelinrichtingen worden echter ook strafrechtelijke maatregelen ten uitvoer gelegd. Tegenwoordig gaat het dan om een 'plaatsing in een inrichting voor jeugdigen' (PIJ). Tabel 4.6 geeft de aantallen en achtergronden van de ex-pupillen.⁵³ Ook de JJI-sector kende een explosieve groei. Het aantal plaatsen in de opvang- en behandelinrichtingen is tussen 1997 en 2004 bijna verdubbeld (Eggen & Van der Heide, 2007: 509-510). De onderzoeksantallen reflecteren de groei van de capaciteit, hoewel de groep uit 1997 ook om een technische reden kleiner is. In dat jaar werden nog niet alle verblijven vastgelegd in TULP-Jeugd, het beleidsinformatiesysteem dat de uitstroomgegevens voor dit onderzoek opleverde.

51 Daarnaast bieden de opvanginrichtingen onderdak aan jeugdige vreemdelingen die wachten op uitzetting, zijn er wachtplaatsen voor jongeren die naar een behandelinrichting gaan en vindt er crisisopvang plaats in het kader van een civielrechtelijke ondertoezichtstelling.

52 Gemakshalve zal ook bij civielrechtelijk geplaatste jongeren worden gesproken van 'recidive' als er zich na het verblijf in de inrichting een strafzaak voordoet. In feite is daar niet noodzakelijk sprake van, omdat zij niet (primair) naar aanleiding van een misdrijf werden geplaatst.

53 Het totale aantal unieke personen bedraagt 18.929.

Tabel 4.5 Onderzoeksaantallen en achtergronden van pupillen van justitiële jeugdinrichtingen; naar jaar van uitstroom (in kolompercentages)

	1997 n=1.476	1998 n=2.006	1999 n=2.122	2000 n=2.367	2001 n=3.007	2002 n=3.224	2003 n=4.186	2004 n=4.937
<i>Sekse</i>								
Man	91,5	88,5	85,2	88,1	87,2	85,4	82,9	80,3
Vrouw	8,5	11,5	14,8	11,9	12,8	14,6	17,1	19,7
<i>Leeftijd bij uitstroom</i>								
15 jaar of jonger	25,7	24,5	29,2	27,2	30,3	28,3	27,7	31,7
16 tot 18 jaar	53,0	54,6	54,5	53,8	52,5	53,4	53,0	51,1
18 jaar of ouder	21,3	20,9	16,1	19,0	17,2	18,2	19,3	17,1
<i>Geboorteland</i>								
Nederland	69,8	67,7	69,2	69,7	70,6	70,1	73,2	75,0
Marokko	10,3	10,4	8,2	7,4	6,5	6,0	4,8	4,0
Nederlandse Antillen/Aruba	5,1	4,2	4,1	5,2	4,6	4,9	4,0	3,8
Suriname	4,4	4,6	4,2	3,6	3,3	3,5	3,1	2,3
Turkije	1,5	1,3	1,3	1,6	1,3	1,0	1,0	1,1
Overige westerse landen	3,0	3,8	4,2	4,4	5,0	4,4	4,7	4,2
Overige niet-westerse landen	6,0	7,8	8,6	8,1	8,7	9,9	9,1	9,5
<i>Etniciteit</i>								
Nederland	46,6	38,9	38,6	40,6	38,7	41,2	44,4	47,1
Marokko	20,1	21,6	20,0	19,9	18,9	18,3	17,2	15,9
Nederlandse Antillen/Aruba	5,2	5,1	5,0	6,4	6,1	6,1	5,4	5,4
Suriname	11,2	12,9	13,2	10,8	11,7	11,4	10,0	8,9
Turkije	5,0	5,6	5,7	5,2	6,1	4,1	4,3	4,7
Overige westerse landen	2,7	3,4	3,2	3,3	3,3	3,4	3,8	3,9
Overige niet-westerse landen	8,9	12,2	13,9	13,6	15,2	15,5	14,8	14,0
Onbekend	0,3	0,1	0,3	0,3	0,1	0,1	0,1	0,1
<i>Type delict</i>								
Openbare orde*	5,3	5,8	6,4	7,0	5,0	6,0	6,7	7,7
Vermogen zonder geweld	18,6	19,8	17,8	18,5	17,0	16,7	16,5	18,2
Vermogen met geweld	25,7	26,3	27,9	31,4	26,3	26,0	19,5	26,5
Geweld	6,2	5,2	6,6	6,6	5,8	6,5	7,6	8,2
Zeden	2,9	1,7	3,2	3,0	2,8	3,3	2,5	4,1
Overig	2,3	3,0	3,3	2,5	2,3	3,3	2,6	2,8
N.v.t./geen delict	38,9	38,1	34,8	30,9	40,8	38,3	44,6	32,2

Tabel 4.5 Onderzoeksaantallen en achtergronden van pupillen van justitiële jeugdinrichtingen; naar jaar van uitstroom (in kolompercentages) (vervolg)

	1997 n=1.476	1998 n=2.006	1999 n=2.122	2000 n=2.367	2001 n=3.007	2002 n=3.224	2003 n=4.186	2004 n=4.937
<i>Strafrechtelijk verleden</i>								
0 eerdere contacten	29,2	31,1	36,1	37,7	35,6	36,2	34,9	40,4
1-2 eerdere contacten	34,8	41,8	41,4	33,2	34,6	35,6	40,0	44,2
3-4 eerdere contacten	17,5	13,0	12,0	15,4	15,2	15,0	13,6	10,8
5-10 eerdere contacten	15,7	11,8	9,1	12,2	12,7	11,7	10,1	4,5
11 of meer eerdere contacten	2,8	2,4	1,5	1,6	1,9	1,5	1,4	0,1
<i>Leeftijd 1e strafzaak</i>								
15 jaar of jonger	67,2	64,5	66,7	64,9	66,4	64,3	64,2	67,0
16 tot 18 jaar	29,3	32,2	30,1	31,2	30,1	31,8	31,9	29,7
18 jaar of ouder	3,5	3,2	3,1	3,8	3,4	3,7	3,8	3,2
<i>Verblijfsduur</i>								
Korter dan 3 maanden	54,8	51,9	51,0	50,7	52,0	52,7	54,8	52,0
3 tot 6 maanden	19,2	16,9	19,0	17,6	19,7	16,0	14,2	16,6
6 tot 12 maanden	11,7	11,6	10,9	11,4	10,4	12,4	11,9	12,0
12 maanden of langer	14,3	19,6	19,0	20,2	18,0	18,9	19,0	19,4
<i>Verblijfstitel</i>								
Voorlopige hechtenis	44,4	44,6	47,4	44,1	46,4	45,6	44,1	43,0
Jeugddetentie	21,6	21,9	18,1	16,2	14,4	16,8	12,8	15,3
PIJ	3,9	5,2	5,4	6,5	4,4	4,7	4,2	4,6
Ots**	17,8	18,6	18,4	16,4	17,5	19,0	22,5	29,3

* Vernieling, lichte agressie en openbare ordedelicten.

** Inclusief voogdij en vrijwillig voortgezette hulpverlening.

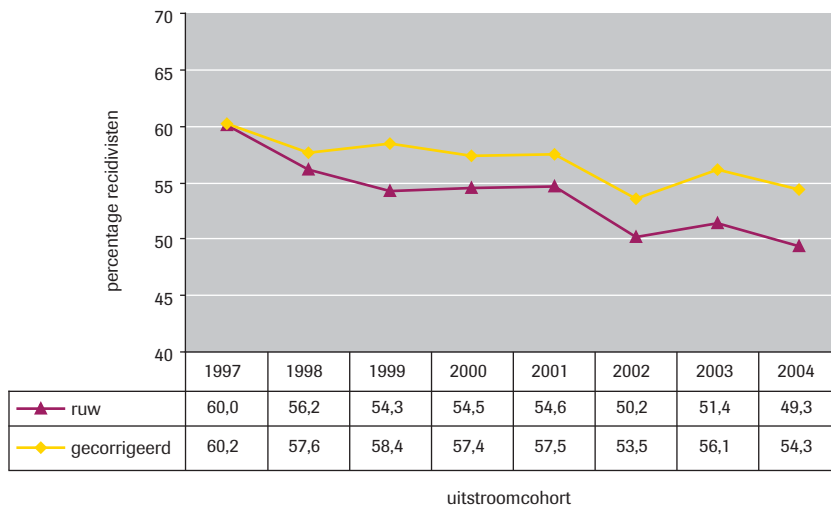
In de onderste rij van de tabel zien we dat het percentage jongeren dat in het kader van een civielrechtelijke maatregel (ots, voogdij of vrijwillig voortgezette hulpverlening) in een JJI werd geplaatst, in de onderzoeksperiode is gestegen. De meeste jongeren die op strafrechtelijke titel – dus naar aanleiding van het plegen van een delict – in een JJI verbleven, stonden terecht voor vermogen met geweld. Het aandeel dat uitstroomde vanuit voorlopige hechtenis, fluctueerde. Het aantal jongeren dat een onvoorwaardelijke vrijheidsstraf uitzat, steeg in absolute zin, maar hun aandeel in de populatie daalde. Het percentage jongeren dat in het kader van een strafrechtelijke maatregel werd opgenomen, schommelde tussen de 5 en 8%. Vanaf het jaar 2000 betrof dit vrijwel uitsluitend plaatsingen in een inrichting voor jeugdigen (PIJ). Daarvoor stroomden nog jongeren uit vanuit 'plaatsingen in het kader van buitengewone behandeling (pibb)' en vanuit 'jeugd-tbr'.

De toename van het aantal oots-ers is gepaard gegaan met een toename van het aandeel meisjes. In 2004 was het percentage meisjes dat uitstroomde meer dan twee keer zo groot als in 1997. Het percentage autochtone jongeren schommelde tussen 1997 en 2004 tussen 39 en 47%. Het aandeel Antillianen/Arubanen en Turken was stabiel, dat van de Marokkaanse en Surinaamse jongeren nam af.⁵⁴ Tabel 4.5 geeft ook een beeld van het strafrechtelijke verleden van de ex-JJI-pupillen. 60% van de jongeren die in 2004 een justitiële jeugdinrichting verlieten, was ooit eerder verdachte in een strafzaak wegens het plegen van een misdrijf. Het geeft aan dat het hier om een relatief zware groep gaat, want bij de jeugdige daders in het algemeen was bijna drie kwart first offender. Het percentage ex-JJI-pupillen zonder strafrechtelijk verleden is in de loop van de onderzoeksperiode wel gestegen, van 29% in 1997 tot 40% in 2004.

Ruwe en gecorrigeerde percentages

De toename van het percentage first offenders, de afname van het aandeel jongeren van Marokkaanse en Surinaamse origine en de toename van het percentage meisjes en oots-ers doen vermoeden dat de ruwe recidivepercentages in de JJI-sector zijn gedaald. Dat blijkt inderdaad het geval te zijn.

Figuur 4.5 Ruwe en gecorrigeerde percentages algemene recidive binnen twee jaar na uitstroom in acht opeenvolgende cohorten van ex-JJI-pupillen



54 De etniciteit van de jongeren is vastgesteld aan de hand van hun geboorteland en dat van de ouders. Dit wordt bijgehouden in TULP.

Figuur 4.5 geeft de ruwe en de gecorrigeerde 2-jarige recidivepercentages voor de acht opeenvolgende uitstroomcohorten weer. De daling is niet monotoon; in 2003 nam de terugval in vergelijking met het jaar daarvoor licht toe. Gerekend over de gehele onderzoeksperiode echter, daalde het ruwe percentage van 60 naar 49%. Gecorrigeerd voor de verschuivingen in tabel 4.5 is de afname iets minder spectaculair: van 60 naar 54%, een afname van 5,5%. De correctie heeft tot gevolg dat de ruwe recidivepercentages naar boven worden bijgesteld. Een deel van de afname moet dus worden toegeschreven aan de besproken veranderingen in de populatie. Gedurende de onderzoeksperiode stroomden er steeds meer jongeren uit die al bij plaatsing in de inrichting een lager samengesteld recidiverisico hadden. Maar ook na correctie voor de gemeten instroomverschillen is er nog sprake van een afname van de recidive. En dat is toch opmerkelijk, want de meeste andere categorieën van justitiablelen lieten in dezelfde periode een stijging van de gecorrigeerde recidivepercentages zien. Alleen bij de onvoorwaardelijke vrijheidsstraffen kon ook een daling worden geconstateerd.

Recidive uitgesplitst naar verblijfstitel

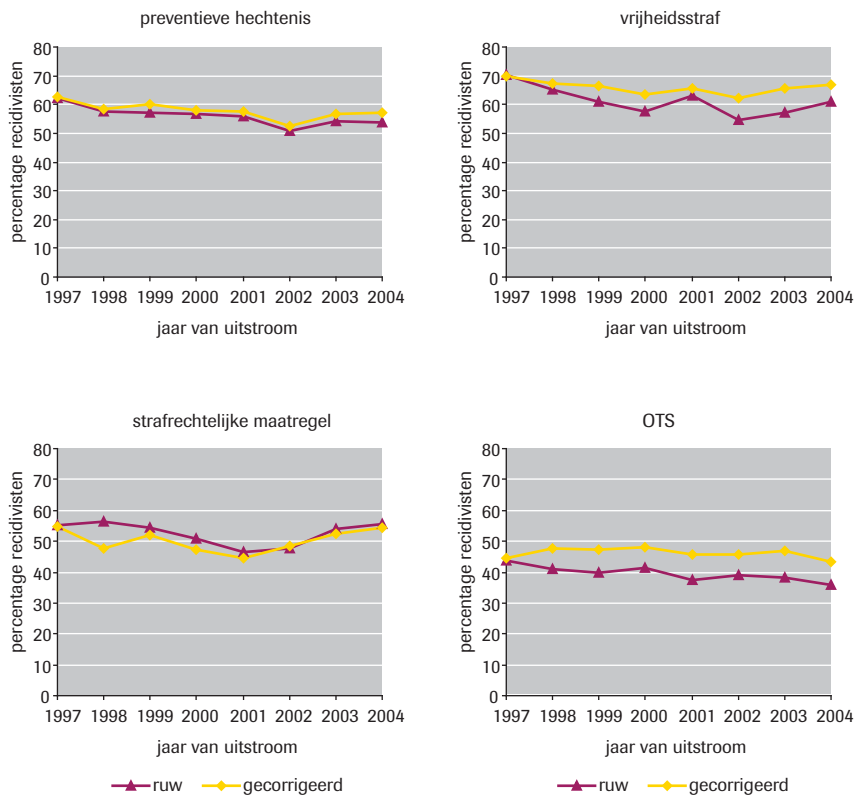
Figuur 4.6 geeft opnieuw de ruwe en gecorrigeerde recidivepercentages, maar nu uitgesplitst naar verblijfstitel. De situatie bij vertrek uit de inrichting was daarbij bepalend. De correcties voor verschuivingen op achtergronden hebben over het algemeen een opwaarts effect. Bij de otsgroep zijn de verschillen tussen de ruwe en gecorrigeerde percentages het grootst. Ondertoezichtstelling met opname in een JJI werd na 1997 in toenemende opgelegd aan jongeren met een relatief laag recidiverisico: meer meisjes en minder jongeren met een strafrechtelijk verleden. Hetzelfde geldt voor de groep die in voorlopige hechtenis de straf uitzat en voor de groep met jeugddetentie.

We zien dat de daling van de gecorrigeerde recidive in de JJI-sector bijna geheel voor rekening komt van de jongeren die een straf kregen opgelegd. De recidive na preventieve hechtenis daalde van 62 naar 57% en na jeugddetentie van 70 naar 67%. De terugval na een strafrechtelijke maatregel schommelde tijdens de onderzoeksperiode tussen 45 en 55%. De terugval na een ots steeg eerst licht en nam daarna weer af tot iets onder het beginniveau. Ots-ers worden in een JJI opgenomen vanwege gedragsstoornissen of een problematische thuissituatie en niet, althans niet primair, naar aanleiding van het plegen van een delict. Toch komt onder deze groep een vrij hoge 'recidive' voor, al zijn de percentages wel lager dan onder de drie overige categorieën.

Hoe de daling van de recidive in de JJI-sector moet worden geïnterpreteerd is niet meteen duidelijk. Vaststaat dat de afname niet kan worden toegeschreven aan verschillen in de verdeling van sekse, leeftijd, etnische achtergrond, type delict en de omvang van het strafrechtelijke verleden. Het effect van deze verschuivingen kon immers worden geneutraliseerd.

Wat hier dan wel een rol speelt, is op grond van de verzamelde gegevens niet met zekerheid te zeggen. Men zou kunnen stellen dat een verblijf in een JJI in de loop van de onderzoeksperiode succesvoller is geworden. Maar dan zal duidelijk moeten worden gemaakt waarom dat alleen voor bestrafte jongeren (in voorlopige hechtenis en na jeugddetentie) geldt en niet voor oters en jeugdigen met een pij-maatregel. Is er in de aanpak van juist deze twee groepen iets gewijzigd dat de gunstige wending van de recidivecijfers kan verklaren? Of is hier iets anders aan de hand en moet de verklaring elders worden gezocht? In de slotparagraaf komen we op deze vraag terug.

Figuur 4.6 Ruwe en gecorrigeerde percentages 2-jarige algemene recidive in acht opeenvolgende populaties van ex-JJI-pupillen; naar verblijfstitel bij uitstroom



4.2.5 Ex-terbeschikkinggestelden

In de jaren '80 en '90 van de vorige eeuw brachten Van Emmerik (1985, 1989) en Leuw (1995, 1999) de terugval in de tbs-sector in kaart. Zij inventariseerden de nieuwe justitiecontacten in vier opeenvolgende cohorten van ex-terbeschikkinggestelden. Elke cohort besloeg een uitstroombperiode van vijf jaar. Inmiddels wordt ook in deze sector gewerkt volgens de methode van de Recidivemonitor. Aan de vier cohorten uit de onderzoeken van Van Emmerik en Leuw zijn twee nieuwe uitstroombperioden van vijf jaar toegevoegd. Tabel 4.6 geeft de aantallen in de zes opeenvolgende cohorten plus enkele karakteristieken van de betreffende daders. Er zijn nu gegevens bekend over een periode tot dertig jaar na het ontslag uit de tbs. De eerste ex-terbeschikkinggestelden stroomden in de jaren 1973-1978 uit, de laatste in de periode 1998-2003. Het aantal terbeschikkinggestelden dat in vijf jaar uitstroomb daalde steeds, maar is na de eeuwwisseling voor het eerst weer toegenomen.⁵⁵ De laatste cohort telt 356 personen, 30% meer dan in de vijf jaar daarvoor. Belangrijk is ook de ontwikkeling in de duur van de maatregel. De tbs van de personen die in de laatste vijf jaar zijn uitgestroomd, duurde gemiddeld gesproken het langst. Ten opzichte van de groep van wie de maatregel in de jaren 1994-1998 werd beëindigd, is de gemiddelde verblijfsduur in de laatste cohort met een kwart toegenomen. In vergelijking met ex-gedetineerden of volwassen daders in het algemeen komen in de tbs relatief weinig daders voor die in het buitenland geboren zijn, al is het aandeel allochtonen de laatste vijftien jaar wel gestegen. Het aantal justitiecontacten voorafgaand aan de tbs (het strafrechtelijke verleden) is in de laatste cohort het hoogst. Ook op dit punt is er sprake van een geleidelijke toename. De leeftijd waarop men voor het eerst met justitie in aanraking kwam, is gedaald. De meeste terbeschikkinggestelden kregen de maatregel opgelegd naar aanleiding van een geweldsdelict. In vergelijking met de voorlaatste cohort is het aandeel vermogen met geweld vrij sterk toegenomen. Het aandeel zedendelinquenten nam in de laatste uitstroombperiode af. In de eerste cohorten kwamen nog veel vermogensdelinquenten voor die bij het plegen van hun delict geen geweld toepasten.

Over het verblijf in de tbs zijn slechts enkele zaken bekend. De uitstroombgegevens van DJI bevatten alleen informatie over eventueel vluchtgedrag en over de wijze van beëindiging van de maatregel. Diagnostische gegevens en informatie over het behandelingsucces ontbreken. Het percentage ex-terbeschikkinggestelden dat tijdens de tbs op enig moment ongeoorloofd afwezig was, is in de periode 1994-1998 scherp gedaald, maar was de laatste perioden stabiel. Het percentage contraire beëindigingen, waarbij de rechter tegen het advies van de behandelaars of de officier van justitie in de maatregel stopt, steeg aanvankelijk, maar is de laatste tien jaar weer

55 Het gaat in dit onderzoek alleen om de intramuraal ten uitvoer gelegde tbs-maatregelen ('met bevel tot verpleging van overheidswege') die door de rechter werden beëindigd.

gezakt. Het percentage beëindigingen vanuit een situatie van proefverlof nam af. Alleen eind jaren '90 nam het aandeel tbs-ers dat met ontslag ging na een periode van proefverlof ten opzichte van de vorige cohort toe. In de laatste cohort is het percentage weer lager.

Tabel 4.6 Kenmerken van ex-terbeschikkinggestelden uitgestroomd in 1974-2003 (in kolompercentages)

Cohort	1	2	3	4	5	6	
Uitstroomperiode	74-78	79-83	84-88	89-93	94-98	99-03	totaal
Aantal	497	394	371	312	275	356	2.205
% man	91,8	95,9	94,6	96,5	94,5	95,5	94,6
Gemiddelde leeftijd bij uitstroom	37,5	32,9	33,7	34,9	34,8	37,4	35,7
% geboren in Nederland	92,3	89,3	86,1	75,8	75,4	73,1	83,2
Leeftijd 1e justitiecontact	30,3	26,6	24,5	23,7	23,1	23,1	25,4
Gem. aantal eerdere justitiecontacten	7,3	7,3	9,5	11,7	11,2	12,1	9,7
<i>Type delict van de uitgangzaak*</i>							
Vermogen geen geweld	46,5	32,4	28,8	25,6	21,4	18,3	30,6
Vermogen met geweld	7,6	11,0	12,1	12,8	14,9	19,4	12,5
Agressie	36,3	55,2	70,1	70,2	73,6	71,6	60,2
Zeden	22,2	22,1	23,7	26,9	20,7	17,1	22,1
Opiumwet	0,6	1,2	0,8	1,3	0,4	1,7	1,0
Overig	3,8	5,5	5,4	8,7	5,4	8,1	5,9
Verkeer	4,0	3,8	3,8	5,8	1,1	2,8	3,6
Gemiddelde duur maatregel in jaren	6,1	4,6	5,5	5,7	6,1	7,6	5,9
% ongeoorloofd afwezig	61,8	60,7	54,8	40,4	8,7	8,1	42,0
% beëindigingen vanuit proefverlof	64,4	59,1	50,1	44,2	52,2	41,0	52,9
% contraire beëindigingen	17,9	37,3	48,0	60,2	51,5	41,6	40,3

* De percentages tellen op tot meer dan 100%, omdat binnen één zaak meer dan één type delict kan voorkomen.

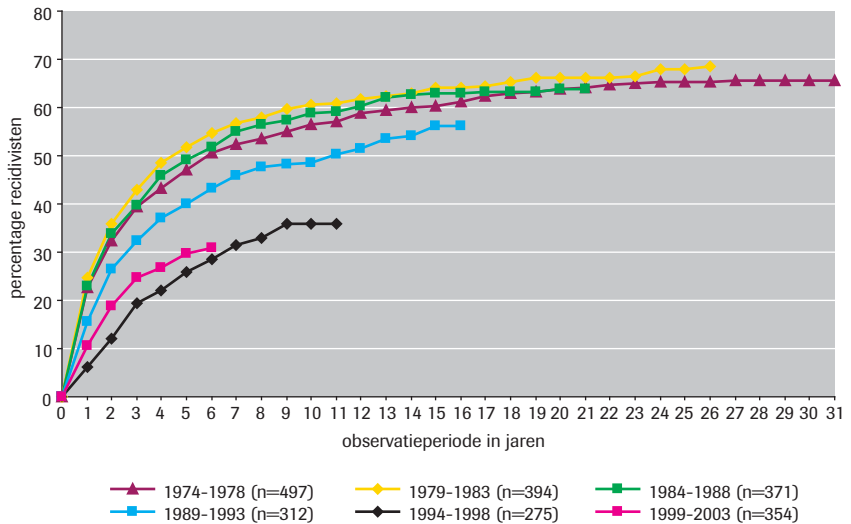
Ruwe recidivepercentages

Een tbs-maatregel kan alleen worden opgelegd voor misdrijven met een strafdreiging van vier jaar of meer. In deze paragraaf wordt daarom vooral aandacht besteed aan de 'ernstige' recidive.⁵⁶ Figuur 4.7 geeft de percentages in de zes opeenvolgende cohorten te zien. De ex-terbeschikkinggestelden uit de eerste cohort konden het langst worden gevolgd. Zij zijn uitgestroomd in de periode 1974-1978; voor hen geldt een observatieperiode van minimaal 26 en maximaal 31 jaar. Van de personen in de

⁵⁶ De prevalentie van algemene recidive is doorgaans 1,2 tot 1,4 keer zo groot als die van de ernstige recidive.

laatste cohort zijn alleen kort na de beëindiging van de maatregel gegevens bekend: ten minste over een periode van 1,5 en ten hoogste over een periode van 6 jaar.⁵⁷

Figuur 4.7 Prevalentie van ernstige recidive in zes opeenvolgende uitstroomcohorten van ex-terbeschikkinggestelden



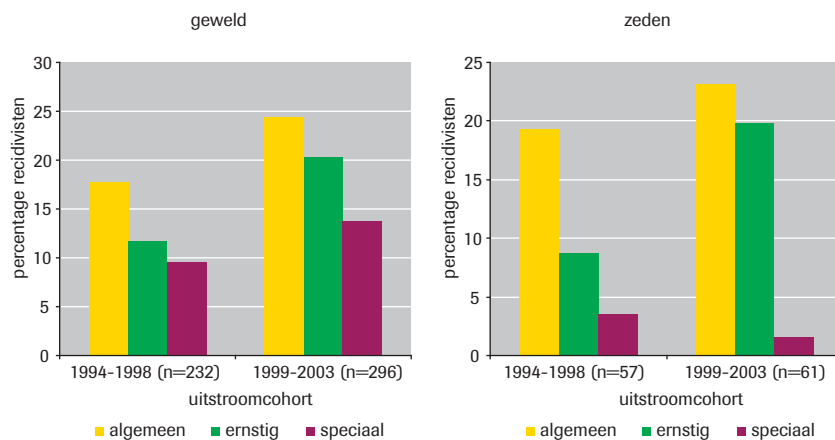
In de eerste drie cohorten is de recidive van een vergelijkbaar niveau, zowel op de korte als op de lange termijn. Twee jaar na beëindiging van de maatregel had ongeveer een derde van de ex-terbeschikkinggestelden één of meer nieuwe justitiecontacten naar aanleiding van een misdrijf met een strafdreiging van vier jaar of meer. Uiteindelijk tendeeft de prevalentie van ernstige recidive in de eerste drie cohorten naar maximaal 70%. Na de jaren '80 van de vorige eeuw daalde de recidive. De 2-jarige ernstige recidive in de vierde cohort, onder ex-terbeschikkinggestelden die hun maatregel in de periode 1989-1993 beëindigd zagen, bedraagt 27%; in de uitstroomperiode daarna, 1994-1998, nog slechts 12%. Na 1998 stijgt de recidive echter weer. Van de laatste cohort werd 19% binnen twee jaar opnieuw voor een ernstig misdrijf vervolgd. Dit is lager dan in de eerste vier perioden, maar hoger dan in de vijfde cohort. Hoe dit op de langere termijn uitpakt, is nog niet bekend. Op grond van het beeld tot dusver mag men verwachten dat het percentage ernstige recidivisten in de voorlaatste cohort niet boven 40% zal uitstijgen. Maar voorzichtigheid is hier op zijn plaats. Aan het eind van de observatieperiode zijn de percentages vastgesteld op een kleiner aantal waarnemingen.

57 92% van de personen uit cohort 6 konden twee jaar of langer worden gevolgd. Tot anderhalf jaar na de beëindiging van de tbs is de recidive vastgesteld op basis van alle observaties.

Box 4.1 Tbs-ers en speciale recidive

Waar vroeger nog veel vermogensdelinquenten de maatregel kregen opgelegd, bestaat de tbs-populatie tegenwoordig vrijwel uitsluitend uit geweldplegers en zedendelinquenten. Dit is een van de redenen waarom de recidive in deze sector is gedaald. De base rate voor gewelds- en zedendelinquenten is over het algemeen lager dan voor plegers van vermogensmisdrijven (zie paragraaf 3.4.2). Figuur 4.8 geeft voor de gewelds- en zedendelinquenten uit de laatste twee cohorten de 2-jarige prevalentie van drie vormen van recidive weer: algemene, ernstige en speciale recidive. Algemene recidive betreft alle nieuwe strafzaken, en ernstige recidive handelt om misdrijven met een strafdreiging van vier jaar of meer. De betekenis van 'speciale recidive' hangt af van het uitgangsdelict: voor de geweldplegers is speciale recidive een nieuw geweldsmisdrijf en voor de zedendelinquenten een nieuw zedendelict. Het onderscheid tussen gewelds- en zedendelinquenten is vrij grof. De indeling is gebaseerd op de wetsartikelen vermeld in de OBD.

Figuur 4.8 Prevalentie van verschillende vormen van strafrechtelijke recidive binnen twee jaar na ontslag in twee uitstroomcohorten van ex-terbeschikkinggestelden; naar soort uitgangsdelict

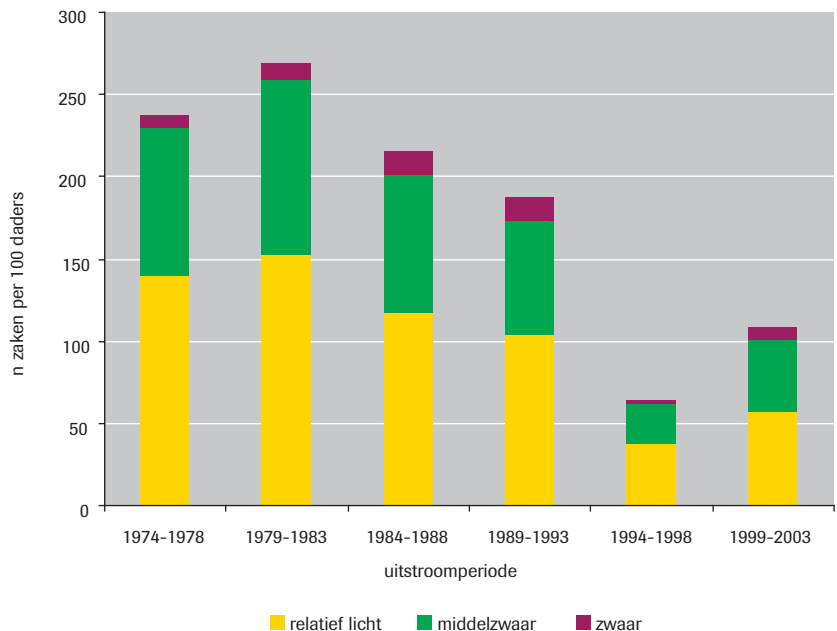


In de figuur zien we terug dat de recidive de laatste jaren is gestegen. Zowel onder de geweldplegers als onder de zedendelinquenten in de tbs-sector is de ernstige recidive in cohort 1999-2003 hoger dan in cohort 1994-1998. Hetzelfde geldt voor de algemene recidive. Bij de geweldplegers is ook de speciale recidive toegenomen. Ten opzichte van de vorige cohort is dus niet alleen de proportie recidivisten gestegen, maar meer specifiek ook het aantal dat opnieuw voor het plegen van een geweldsdelict werd vervolgd. Bij de zedendelinquenten komt speciale recidive duidelijk minder vaak voor en is er ook geen sprake van een stijging.

Omvang van de recidive

De ontwikkeling van de recidive in de tbs-sector kan ook worden geïllustreerd aan de hand van de omvang van de hernieuwde criminaliteit. Zoals uiteengezet in hoofdstuk 4 wordt bij elke meting van de Recidivemonitor naast de prevalentie ook de frequentie van nieuwe justitiecontacten berekend. Figuur 4.9 drukt het totale aantal recidivezaken in de opeenvolgende cohorten uit, onderscheiden naar de ernst van gepleegde delicten. Het gaat om justitiecontacten binnen twee jaar na het ontslag. De aantallen gelden voor groepen van honderd daders. Voor elke groep van honderd ex-terbeschikkinggestelden is nagegaan hoeveel nieuwe strafzaken er na twee jaar aan de justitiële documentatie zijn toegevoegd. De recidiveomvang laat globaal hetzelfde beeld zien als de recidiveprevalentie. Na een lichte stijging daalt de recidive tot in de voorlaatste uitstroomperiode, waarna weer een stijging inzet. Deze ontwikkeling geldt zowel voor de recidive naar aanleiding van middelzware delicten, met een strafdreiging van vier tot acht jaar, als voor de recidive naar aanleiding van lichtere misdrijven. De zeer ernstige recidive met delicten met een strafdreiging van acht jaar of meer volgt een iets ander verloop. Deze stijgt langer door, maar daalt dan evengoed. Ook voor de zeer ernstige recidive ligt het laagste punt bij cohort 1994-1998.

Figuur 4.9 Omvang van de strafrechtelijke recidive binnen twee jaar na ontslag in zes opeenvolgende cohorten van ex-terbeschikkinggestelden

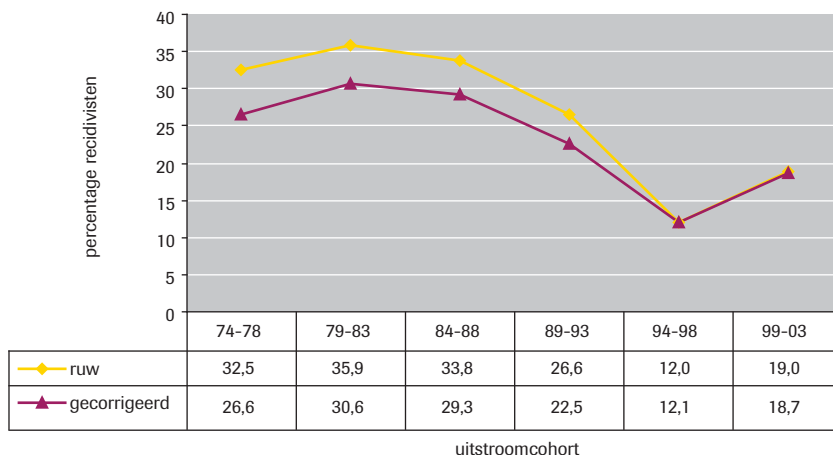


Gecorrigeerde cijfers

Figuren 4.7 en 4.9 maken duidelijk dat de recidive in de tbs-sector de laatste jaren weer is gestegen, na een sterke daling in de jaren '80 en '90 van de vorige eeuw. Zegt deze ontwikkeling iets over de invulling van de maatregel? Misschien, maar misschien ook niet. De hoogte van de recidive houdt allicht verband met het succes van de behandeling tijdens de tbs, maar is ook afhankelijk van andere factoren, zoals de achtergronden van de daders, de aard en ernst van hun stoornissen en hun *post exit* situatie. In dit onderzoek zijn geen stoornis- en behandelgegevens beschikbaar. We weten niet waarop de behandeling van de ex-terbeschikkinggestelden was gericht en kennen evenmin de omstandigheden waarin zij na ontslag verkeerden. We weten wel dat de populatie van ex-terbeschikkinggestelden in een aantal opzichten van samenstelling is veranderd. Tabel 4.6 geeft enkele van de verschuivingen weer. De mate waarin deze van invloed zijn geweest op de recidivecijfers, kan wel worden onderzocht.

Figuur 4.10 geeft voor de zes cohorten de ernstige recidive te zien, twee jaar na het ontslag uit de tbs. Naast de ruwe percentages zijn nu ook de gecorrigeerde percentages afgebeeld. De gecorrigeerde cijfers laten zien hoe hoog de recidive zou zijn geweest indien de ex-terbeschikkinggestelden in termen van sekse, leeftijd, geboorteland, het type uitgangsdeldict en hun delictgeschiedenis in alle uitstroomperioden dezelfde kenmerken zouden hebben gehad als de ex-terbeschikkinggestelden uit de periode 1999-2003. Ze geven met andere woorden aan welke uitstroomresultaten de sector zou hebben gekend indien er op deze punten geen verschillen tussen de zes cohorten waren geweest.

Figuur 4.10 Ruwe en gecorrigeerde percentages ernstige recidive binnen twee jaar na ontslag in zes opeenvolgende cohorten van ex-terbeschikkinggestelden



De correctie voor instroomverschillen op de achtergrondkenmerken heeft in de eerste vier cohorten een verlaging van de recidivepercentages tot gevolg. Kennelijk omvatten deze uitstroomperioden relatief veel ex-terbeschikkinggestelden die al bij instroom in de tbs, dus nog voor aanvang van hun behandeling, een hoger recidiverisico hadden. Het verschil hangt vooral samen met de factor delicttype. In de eerdere cohorten komen veel vermogensdelinquenten voor (zie tabel 4.6). Dezen recidiveren over het algemeen vaak en dat verklaart voor een deel althans waarom de ruwe percentages in de eerdere uitstroomcohorten hoger zijn dan in de latere cohorten. Als de invloed van deze en andere instroomverschillen wordt gecorrigeerd, neemt het verschil in recidive tussen de zes cohorten af. Maar ook na correctie voor de gemeten instroomverschillen blijft er een duidelijke ontwikkeling in de uitstroomresultaten te zien. Na een lichte stijging in de jaren '70, neemt de 2-jarige ernstige recidive in de jaren '80 en '90 af. Na 1998 neemt de recidive weer toe. Opvallend is dat in de voorlaatste cohort het gecorrigeerde percentage nauwelijks afwijkt van het ruwe recidivepercentage. Dit betekent dat deze groep vergelijkbaar is met de laatste cohort, althans wat betreft het risicoprofiel berekend op basis van de kenmerken opgenomen in de OBJD. De stijging van het recidivepercentage na 1998 kan *niet* worden toegeschreven aan verschuivingen op deze statische moderatorvariabelen. Misschien vormde de laatste cohort in andere opzichten een zwaardere groep dan de groep ex-terbeschikkinggestelden uit de periode 1994-1998 en is daarom de recidive weer gestegen. In de slotparagraaf komen we terug op de betekenis van dit resultaat.

4.3 Recidive in internationaal perspectief

Nu we weten hoe de terugval van justitiabelen zich in Nederland in grote lijnen heeft ontwikkeld, rijst de vraag hoe deze zich verhoudt met de recidive in het buitenland. Uit een recent gehouden postenquête (Wartna & Nijssen, 2006) blijkt dat minstens veertien Europese landen op nationaal niveau cijfers beschikbaar hebben. Tabel 4.7 geeft een overzicht van deze landen en de belangrijkste publicaties op dit terrein.

De verschillen in de opzet van het nationale recidiveonderzoek zijn groot. In Engeland werden de recidivepercentages tot voor kort gewoonlijk berekend over de periode tot twee jaar na het opleggen van de straf of het vertrek uit de gevangenis. In het Duitse onderzoek hanteert men een observatietermijn van vier jaar, in het Franse en Oostenrijkse onderzoek werden de personen gedurende een periode van vijf jaar gevolgd. Sommige landen houden jaarlijks metingen, zoals sinds 2005 in Nederland het geval is. Ook Engeland & Wales, Schotland, Zweden en Noorwegen hebben cijfers over een langere periode.

Tabel 4.7 Nationale recidivestudies in Europa

Land	Publicatie	Dadergroep
Oostenrijk	Pilgram (1994)	Volwassenen en jeugdigen bestraft in 1983
Denemarken	Prison & Probation Department (2001)	Personen vrijgelaten uit detentie of veroordeeld tot een voorwaardelijke straf in 1996-1999
Engeland & Wales	Prime (2002)	Personen vrijgelaten uit detentie of veroordeeld tot een toezichtstraf in 1999
	Home Office (2005)	Jeugdigen (10-17) veroordeeld of vrijgelaten uit detentie in 2003
	Ministry of Justice (2008)	Volwassenen vrijgelaten uit detentie of onder toezicht gesteld van de reclassering in 2006
Finland	Hypén (2004)	Onvoorwaardelijk gestraften vrijgelaten in 1993-2001
Frankrijk	Kensey & Tournier (2005)	Gedetineerden vrijgelaten van 1996/5 tot 1997/4
Duitsland	Jehle, Heinz & Sutterer (2003)	Personen bestraft of vrijgelaten uit detentie in 1994
IJsland	Baumer et al. (2002)	Gedetineerden vrijgelaten van 1994/1 tot 1998/11
Ierland	O'Donnell, Baumer & Hughes (2008)	Gedetineerden vrijgelaten van 2001/1 tot 2004/11
Noord-Ierland	McMullen & Rudy (2005)	Volwassenen met een extramurale straf of vrijgelaten uit detentie in 2001
	Decodts (2005)	Jeugdigen (10-16) met een extramurale straf of vrijgelaten uit detentie in 2001
Norwegen	Statistics Norway (2006)	Politieverdachten uit 1996
Schotland	Scottish Executive (2005)	Volwassenen vrijgelaten uit detentie of bestraft met een extramurale sanctie in 1999
Zweden	National Council (2005)	Personen berecht in 1999
Zwitserland	Storz (1997)	Volwassen gedetineerden vrijgelaten in 1982-1990
Nederland	Wartna et al. (2008)	Volwassenen en jeugdigen (12-18) bestraft in 1997-2004

De meeste studies hebben betrekking op alle daders die in de selectieperiode zijn vervolgd. Sommige landen volgen echter bepaalde subgroepen. Zo zijn in Frankrijk, Finland, IJsland en Zwitserland alleen cijfers bekend over ex-gedetineerden.⁵⁸ In de noordse landen (Noorwegen, Zweden, Finland, Denemarken en IJsland) is een recidivemonitor in de maak gericht op personen die waren gedetineerd of te maken kregen met de reclassering.

58 In Zwitserland zijn ook cijfers bekend over andere sancties, maar alleen voor first offenders.

Naast verschillen in de onderzoeksgroep, de selectieperiode en de observatietermijn zijn er verschillen in de operationele definitie van het recidivebegrip. Ook dit staat een goede vergelijking van de onderzoeksresultaten in de weg. In het Finse onderzoek van Hypén (2004) wordt onder 'recidive' bijvoorbeeld een heropname in een penitentiaire inrichting verstaan. In het noordse vijflandenonderzoek is het recidivecriterium ruimer en gaat het om gevallen waarin de persoon opnieuw in contact is gekomen met de *Prison & Probation Service*.

In tabel 4.8 staan de nationale recidivecijfers vermeld van de zeven Europese landen die zich richten op de totale populatie van 'adjudicated offenders'. We zien dat in het onderzoek in Noorwegen recidive een afwijkende betekenis heeft. In het Noorse onderzoek is sprake van recidive als de dader opnieuw als verdachte bij de politie is ingeschreven. In de meeste andere landen is een 'nieuwe veroordeling' het criterium. De vraag is echter of met deze term steeds hetzelfde wordt bedoeld. Tellen in Engeland dezelfde delicten mee als in Schotland of in Zweden? Dat staat niet zonder meer vast. In Nederland wordt ongeveer de helft van alle misdrijfzaken afgedaan door het Openbaar Ministerie. Als het onderzoek zich in Nederland zou beperken tot die gevallen waarin het kwam tot een veroordeling door de rechter, dan zou de recidive een stuk lager uitvallen. Ook in Duitsland worden de OM-afdoeningen meegenomen, zij het alleen voor minderjarige daders (Jehle, Heinz & Sutterer, 2003).

Tabel 4.8 Recidivepercentages in zeven Europese landen

Land	Dadergroep	Selectieperiode	Leeftijd in jaren	N	Recidivecriterium	Percentage recidivisten na x jaar						
						1	2	3	4	5	6	
Oostenrijk	Bestrafte volwassenen	1983	≥ 14	69.267	Veroordeling							38,0
Duitsland	Personen bestraft of vrijgelaten uit detentie	1994	≥ 14	947.382	Veroordeling							35,7
Nederland	Personen bestraft voor een misdrijf	1997	≥ 12	153.834	Veroordeling	18,1	26,7	32,2	36,1	39,3	41,8	
Zweden	Personen veroordeeld voor een delict	1999	≥ 15	76.700	Veroordeling	22,0		36,0				
Noorwegen	Politieverdachten	1996	≥ 15	65.086	Beschuldiging	19,6	30,5	37,4	41,5	43,4		
Schotland	Personen veroordeeld of vrijgelaten uit detentie	1999	≥ 16	45.245	Veroordeling	31,0	42,0	49,0	53,0			
Engeland/Wales	Personen vrijgelaten of bestraft met een toezichtstraf	1999	≥ 10	13.316	Veroordeling		48,0					

Elk land put zijn gegevens uit een domein van delicten dat niet zonder meer overeenkomt met dat van andere landen. Zolang niet in detail duidelijk is wat elk land onder 'recidive' verstaat, kan op grond van tabel 4.8 slechts provisorisch een vergelijking worden gemaakt. De recidivepercentages van Nederland zijn afkomstig uit het onderzoek van 2007 en hebben betrekking op de volwassen en minderjarige daders vervolgd in 1997. Ze zijn ongeveer even hoog als in Oostenrijk en Duitsland. In Noorwegen liggen de percentages iets hoger, maar dat kan te maken hebben met het feit dat in het Noorse onderzoek een ruimere definitie werd gehanteerd. De recidivepercentages in Schotland en Engeland & Wales zijn duidelijk hoger dan in de eerste vijf landen. Engeland & Wales spannen de kroon. Bijna de helft van de vervolgte daders heeft binnen twee jaar een nieuwe veroordeling wegens een 'standard list offence'.

Tabel 4.9 geeft de recidivepercentages voor de ex-gedetineerden in acht Europese landen en de Verenigde Staten. De cijfers voor de Verenigde Staten zijn afkomstig uit het onderzoek van Langan en Levin (2002). Zij volgden een groep van bijna 34.000 personen die in 1994 werden ontslagen uit de *state prisons* van vijftien van de vijftig Amerikaanse staten. Alle landen in de tabel gebruikten een 'nieuwe veroordeling' als criterium voor recidive. Hoewel men zou denken dat de dadergroep nu homogener is, lopen de resultaten sterker uiteen dan in tabel 4.8. Twee jaar na ontslag uit de gevangenis bedraagt het recidivepercentage in Zwitserland 26% en in Schotland, dat samen met Engeland & Wales opnieuw hoog scoort, 60%. De percentages van de ex-gedetineerden uit Nederland zijn maar iets lager dan die van het Verenigd Koninkrijk. Bij het onderzoek onder vervolgte daders in het algemeen neemt Nederland een middenpositie in, maar wat betreft de recidivecijfers voor ex-gedetineerden lijken we tot de Europese top te behoren.

Tabel 4.9 Recidive onder gedetineerden in acht Europese landen en in de VS

Land	Ontslagperiode	N	Leeftijd	Detentie ratio*	% gedetineerden opnieuw veroordeeld na x jaar					
					1	2	3	4	5	6
Zwitserland	1988	6.393	≥ 18	79	12,0	26,0	34,0	40,0	45,0	48,2
Frankrijk	1996-1997	2.859	≥ 13	89					51,9	
IJsland	1994-1998	1.176	≥ 18	44			37,0		53,0	
VS (15 staten)	1994	33.796	≥ 18	600	21,5	36,4	46,9			
Noord-Ierland	2001	703	≥ 17	52		45,0				
Duitsland	1994	946.136	≥ 14	90				56,4		
Nederland	1997	18.281	≥ 18	85	43,0	54,4	60,9	65,2	68,3	70,8
Engeland & Wales	2001	14.569	≥ 18	127		58,2				
Schotland	1999	5.738	≥ 16	120	46,0	60,0	67,0	71,0		

* Bron: International Centre for Prison Studies / World Prison Brief

De vraag is uiteraard in hoeverre deze cijfers vergelijkbaar zijn. Is de groep van ex-gevangenen uit Schotland in alle opzichten vergelijkbaar met die uit Nederland? Waarschijnlijk niet. In het Verenigd Koninkrijk worden 'subsidiar gehechten' buiten het onderzoek gelaten, maar in het Nederlandse onderzoek tellen de personen die hun verkeersboete niet hebben betaald gewoon mee. Ook de verschillen in selectieperioden beïnvloeden de uitkomsten. De groep die in Zwitserland werd gevolgd, stamt uit 1988. Misschien zijn de terugvalpercentages nu hoger. Er doen zich bovendien verschillen voor in de minimumleeftijd van de personen in het onderzoek. Ook dat zal een rol hebben gespeeld, want zoals bekend is leeftijd een sterke voorspeller van latere recidive.

Als een land restrictief is in het gebruik van de onvoorwaardelijke vrijheidsstraf, vallen de recidivepercentages na een detentie als vanzelf hoger uit omdat alleen de personen met een hoog recidiverisico naar de gevangenis worden gestuurd. Misschien is dat wel de reden dat de resultaten voor de Verenigde Staten beter zijn dan voor de meeste Europese landen. Afgaand op de 'detentieratio', het aantal gevangenen per 100.000 inwoners, zijn de VS veel minder terughoudend in het toepassen van gevangenisstraf. Maar misschien speelt hier ook wel een rol dat niet alle staten meededen aan het onderzoek. De VS kennen niet één systeem waarin in alle justitiële gegevens zijn terug te vinden. De informatie moet uit diverse, gelaagde systemen worden gefilterd en samengesteld. Het risico dat men strafzaken mist is hierdoor groter, met als gevolg dat de recidive wordt onderschat.

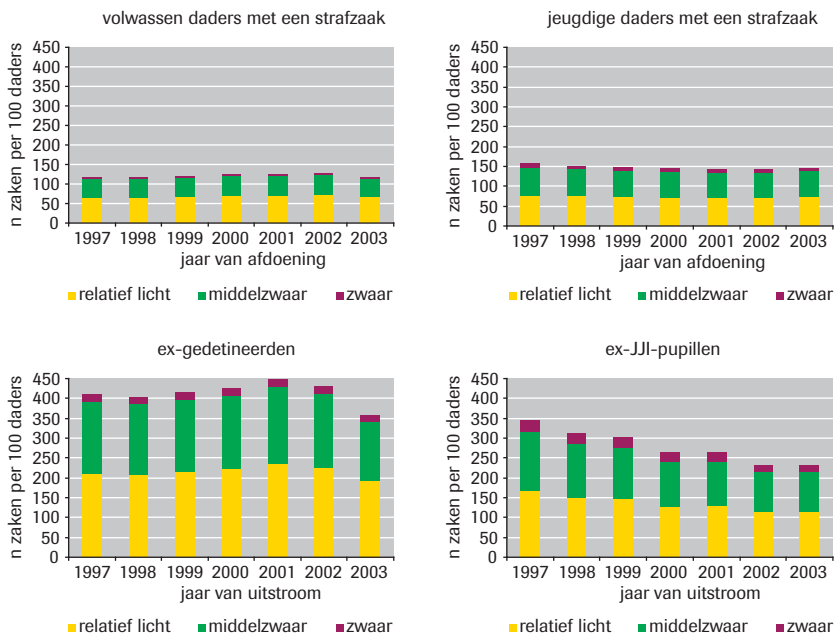
De hoogte van de detentieratio biedt geen verklaring voor de resultaten in Zwitserland, Frankrijk, IJsland en ook Noord-Ierland. In deze landen is het aantal gevangenen per 100.000 inwoners relatief laag, maar de recidivepercentages ook. Frankrijk, Duitsland en Nederland kennen globaal dezelfde detentieratio's. Toch is het recidivepercentage onder ex-gedetineerden in Frankrijk veel lager dan in Duitsland en Nederland. Onduidelijk blijft waarom. IJsland kent het laagste aantal gevangenen op de bevolking, maar de recidive op dit eiland is niet bijzonder hoog. Baumer (2002) maakte eerder een internationale vergelijking. Zijn conclusie was dat de terugvalpercentages in IJsland 'niet merkbaar lager zijn dan gemeten in andere landen' (p. 52). Afgaand op de cijfers in tabel 4.9 lijkt deze conclusie echter niet juist. IJsland doet het beter dan Duitsland, Nederland, Amerika en het Verenigd Koninkrijk.

4.4 Slot: mogelijke verklaringen

De korte trip naar het buitenland heeft nog eens duidelijk gemaakt hoe belangrijk het is om de recidive steeds op dezelfde wijze te meten. De veertien landen uit tabel 4.7 hebben zich verenigd in de *European research group on national reconviction rates* (ERNR), een werkgroep die nagaat

welke mogelijkheden en belemmeringen er zijn om tot een Europese standaard te komen. In Nederland maakt de aanpak van de Recidivemonitor het mogelijk de uitkomsten van verschillende metingen tegen elkaar af te zetten. Figuur 4.11 laat dit nog eens zien. De figuur toont de ontwikkeling van de *omvang* van de terugval in vier van de vijf vaste onderzoeksgroepen van de Recidivemonitor. De omvang van de recidive is het totale aantal nieuwe justitiecontacten dat een groep in een bepaalde periode heeft opgebouwd. In dit geval gaat het om een periode van twee jaar. Voor iedere cohort wordt per groep van honderd daders het aantal nieuwe justitiecontacten afgebeeld, onderscheiden naar de ernst van de gepleegde delicten.

Figuur 4.11 Omvang van de strafrechtelijke recidive binnen twee jaar na oplegging van de straf c.q. vertrek uit de inrichting

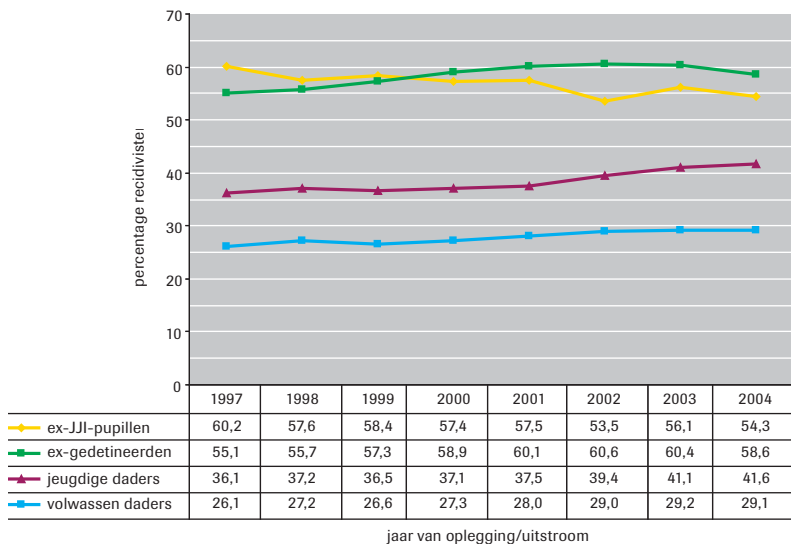


Onder jeugdige en volwassen daders is de omvang van de recidive betrekkelijk stabiel en van een vergelijkbaar niveau. Dit laatste is verrassend, want onder jeugdigen is de prevalentie van recidive hoger dan onder volwassenen. Onder volwassen daders zijn er minder recidivisten, maar de recidivisten die er zijn hebben in twee jaar tijd gemiddeld meer nieuwe strafzaken opgebouwd dan de recidivisten onder de jeugdige daders. Bij ex-gedetineerden komt de meeste recidive voor, meer nog dan onder jongeren uit justitiële jeugdinrichtingen. In beide sectoren doen zich

grotere schommelingen voor dan bij jeugdige en volwassen daders in het algemeen. Een deel daarvan is toe te schrijven aan instroomverschillen: verschuivingen in de achtergronden van de personen die in een justitiële inrichting terecht komen.

Om het effect van instroomverschillen tegen te gaan worden in de Recidivemonitor ruwe percentages gecorrigeerd. Figuur 4.12 geeft de prevalentie van de 2-jarige, algemene recidive voor dezelfde groepen als hiervoor: de volwassen daders, de jeugdige daders, de ex-gedetineerden en de jongeren die zijn uitgestroomd uit een justitiële jeugdinrichting.⁵⁹ De figuur geeft de *gecorrigeerde* percentages. Er is rekening gehouden met registratie-effecten en met verschuivingen in de sekse, leeftijd en het geboorteland van de daders, het type delict dat zij pleegden en de omvang van hun strafrechtelijke verleden.

Figuur 4.12 Gecorrigeerde percentages algemene recidive twee jaar na oplegging van de straf c.q. vertrek uit de inrichting; naar jaar van oplegging/uitstroom



In drie van de vier groepen is de recidive in de periode 1997-2004 gestegen. Ook voor de ex-terbeschikkinggestelden zijn de terugvalpercentages in de jaren rond de eeuwwisseling toegenomen (zie figuur 4.10). Alleen in

59 De resultaten van de ex-terbeschikkinggestelden passen niet in het plaatje, omdat zij in verband met de lage aantallen worden ondergebracht in uitstroomperiodes van vijf jaar. Voor volwassen en jeugdige daders samen schommelde het 6-jarige recidivepercentage in de periode 1997-2004 tussen 42 en 46%. Dat ligt in de buurt van de percentages die Van der Werff meldde in haar recidivestudies over de veroordeelde daders uit 1966 en 1977 (zie paragraaf 2.4). De percentages zijn echter niet goed vergelijkbaar. Van der Werff (1978, 1986) rekende alleen de nieuwe veroordelingen mee en niet de afdoeningen door het OM, zoals bij de Recidivemonitor het geval is.

de JJI-sector is de tweejarige recidive gedaald, van 60% voor de groep uit 1997 naar 54 voor de groep uit 2004.

De toename van de recidive in de diverse sectoren van ons strafrecht-systeem kan *niet* worden toegeschreven aan veranderingen in de gemeten statische criminogene factoren. Evenmin kan worden gesteld dat de percentages zijn gestegen als gevolg van registratie-effecten. Het onderzoek heeft in elk geval aannemelijk gemaakt dat er sprake is van reële groei. De recidivecijfers zijn niet gestegen omdat er meer delicten worden aangegeven of worden opgelost, of omdat er nu meer aandacht uitgaat naar daders die al eerder in contact kwamen met politie en justitie. Voor al deze zaken is gecontroleerd.

Hoe moeten de resultaten van de Recidivemonitor dan wel worden geduid? Deze vraag is lastig te beantwoorden. Er zijn vele factoren die de hoogte van de recidive bepalen en daarvan konden in dit onderzoek slechts enkele worden geneutraliseerd. Een sluitende verklaring voor de nettostijging van het recidiveniveau is dus niet te geven. Maar er is wel iets te zeggen over de richting waarin men het zou kunnen zoeken. Grofweg zijn er vier mogelijkheden:

- 1 De Nederlandse bevolking, of een specifiek deel daarvan, is tussen 1997 en 2004 criminel geworden. Er is meer criminaliteit en dus ook meer recidive. Niet alle recidivepercentages zijn in gelijke mate gestegen, maar dit heeft te maken met een herverdeling van de sancties. Buiten de invloed van de in dit onderzoek gecontroleerde factoren om krijgen 'lichtere' daders tegenwoordig 'zwaardere' sancties opgelegd, en andersom. Daarom stijgen de percentages bij de ene sanctie minder hard dan bij de andere en is er bij de vrijheidsstraffen voor jeugdigen zelfs een afname van de strafrechtelijke recidive te zien.
- 2 Er hebben zich wijzigingen voorgedaan in het optreden van politie, justitie en de organisaties die verantwoordelijk zijn voor de uitvoering van straffen en maatregelen. De invulling van de strafrechtelijke sancties in Nederland is veranderd, wat leidt tot meer en in sommige gevallen juist tot minder recidive. Het kan ook zijn dat soortgelijke sancties tegenwoordig worden toegepast op groepen daders die minder of juist meer gevoelig blijken te zijn voor de afschrikwekkende dan wel resocialiserende werking ervan. Met andere woorden: de effecten van de sancties zijn veranderd omdat zij nu bij andere daders worden ingezet.
- 3 De stijging van de recidive is reëel, maar moet worden toegeschreven aan selectie-effecten waarvoor in dit onderzoek niet kon worden gecorrigeerd. Indien bijvoorbeeld naast de rol van sekse, leeftijd, geboorteland, type delict en delictgeschiedenis een laag opleidingsniveau indicatief zou zijn voor de kans op een nieuw justitiecontact, dan zou een verlaging van het gemiddelde opleidingsniveau de hogere recidivecijfers kunnen verklaren. De daderbevolking is in dat geval niet intrinsiek meer crimineel geworden (zoals onder 1 wordt verondersteld), er zijn simpelweg meer daders vervolgd met een laag opleidingsniveau.

- 4 De stijging van de recidive is niet reëel. Ze moet worden toegeschreven aan registratie-effecten waarvoor in dit onderzoek niet kon worden gecorrigeerd. Er worden niet méér delicten gepleegd, de delicten die er zijn, worden – bijvoorbeeld omdat de capaciteit van de rechterlijke macht is toegenomen – sneller dan voorheen tot een strafzaak gevormd. Dit leidt tot een toename van het aantal strafzaken en tot hogere recidivecijfers zonder dat de herhalingscriminaliteit feitelijk is toegenomen.

Welk van de vier mogelijkheden opgaat, kan op basis van het onderzoek niet worden bepaald. De vier verklaringswijzen – ‘meer criminaliteit’, ‘minder effectief optreden’, ‘niet-gemeten instroomverschillen’ en ‘overige registratie-effecten’ sluiten elkaar niet uit. Wellicht zijn alle voor een deel juist. Per sector kan het beeld bovendien verschillend zijn.

Tbs

In de tbs-sector is na een daling in de jaren '80 en '90 het percentage recidivisten de laatste jaren weer gestegen. Hier ligt het derde verklaringstype het meest voor de hand: er stromen tegenwoordig andere terbeschikkinggestelden uit. De uitkomst van het onderzoek kan worden verklaard door te verwijzen naar mechanismen die de in- en uitstroom van terbeschikkinggestelden hebben bepaald. Kröger en Van Beek (2006) beschrijven hoe de vermaatschappelijking van de geestelijke gezondheidszorg de tbs-populatie heeft veranderd. Naast daders met een persoonlijkheidsstoornis die altijd al de klinieken bevolkten, werden de laatste twintig jaar steeds meer personen met een zware psychische aandoening opgenomen. Zelf hebben we gezien dat de maatregel vanaf het midden van de jaren '80 vrijwel uitsluitend nog wordt opgelegd aan plegers van gewelds- en/of zedenmisdrijven. Deze delinquenten hebben over het algemeen een lagere *base rate* dan personen die alleen vermogensdelicten plegen. Hun a-priori recidivekansen zijn kleiner en dat is een van de redenen waarom de terugval in de sector is gedaald.

Een andere factor van belang is de verhoogde aandacht voor risicomanagement. Waar vroeger de inschatting van het behandelingsucces werd gebaseerd op het klinische oordeel van de behandelaar, vindt de taxatie van het recidivegevaar tegenwoordig op meer gestructureerde wijze plaats. De ontwikkeling van standaardinstrumenten als de HKT-30, de HCR-20 en de PCL-R stelt het veld steeds beter in staat om tijdens de behandeling een reële inschatting te doen van de kans op recidive (Brand & Van Emmerik, 2006; De Vogel, 2005; Philipse, 2005; Hildebrand et al., 2006). De kennis die deze nieuwe benadering heeft opgeleverd leidt tot andere afwegingen bij verlengingsbeslissingen. Men is voorzichtiger geworden om de maatregel te beëindigen. Temeer omdat de sector intussen sterker in de schijnwerpers is komen te staan. De belangstelling voor de prestaties van de inrichtingen is gegroeid. Gevallen van ernstige

recidive worden breed uitgemeten in de landelijke media. Politici besteden er veel aandacht aan en twijfelen openlijk aan de houdbaarheid van het tbs-systeem (RMO, 2006). Bevreesd voor de commotie die bij elk geval van recidive kan ontstaan, zijn behandelaars terughoudender geworden in hun advies aan de rechter. Dat blijkt ook uit de ontwikkeling in de duur van de maatregel en uit de uitstroomaantallen. Terwijl steeds meer daders tbs krijgen opgelegd – het aantal plaatsen in de tbs-inrichtingen is sinds de jaren '80 verdrievoudigd – daalde tot voor kort het aantal uitstromers. Alleen in de laatste vijf jaar is het aantal ex-terbeschikkinggestelden toegenomen (zie paragraaf 4.2.5). Terbeschikkinggestelden zitten steeds langer vast. De ontwikkeling van 'longstay-afdelingen' in een aantal tbs-klinieken getuigt ook van het striktere ontslagbeleid. Een klein deel van de terbeschikkinggestelden reageert niet op behandeling en wordt 'blijvend delictgevaarlijk' geacht (Perik, 2000; Roede, 2002; De Kogel & Verwers, 2003; De Kogel, Verwers & Den Hartogh, 2005). Deze terbeschikkinggestelden kunnen niet eens mislukken, want ze zullen niet in vrijheid worden gesteld.

In de achterliggende decennia heeft zich een verschuiving voorgedaan in de wijze waarop aan de dubbele opdracht van de tbs wordt gewerkt. De nadruk is meer komen te liggen op de bescherming van de maatschappij. Meer bewust van de risico's die er zijn, heeft het veld een periode achter de rug waarin alleen de patiënten met een gunstige prognose werden ontslagen. Als gevolg daarvan is de recidive in de jaren 1994-1998 naar een laagterecord gezakt. Maar het keerpunt is bereikt. De 'voorraad' van perspectiefrijke tbs-ers raakte op. Men moet individuele terbeschikkinggestelden uitzicht bieden op herstel en hen op een gegeven moment laten gaan. De terbeschikkinggestelden uit de laatste cohort hadden *in vergelijking met* hun voorgangers minder goede vooruitzichten, maar hun vooruitzichten waren niet slecht, men kon hen niet blijven vasthouden. Paradoxaal genoeg is dus ook de recente toename van de recidive onder ex-terbeschikkinggestelden het gevolg van de roep om meer veiligheid en niet, althans niet per se, het resultaat van een 'falende behandeling'.

Het gevangeniswezen

De enorme groei die het Nederlandse gevangeniswezen in de afgelopen dertig jaar heeft doorgemaakt, heeft natuurlijk de nodige organisatorische aanpassingen gevergd. Ook in de onderzoeksperiode nam het aantal plaatsen in gevangenis en huizen van bewaring belangrijk toe en kreeg het veld te maken met een aantal bestuurlijke maatregelen. In 2000 leek er even sprake te zijn van overcapaciteit en werd een aantal cellen afgestoten (DJI, 2001). Kort daarna echter nam het aantal heenzendingen weer toe en moest versneld noodcapaciteit worden gerealiseerd. Deze viel voornamelijk toe aan gearresteerde drugkoeriers. Zij werden opgenomen in nieuwe

detentiecentra met een sober regime.⁶⁰ Dekker, Jongejan en Spek (2003) beschrijven hoe de discussie over de 'bolletjesslikkers' de opmaat vormde voor de invoering van het gebruik van meermanscellen. Sinds september 2004 is het samenplaatsen van gedetineerden in alle penitentiaire inrichtingen toegestaan. Tijdens de onderzoeksperiode werd het principe van 'twee-op-één-cel' echter nog niet breed toegepast.

Wat wel een rol kan hebben gespeeld, zijn de bezuinigingen. Vanaf 1999 kreeg het gevangeniswezen te maken met een reeks van bezuinigingsmaatregelen en -voorstellen. In 2000 werd in alle typen inrichtingen het dagprogramma ingekort (Peters, 2005). In 2003 zorgde een wijziging van de penitentiaire maatregel voor een verdere versobering van het regime. Voor gedetineerden in 'algehele gemeenschap' verschoof het minimum van aan te bieden activiteiten van 78 naar 59 uur per week, voor gedetineerden in 'beperkte gemeenschap' van 48 naar 18 uur. Deze veranderingen traden echter pas na juli 2004 in werking. Zij kunnen niet, althans niet in directe zin, van invloed zijn geweest op de uitkomsten van dit onderzoek.

We hebben gezien dat een deel van de schommelingen in het recidivepercentage van ex-gedetineerden kan worden toegeschreven aan verschuivingen in de samenstelling van de populatie. Na de eeuwwisseling namen het aandeel vrouwen, het aandeel first offenders en het aandeel opiumwetovertreeders onder de ex-gedetineerden toe. Misschien is de base rate ook om andere redenen afgenomen en zijn de gecorrigeerde recidivepercentages in 2003 en 2004 daarom lager dan in 2002. De hoge terugval onder ex-gedetineerden is al langere tijd een punt van zorg. In het kader van beleidsprogramma's als *Modernisering Sanctietoepassing*, *Terugdringen Recidive* en *Aanpak Veelplegers* zijn maatregelen afgekondigd om de effectiviteit van de opgelegde straffen te verhogen. De resultaten van deze programma's kunnen, als zij al op landelijk niveau zichtbaar zijn, op zijn vroegst echter pas na 2008 worden gemeten, in termen van recidive althans. De maatregelen zijn namelijk in 2003 en 2004 in gang gezet. Het duurt enige tijd voordat zij zijn uitgemond in concrete interventies voor individuele gedetineerden. De daling van de recidive die figuur 4.12 na 2002 vertoont lijkt dus geen direct gevolg te kunnen zijn van het ingezette beleid. Het is meest waarschijnlijk toch een selectie-effect. Het aantal beschikbare cellen in Nederland is sterk toegenomen, verdachten van misdrijven worden steeds vaker ingesloten. De gedetineerdenbevolking groeit en is tegelijkertijd 'lichter' geworden. Ook buiten de in dit onderzoek gemeten controlevariabelen zoals sekse en het aantal eerdere justitiecontacten om, is de base rate voorafgaand aan de detentie lager komen

60 Sinds 1996 bestaat er in de huizen van bewaring een speciaal regime oorspronkelijk bedoeld voor 'kortverblijvenden': arrestanten, subsidiaire gehechten en gedetineerden uit de zogenoemde overlastcategorie. Het regime kent geen activiteiten gericht op terugkeer naar de samenleving en er is geen avondprogramma.

te liggen en daarom is de recidive, in weerwil van de (dreigende) bezuinigingen, de laatste twee jaar van de onderzoeksperiode gedaald.

Justitiële jeugdinstellingen

De JJI-sector maakte tijdens de onderzoeksperiode globaal dezelfde ontwikkelingen door als het gevangeniswezen. Bestuurlijk gezien waren schaalvergroting, clustervorming en kostenbeheersing de belangrijkste thema's. JJI is de enige sector waarin tussen 1997 en 2004 een daling van de recidive te zien is geweest. Het recidiveniveau schommelt; de afname van het percentage recidivisten is niet constant. Gerekend over de gehele onderzoeksperiode neemt de terugval echter met enkele procenten af. Twee zaken heeft het onderzoek onder ex-pupillen duidelijk gemaakt. Ten eerste is gebleken dat de justitiële jeugdinstellingen steeds meer worden bevolkt met jongeren met een relatief laag recidiverisico. Het aandeel meisjes en het aandeel first offenders stegen. Aan het eind van de onderzoeksperiode nam het aantal civielrechtelijke ondertoezichtgestelden plotseling sterk toe.⁶¹ Een deel van de afname van de recidive moet worden toegeschreven aan de verschuivingen op deze kenmerken. Maar voor een ander deel moet de oorzaak elders worden gezocht, want ook de gecorrigeerde percentages namen af.

Het tweede aspect dat de meting onder JJI-pupillen aan het licht heeft gebracht, is dat alleen onder gestrafte jongeren de recidive is afgenomen. Zou de dalende recidive na een voorlopige hechtenis of na een jeugd detentie het neveneffect kunnen zijn van de verharding van het strafklimaat waarop zoveel auteurs hebben gewezen? Onder druk van veranderende opvattingen over de aanpak van criminaliteit en de mogelijkheden die het nieuwe jeugdstrafrecht daartoe biedt, zouden ook kinderrechtshouders strenger zijn gaan straffen (Van der Laan, 1995; Junger-Tas, 1997; De Haan et al., 1999; SCP, 1999; Boutellier, 2002; Buruma, 2005; Downes & Van Swaaningen, 2006). In Nederland worden steeds meer jongeren vervolgd en ingesloten in een justitiële jeugdinstelling. Waar vroeger werd volstaan met een standje van de politie, wordt nu vervolging ingesteld. Waar vroeger beleidssepsen en geldstraffen werden opgelegd, gaat de rechter nu over tot taakstraffen of jeugddetentie. Er is geen hard bewijs dat er tegenwoordig strenger wordt gestraft. Het zou ook kunnen dat de delicten die worden vervolgd per saldo ernstiger geworden zijn. Maar het feit dat de zwaardere sancties de laatste jaren van de onderzoeksperiode werden opgelegd aan jeugdigen en volwassenen die ten tijde van de uitgangzaak een relatief laag statisch recidiverisico vertonen, is wel een sterke aanwijzing dat er zich in de straftoemeting een zekere inflatie heeft voorgedaan. Steeds meer 'lichtere' jongeren komen in een justitiële jeugdinstelling terecht.

⁶¹ In veel gevallen ging het om crisisplaatsingen in opvanginstellingen. Het 'samenplaatsen' van civiel- en strafrechtelijke jongeren stuitte op veel verzet. Na verschillende onderzoeken werd besloten de twee groepen in aparte voorzieningen onder te brengen (zie o.m. Goderie et al., 2004).

Het gevolg is dat de recidive is gedaald. Niet omdat de aanpak in de opvanginrichtingen nu succesvoller is. Niet omdat dit segment van de daderpopulatie intrinsiek minder crimineel is geworden. Wel omdat de populatie in meer opzichten dan in dit onderzoek werden gemeten, van samenstelling is veranderd en de kansen op recidive bij instroom daarvoor al lager waren. Een groeiend deel van de JJI-populatie valt niet terug, maar zou allicht ook bij een minder zware aanpak niet hebben gerecidiveerd.

De uitkomsten van het landelijke recidiveonderzoek kunnen worden verklaard door te verwijzen naar niet gemeten instroomverschillen. Of het nu gaat om ex-terbeschikkinggestelden, om volwassenen die verbleven in een penitentiaire inrichting of om jongeren die zijn uitgestroomd uit een justitiële jeugdinstelling, er is altijd wel aanleiding om te veronderstellen dat de ene uitstroomcohort niet in alle relevante opzichten vergelijkbaar was met de andere en dat de reden moet zijn dat er zich verschillen in de recidivepercentages voordoen. We hebben in dit onderzoek een aantal relevante factoren kunnen uitsluiten. Er is rekening gehouden met registratie-effecten en er werd gecorrigeerd voor verschuivingen in sekse, leeftijd en geboorteland van de daders, het type delict dat zij pleegden en hun delictgeschiedenis. De rol van deze factoren mag niet worden onderschat. Bij onderzoek in binnen- en buitenland komen zij vrijwel altijd als de beste voorspellers van latere recidive uit de bus. Maar er zijn, vooral bij meer specifieke groepen daders, naast deze grootheden steeds andere factoren denkbaar die van invloed zijn op de hoogte van de recidive. Als deze niet worden meegenomen in het onderzoek, blijft onduidelijk hoezeer zij de uitkomsten hebben bepaald en kan hun invloed dus ook niet worden geneutraliseerd.

Het feit dat er altijd 'overige verschillen' kunnen worden aangevoerd, wil echter *niet* zeggen dat de andere verklaringsmechanismen niet opgaan. Het kan zijn dat de recidivepercentages in de laatste jaren van de onderzoeksperiode hoger uitvallen omdat de criminaliteit, of meer specifiek de jeugdcriminaliteit, is gestegen. Volgens Wittebrood en Nieuwbeerta (2006) is dat overigens niet het geval. Zij analyseerden het verband tussen de omvang van de criminaliteit zoals geschat op basis van slachtoffer-enquêtes en de geregistreerde criminaliteit zoals blijkt uit cijfers over bij de politie aangegeven delicten. Hun conclusie is dat de stijging van de criminaliteitscijfers die we in de achterliggende decennia in Nederland hebben gezien, vooral te danken is aan een betere registratie: de politie maakt vaker een proces-verbaal op. Volgens Wittebrood en Nieuwbeerta is er nauwelijks meer criminaliteit, maar wordt er meer criminaliteit geregistreerd. Dit klinkt plausibel en past ook goed in het beeld van 'net-widening' en de veelgehoorde opvatting dat het strafklimaat in Nederland is verhard. Toch valt er op het eerste deel van de conclusie wel het een en ander af te dingen. De stelling dat de feitelijke criminaliteit niet of

nauwelijks is toegenomen, berust namelijk op de assumptie dat slachtofferenquêtes wél een betrouwbare weergave zouden bieden van de omvang van de criminaliteit. En dat is erg onwaarschijnlijk, want waarom zou uitgerekend deze bron een vaste afspiegeling vormen van de feitelijke criminaliteit en andere bronnen niet? Slachtofferenquêtes hebben evengoed te lijden onder selectieve uitval, non-response en vertekening. In alle schakels van de 'kennisketen' doen zich registratie-effecten voor, ook bij de afname van slachtofferenquêtes (zie box 3.2).

Als we criminaliteit niet goed kunnen meten, weten we dus ook niet of ze is toegenomen. De analyse van Wittebrood en Nieuwbeerta toont aan dat op grond van politie- en justitiecijfers geen generaliserende uitspraken kunnen worden gedaan over de omvang van de criminaliteit. Toegepast op de situatie van de Recidivemonitor betekent dit dat op grond van de algemene cijfers uit dit project geen absolute uitspraken kunnen worden gedaan over de effecten van het strafrechtelijke beleid en de sancties die werden opgelegd. De politiek wil graag weten of het beleid succesvol is geweest, maar de trends die de Recidivemonitor blootlegt, geven daarvan hoogstens een aanwijzing. Het kan zijn dat de recidivepercentages onder jeugdigen met een vorm van vrijheidsstraf zijn afgenomen omdat deze interventies nu effectiever zijn, en het is evengoed denkbaar dat de terugval bij bijna alle andere sancties is gestegen omdat deze nu minder aanslaan. Het punt is dat dit met het hier gehanteerde onderzoeksdesign niet ondubbelzinnig kan worden aangetoond, omdat er altijd alternatieve verklaringen voorhanden zijn. Zolang niet alle relevante factoren gemeten zijn en niet is weerlegd dat de criminaliteit in de betreffende dadergroep autonoom van omvang is veranderd, valt niet te bewijzen dat een verlaging van de recidive het werk is van de sancties die werden opgelegd. Platte recidivecijfers geven niet aan hoe effectief sancties zijn. De Recidivemonitor werkt als een thermometer: het instrument laat zien hoe hoog de recidive is, maar maakt – in zijn huidige vorm – niet duidelijk waarom deze is gestegen of gedaald. Door de ruwe cijfers te corrigeren sluiten we enkele verklaringen uit, maar niet allemaal. Om harde uitspraken te kunnen doen over de effecten van strafrechtelijke interventies zijn meer gegevens nodig, en vooral: een meer gericht onderzoeksdesign, het onderwerp van het volgende hoofdstuk.

5 Recidive als maat voor effectiviteit

De rol van recidivemetingen bij de evaluatie van straffen en maatregelen

De percentages uit hoofdstuk 4 geven aan hoe hoog de recidive is, maar maken niet duidelijk welke invloed de straffen en maatregelen op het niveau van de terugval hebben gehad. In dit hoofdstuk pikken we de draad op die we aan het slot van hoofdstuk 2 hebben laten liggen. We gaan na onder welke omstandigheden op grond van recidivemetingen uitspraken kunnen worden gedaan over de effectiviteit van strafrechtelijke interventies.

Om aan te geven om welke klasse van straffen en maatregelen het gaat zal in paragraaf 5.1 de term *daderprogramma's* worden geïntroduceerd. Een daderprogramma is niet eenzelfde soort interventie als de invoering van een nieuwe wet of een pakket aan beleidsmaatregelen, en in het onderzoek naar de effecten ervan zien we dat terug. De evaluatie van de geïntegreerde aanpak van de georganiseerde criminaliteit in Amsterdam (Nelen, 2008), om maar een voorbeeld te noemen, zal anders verlopen dan de evaluatie van een leerproject voor jongvolwassen geweldplegers. Dat heeft te maken met de herhaalbaarheid van de interventie. Het succes van een daderprogramma kan men 'tellen' en 'vergelijken', omdat het daderprogramma kan worden gezien als een reeks van replicaties van dezelfde ingreep. Men past de interventie steeds opnieuw toe op verschillende individuen en meet de uitkomst in een vergelijkbare maat. Bij het onderzoek naar dit type interventies kan men profiteren van de verworvenheden van de experimentele traditie waarin is nagedacht over het waarborgen van interne validiteit en het doen van statistische inferenties. In andere situaties behoort een experimenteel onderzoeksdesign niet tot de mogelijkheden. Voor brede wetsevaluaties of beleidsdoorlichtingen waarin de 'interventies' die men onderzoekt niet geïsoleerd en herhaald kunnen worden, moet een andere strategie worden bedacht. En een andere oplossing voor het vraagstuk van de causaliteit.

In paragraaf 5.2 gaan we in op de 'evaluatieketen'. Het succes van een daderprogramma kan worden vastgesteld via een recidivemeting, maar deze meting vormt slechts een onderdeel, het sluitstuk, van een meeromvattende reeks van onderzoeksactiviteiten. De evaluatieketen bevat drie schakels: de planevaluatie, de procesevaluatie en de productevaluatie. Pas in de laatste fase draait het om het vaststellen van de uiteindelijke outcome, de recidive onder de deelnemers van het programma. Daaraan voorafgaand is plaats voor het doen van kwalitatief onderzoek waarin de aandacht uitgaat naar de mechanismen die met de interventie in werking worden gezet en die op individueel niveau tot een verlaging van de kans op recidive zouden moeten leiden. In paragraaf 5.3 concentreren we ons op de laatste schakel van de keten: het meten van de recidive. In hoofdstuk 3 kwamen de technische aspecten daarvan al aan bod. In dit hoofdstuk wordt aandacht besteed aan de aanvullende voorwaarden die

nodig zijn om de effectiviteit van de daderprogramma's vast te stellen. Paragraaf 5.4 bevat een samenvatting van het geheel en enkele afsluitende opmerkingen.⁶²

5.1 Daderprogramma's: sancties gericht op recidivereductie

In hoofdstuk 1 hebben we gezien dat er grofweg drie manieren zijn om door middel van strafrechtelijke interventies te werken aan de reductie van recidive: afschrikking, incapacitatie en resocialisatie (zie figuur 1.1). De eerste jaren na de eeuwwisseling lag het accent in Nederland duidelijk op de eerste twee. Het aantal plaatsen in justitiële inrichtingen nam verder toe en het regime binnen de muren versoberde.⁶³ Het nationale Veiligheidsprogramma dat in 2002 van start ging, richtte zich sterk op de problematiek van de 'veelplegers', daders die bij herhaling in aanraking komen met politie en justitie. Het kabinet kondigde extra voorzieningen aan om naast de al bestaande inrichtingen voor de strafrechtelijke opvang van verslaafden (SOV) ook andere dan verslaafde veelplegers 'langere tijd te kunnen vasthouden' (MvJ/BZK, 2002: 32). Met de introductie van inrichtingen voor stelselmatige daders (ISD's) zijn die voorzieningen er ook gekomen. Er zijn nu meer mogelijkheden om criminele carrières voor langere tijd te onderbreken. De gedachte om delinquenten via behandelen begeleidingsprogramma's te resocialiseren, is niet geheel verdwenen, maar wordt in de nota *Naar een veiliger samenleving* wel naar de achtergrond verwezen. Alleen als via 'screening' vast is komen te staan dat de justitiabele nog te beïnvloeden is én de straf lang genoeg duurt om gedragsverandering te bewerkstelligen, is er ruimte voor het uitvoeren van een programma gericht op maatschappelijke reïntegratie. Aan de pogingen tot resocialisatie worden dus restricties verbonden. Het is geen algemeen uitgangspunt meer bij de inrichting van straffen en maatregelen in Nederland.

In het vervolg op het Veiligheidsprogramma – het project genaamd 'Veiligheid begint bij Voorkomen' (2007) – lijkt de insteek anders te zijn. Het kabinet-Balkenende IV wil de recidive onder jeugdige justitiabelen en ex-gedetineerden met 10% verlagen. In een brief aan de Tweede Kamer spreken de minister en de staatssecretaris van 'persoonsgericht ingrijpen', 'gedragsverandering' en 'individuele criminogene factoren' (DSP, 2008). Men wil slimmer gebruikmaken van de voorwaardelijke vrijheidsstraf, men wil de nazorg van ex-gedetineerden beter organiseren en verder bouwen aan een 'evidenced-based' aanbod van interventies. Tijdens de looptijd van het Veiligheidsprogramma werd ondanks de harde toon nog

62 Dit hoofdstuk bevat passages afkomstig uit Wartna, B.S.J. (2005). *Evaluatie van daderprogramma's*. Den Haag: Boom Juridische uitgevers.

63 Zie o.m. Boone & Moerings (2007) voor een bespreking van de 'cellenexplosie'.

wel degelijk geïnvesteerd in randvoorwaarden voor trajecten gericht op maatschappelijke reïntegratie. In het kader van het deelprogramma *Terugdringen Recidive* werden maatregelen genomen om ‘gericht en systematisch’ aan de resocialisatie van justitiabelen te kunnen werken.

Een van de maatregelen was de ontwikkeling van de RISc. In hoofdstuk 3 kwam dit al ter sprake. De RISc is een algemeen diagnose-instrument dat door reclasseringswerkers wordt gebruikt wanneer zij de rechter advies uitbrengen over de aanpak van een verdachte (Vinke et al., 2004). Met behulp van dit instrument wordt het recidivegevaar ingeschat, brengt men de criminogene factoren van een verdachte in kaart en wordt een indruk verkregen van de mogelijkheden om deze te beïnvloeden. Naast de ontwikkeling van de RISc heeft men kwaliteitseisen voor effectieve interventies geformuleerd en werd er een ‘erkenningcommissie’ in het leven geroepen. Deze groep van experts beoordeelt aanvragen van uitvoeringsorganisaties voor interventieprogramma’s bedoeld om de recidive te verlagen. Na advies van de commissie besluit de Minister van Justitie of de plannen worden gefinancierd.

De kwaliteitseisen die de commissie bij de beoordeling van de plannen hanteert, staan opgesomd in box 5.1. In de eisen zien we de basisprincipes van de *what works*-benadering terugkomen. De Nederlandse criteria zijn ontleend aan de interventiepraktijk in Canada en Engeland, waar al langer accreditatiesystemen worden gebruikt (Programma TR, 2005). Wat opvalt is het gebruik van de term ‘gedragsinterventies’ en de eis dat een programma gericht moet zijn op het aanleren van vaardigheden (onder nr. 5 in de lijst). De *what works*-benadering is verankerd in een onderzoekstraditie waarin sociaalpsychologische theorieën van crimineel gedrag de boventoon voeren. Crimineel gedrag is aangeleerd en kan via dezelfde mechanismen waarmee het tot stand is gekomen – modellering, bekrachtiging en cognitieve herstructurering – worden afgeleerd. Volgens sommigen zou uit meta-analyses blijken dat een sociaalpsychologische benadering van crimineel gedrag superieur is aan andere theoretische invalshoeken, zoals sociologische en psychopathologische benaderingen en de rationele keuzetheorie (zie o.m. Gendreau, Little & Goggin, 1996; Bonta, Law & Hanson, 1998). De vraag is echter of een meta-analyse doorslaggevend kan zijn als het gaat om het blootleggen van de wortels van crimineel gedrag. Meta-analyses maken duidelijk wat de belangrijkste voorspellers zijn, maar voorspellers zijn nog geen oorzaken. Het kan zijn dat vaardigheidstrainingen een bewezen werking hebben, maar dat wil nog niet zeggen dat elke andere methode ineffectief is. Het toedienen van een voedingssupplement aan geweldplegers is geen gedragsinterventie, een farmaceutische behandeling is niet gericht op het aanleren van vaardigheden. Niettemin lijkt dit type interventie voldoende theoretisch te zijn onderbouwd. Het is in elk geval denkbaar dat het op termijn een recidive-beperkende werking heeft.

Box 5.1 Kwaliteitscriteria voor interventies gericht op recidivereductie

- 1 *Theoretische onderbouwing*: de gedragsinterventie is gebaseerd op een expliciet veranderingsmodel waarvan de werking wetenschappelijk is aangetoond.
 - 2 *Selectie van justitiabelen*: het type justitiabele waarop de gedragsinterventie zich richt wordt duidelijk gespecificeerd en geselecteerd.
 - 3 *Dynamische criminogene factoren*: de gedragsinterventie is gericht op het veranderen van risicofactoren die samenhangen met het criminele gedrag.
 - 4 *Effectieve (behandel)methoden*: er worden (behandel)methoden toegepast die aantoonbaar effectief of veelbelovend zijn.
 - 5 *Vaardigheden en protectieve factoren*: de aanpak is mede gericht op het leren van praktische, sociale en probleemoplossende vaardigheden.
 - 6 *Fasering, intensiteit en duur*: de intensiteit en duur van de gedragsinterventie sluit aan bij de problematiek van de deelnemer.
 - 7 *Betrokkenheid en motivatie*: betrokkenheid van de deelnemer bij de gedragsinterventie en motivatie voor deelname moeten worden bevorderd en gestimuleerd.
 - 8 *Continuïteit*: er moeten duidelijke verbindingen zijn tussen de gedragsinterventie en de totale begeleiding van de justitiabele.
 - 9 *Interventie-integriteit*: de gedragsinterventie wordt uitgevoerd zoals het bedoeld is.
 - 10 *Evaluatie*: een doorlopende evaluatie geeft inzicht in de effectiviteit van de gedragsinterventie.
-

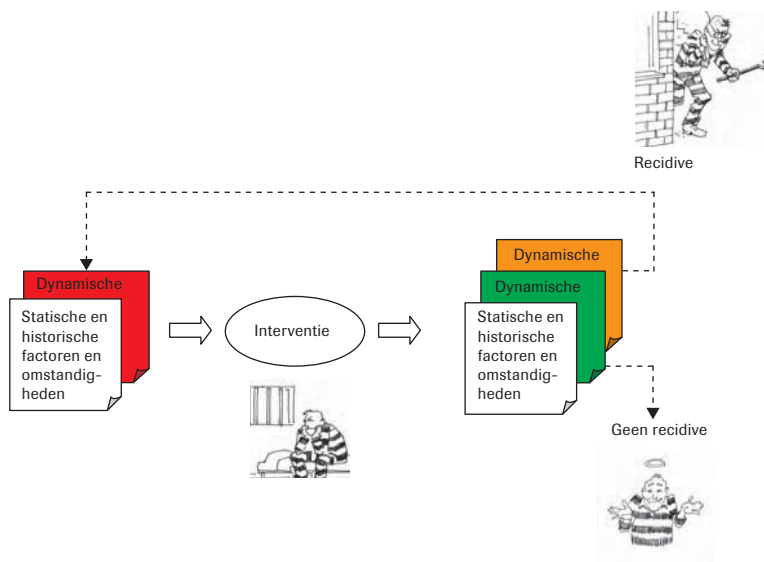
Bron: Erkenningscommissie (2005)

In de zoektocht naar geschikte interventies mag geen enkel theoretisch perspectief worden uitgesloten, tenzij het onvoldoende is uitgewerkt of ondubbelzinnig is weerlegd. In dit hoofdstuk zal de term 'gedragsinterventie' dan ook worden vermeden. Om alle opties open te houden wordt gesproken van 'daderprogramma', een meer omvattende en neutrale term. *Elk initiatief – project, traject, interventie, gedragsinterventie, programma, sanctie, executiemodaliteit – dat zich laat beschrijven als een gestructureerd geheel van voorzieningen en methodische handelingen gericht op speciale preventie via een gerichte aanpak van criminogene factoren van individuele justitiabelen, is een daderprogramma.* In deze klasse van strafrechtelijke interventies zijn we hier primair geïnteresseerd.

Niet alle strafrechtelijke interventies gericht op speciale preventie zijn daderprogramma's. Sancties die bedoeld zijn als afschrikking of uitgaan van pure incapacitatie voldoen niet aan de beschrijving. Er moet sprake zijn van een element van resocialisatie en behandeling. Een arbeidstoeleidingsproject voor langgestraften is een goed voorbeeld van een daderprogramma. Het programma wil de deelnemers behoeden voor een terugval door het vergroten van hun *employability*. Een korte vrijheidsstraf uitgevoerd in een sober regime is geen daderprogramma, omdat er geen

behandeling van criminogene factoren plaatsvindt. Welke criminogene factoren worden aangepakt hangt af van het theoretisch perspectief, de doelgroep van de interventies en de opzet van het programma. Bij veel programma's wordt in feite een keuze gemaakt uit de tien leefgebieden die in het kader van de RISC-screening worden onderscheiden (zie 3.4.1). Een arbeidstoeleidingsproject richt zich voornamelijk op 'opleiding, werk en leren'. Een interventie als ETS-COVA tracht criminele denkpatronen te doorbreken en grijpt aan op cognitieve vaardigheden (Van Poppel & Tackoen, 2004; Tierolf, 2007). De strafrechtelijke opvang van verslaafden (SOV) was een goed voorbeeld van een *multimodaal* daderprogramma. De resocialisatiefase van de SOV bestond niet alleen uit de behandeling van verslavingsproblemen, maar was ook gericht op het verkrijgen van werk, huisvesting en het verbeteren van de financiële situatie (Koeter & Bakker, 2007). Inmiddels is de SOV zoals bekend opgegaan in de ISD-maatregel.

Figuur 5.1 Algemeen interventiemodel voor daderprogramma's gericht op speciale preventie



Figuur 5.1 geeft de belangrijkste formele kenmerken van daderprogramma's nog eens schematisch weer. Daderprogramma's richten zich op factoren of omstandigheden waarvan wordt aangenomen dat zij een beslissende rol spelen in de genese van crimineel gedrag. De interventie grijpt aan op dynamische factoren die vatbaar zijn voor behandeling. Aan statische en historische factoren kan niet worden gesleuteld. Is de behandeling succesvol, dan komt het licht op groen te staan: de omstandigheden van de individuele deelnemer zijn verbeterd. Het tij is ten gunste gekeerd, er is sprake van een geslaagde resocialisatie. Beklijft de behande-

ling niet of onvoldoende, dan is er een kans dat hij terugvalt in crimineel gedrag.

5.2 De evaluatieketen: schakels in het onderzoek naar daderprogramma's

Daderprogramma's komen in vele gedaanten voor. Ze verschillen naar vorm, inhoud, aanpak, duur en doelgroep, maar hebben een gemeenschappelijk doel: de vermindering van recidive. Het succes van een daderprogramma wordt dan ook bepaald door de mate waarin de ex-deelnemers terugvallen in crimineel gedrag. De evaluatie bestaat echter uit meer dan een recidivemeting. We willen niet alleen weten óf de interventie werkt maar ook hoe en waarom. Aan de recidivemeting gaat dus ander onderzoek vooraf. De recidivemeting brengt de uiteindelijke uitkomsten in kaart, maar om die te begrijpen is meer kennis nodig: over de opzet van het programma, over de context waarin het wordt uitgevoerd, over de realisatie van de tussenliggende behandeldoelen en over de processen en mechanismen die met de interventie in gang worden gezet om deze doelen te bereiken.

Tabel 5.1 Fasen van onderzoek bij de evaluatie van daderprogramma's

Soort evaluatie	Fase van uitvoering	Beleidsdoelstelling	Hoofdvragen van onderzoek
Planevaluatie	Vóór invoering van het programma	'Go/no go'-beslissing	Waarom zou de interventie werken? Welke effecten mogen van het programma worden verwacht in termen van de criminogene factoren van de deelnemers? ... in termen van recidivereductie? ... in termen van overige verwachte kosten en baten?
Procesevaluatie	Tijdens of vlak na de invoering van een programma	Vaststelling van (de omgeving van) het programma	Hoe werkt de interventie in de praktijk? Wat zijn (volgens deelnemers en behandelaars) de ideale omstandigheden in termen van bedrijfsprocessen en samenwerkingsrelaties? ... voor de beoogde veranderingen in de leefsituatie van de deelnemers?
Productevaluatie	Enige tijd na de invoering van een programma	Vaststellen van het beleid inzake voortzetting of uitbreiding van het programma	Werkt het? In hoeverre voldoet het programma aan de verwachtingen met betrekking tot de gestelde doelen... ... in termen van de te leveren diensten en producten (output)? ... in termen van verbeteringen in de criminogene factoren van de deelnemers (impact)? ... in termen van recidivereductie (outcome)?

Een recidivemeting is, ook gezien in de tijd, het sluitstuk van een ‘evaluatieketen’, een reeks van activiteiten waarin op systematische wijze informatie over de programma’s wordt verzameld en verwerkt. Tabel 5.1 geeft aan uit welke schakels de evaluatieketen grofweg bestaat. Het onderzoek naar de effecten van daderprogramma’s kent drie fasen: de plan-, de proces- en de productevaluatie.⁶⁴ Alle fasen worden doorlopen, anders blijft een deel van de vragen onbeantwoord. De planevaluatie vindt idealiter plaats vóór de invoering van een programma. In deze fase worden de plannen doorgelicht en wordt zo nauwkeurig mogelijk vastgesteld welke effecten men van de interventie mag verwachten. De procesevaluatie vindt plaats als het daderprogramma van start is gegaan. In deze fase draait het voornamelijk om de vraag of de interventie op correcte wijze wordt uitgevoerd. Is dat het geval, dan kan men overgaan tot de productevaluatie. In deze fase richt de aandacht zich op het vaststellen van de feitelijke effecten.

5.2.1 *Planevaluatie: waarom zou een interventie werken?*

Een planevaluatie handelt om de vraag welke effecten van een (bijstelling van een) programma mogen worden verwacht. Deze fase van onderzoek gaat als het goed is aan de invoering van een programma vooraf. Hierop zijn echter uitzonderingen denkbaar. Ook als het programma al enige tijd draait, is het mogelijk een planevaluatie te doen. De centrale vraag verandert daardoor niet: welke resultaten zal het programma op grond van de plannen – dus op grond van de theoretische veronderstellingen – weten te behalen?

Een planevaluatie moet de beleidsmakers duidelijk maken of het verstandig is om een bepaald daderprogramma te implementeren. In feite is dat natuurlijk ook de opdracht van de eerdergenoemde erkenningscommissie. Het werk van deze commissie mag dan ook een planevaluatie worden genoemd. De commissie bekijkt de plannen die zijn ingediend en geeft daar een oordeel over. Er zijn verschillende manieren om een planevaluatie uit te voeren. Een onmisbaar element is het reconstrueren van de onderliggende ‘programmatheorie’. Daarnaast kan men gebruikmaken van een groep technieken die als kosten-batenanalyse kan worden aangeduid. Op beide benaderingen gaan we kort in.

64 Dezelfde indeling vinden we bij Swanborn (1999), zij het dat deze auteur de planevaluatie ziet als de ontwerpfase van de interventie (of het beleid) en niet op de eerste plaats als onderzoek naar de plannen op zich, zoals hier de opvatting is. We spreken van een evaluatieketen omdat de drie fasen normaal gesproken in een vaste volgorde worden afgewerkt. Soms is het echter nodig verschillende plan- en procesevaluaties uit te voeren alvorens over te gaan tot de productevaluatie, bijvoorbeeld wanneer de opzet van een programma een belangrijke wijziging heeft ondergaan.

Reconstructie van de programmatheorie

Aan elk daderprogramma liggen assumpties ten grondslag; over de doelgroep, over de aard en oorzaken van hun criminele gedrag en over de mechanismen die in werking worden gezet met het pakket aan te plegen interventies. De juistheid van de assumpties bepaalt het succes van het programma. Niet alleen de onderliggende criminologische theorie moet kloppen, ook de uitwerking daarvan is van groot belang. Ook bij de vertaling van de theorie in een interventiestrategie en de omzetting dáárvan in concrete plannen en handelingsvoorschriften worden keuzes gemaakt waarvan de houdbaarheid moet worden onderzocht (cf. Leeuw, 2003).

Met een planevaluatie wordt het geheel van assumpties, de onderliggende 'logica' of 'rationale' van de interventie, in kaart gebracht. Althans, voor zover dat nog niet is gebeurd. De eerste eis voor daderprogramma's die in aanmerking willen komen voor erkenning, is namelijk dat de interventie voldoende is onderbouwd (zie box 5.1). De interventie dient gebaseerd te zijn 'op een expliciet veranderingsmodel waarvan de werking wetenschappelijk is aangetoond'. Als een daderprogramma wordt aangemeld, is – als het goed is – een deel van het werk dus al gedaan. Uit de projectbeschrijving moet blijken welke processen en mechanismen in de ogen van de behandelaars een rol spelen bij het ontstaan van crimineel gedrag, op welke criminogene factoren de interventie aangrijpt en hoe men deze precies tracht te beïnvloeden om verdere recidive te voorkomen.

In de planevaluatie worden de plannen van een daderprogramma maximaal ontvouwd om te bezien of zij een kans van slagen hebben. Alle veronderstellingen – niet alleen van de formele, criminologische theorie maar ook van de uitwerking in de praktijk – worden gearticuleerd en getoetst aan de kennis en inzichten vastgelegd in de wetenschappelijke literatuur. De eis van de commissie dat het veranderingsmodel een bewezen werking moet hebben, is erg streng. Het zou betekenen dat er geen nieuwe interventies bij de commissie kunnen worden aangemeld. Gelukkig wordt in de toelichting van de eis het een en ander gerelativeerd. Indien er nog geen wetenschappelijk bewijs voorhanden is, dient de werking van het daderprogramma 'aannemelijk' te worden gemaakt, door zo veel mogelijk te verwijzen naar relevante wetenschappelijke bevindingen (Programma TR, 2005). Zo blijft er – terecht – ruimte voor nieuwe initiatieven.

Meer in het algemeen draait het in deze eerste fase van de evaluatie om 'aannemelijkheid'. Centraal staat niet de feitelijke effectiviteit van een interventie, maar de *mogelijke* werking. De reconstructie van de programmatheorie is een kwalitatieve analyse van de kansen van een daderprogramma. Het is van belang dat de plannen voor interventies goed worden doorgelicht, zodat het risico op mislukking zo klein mogelijk gehouden wordt. Als de programmatheorie niet klopt, werkt de interventie niet – althans niet zoals die is bedoeld – en is het wellicht beter om van invoering af te zien.

Box 5.2 Aandachtspunten voor de planevaluatie

In de fase van de planevaluatie worden de assumpties die ten grondslag liggen aan de uitvoering van een daderprogramma kritisch onderzocht. Voorbeelden van vragen die bij de reconstructie van de 'programmatheorie' aan de orde kunnen komen, zijn:

- In welke specifieke termen wordt de doelgroep van het daderprogramma omschreven?
 - Op welke criminogene factoren richt het programma zich en hoe zijn deze factoren van invloed op het (crimineel) gedrag van de deelnemers?
 - Waaruit bestaat de behandeling van deze criminogene factoren? Is er sprake van een eenvormig geheel (een module) of bestaat de interventie uit verschillende onderdelen die in wisselende samenstelling worden aangeboden (een 'multimodaal' programma)?
 - In welke setting of context vindt de interventie plaats? Maakt zij onderdeel uit van een traject of van een ander groter geheel? Waarom zou de interventie in deze context werken?
 - Welke (ketens van) gedragsmechanismen moeten ervoor zorgen dat de interventies die men pleegt, leiden tot de gewenste veranderingen en het uitblijven van recidive? Waarom zouden deze mechanismen in werking treden?
 - Zijn er, hoewel wellicht in een andere context, vergelijkbare programma's of behandelmethoden bekend voor de betreffende doelgroep en welke resultaten geven deze interventies te zien?
 - Welke verwachtingen koesteren de direct betrokkenen van het programma? Wat motiveert hen om aan het programma mee te werken? Onder welke omstandigheden werkt het volgens hen het best?
 - Welke recidivereductie mag van dit daderprogramma worden verwacht? Waarop is deze verwachting gebaseerd?
 - Welke positieve en negatieve effecten zijn er naast recidivereductie te voorzien? Op welk niveau zullen deze neveneffecten zich voordoen?
-

Kosten-batenanalyse

Een aanvullende manier om vooraf een inschatting te maken van de kwaliteit van een daderprogramma, is het uitvoeren van financiële analyses. Aandacht voor de kostenaspecten van strafrechtelijke interventies bij evaluatieonderzoek is betrekkelijk nieuw. In de Verenigde Staten en in het Verenigd Koninkrijk is inmiddels de nodige ervaring opgedaan en zijn de gebruikte methodieken beschreven (zie bijvoorbeeld Aos et al., 1999; Dhiri & Bran, 1999). In Nederland ontwikkelde onderzoeksbureau Ecorys een eerste versie van een algemeen model voor kosten-batenanalyses van justitiële interventies (Versantvoort et al., 2005). Het bureau onderscheidt drie varianten van kosten-batenanalyses:

- 1 Kosteneffectiviteitanalyse: deze methode wordt gebruikt als het belangrijkste effect van een interventie (i.c. recidivereductie) wel in een maat, maar niet in een geldwaarde kan worden uitgedrukt.
- 2 Kosten-batenanalyse in engere zin: deze techniek wordt toegepast als de belangrijkste baten van een interventie wel in geldwaarden kunnen worden uitgedrukt.
- 3 Multicriteria-analyse: deze methode geldt als de baten van een interventie uiteenlopen en niet (goed) te kwantificeren zijn.

Bij kosteneffectiviteitanalyses worden de verwachte kosten van daderprogramma's afgezet tegen de verwachte baten in termen van reductie van recidive. Op basis van ervaringen met soortgelijke interventies in binnen- en buitenland schat men het percentage recidivedelicten dat men via het programma hoopt te kunnen voorkomen.⁶⁵ De verhouding tussen dit percentage en de (meer)kosten die per deelnemer worden gemaakt, geeft aan wat de verwachte kosteneffectiviteit van een programma is. Het is de prijs die voor de recidivereductie moet worden betaald en dus een maat voor de prijs-kwaliteitverhouding.

Bij een kosten-batenanalyse in engere zin (KBA) worden ook de effecten van een interventie uitgedrukt in financiële waarden. Als een daderprogramma recidiveverlagend werkt, betekent dit dat de samenleving zich kosten bespaart. Recidivereductie houdt in dat er minder delicten worden gepleegd, dat er minder maatschappelijke schade ontstaat en dat politie en justitie minder uitgaven behoeven te doen. In Nederland heeft het economisch onderzoeksbureau SEO schattingen gemaakt van de maatschappelijke kosten van bepaalde delicten (Groot et al., 2007). Met deze 'delictprijzen' kan worden berekend welk bedrag naar verwachting zal worden terugverdiend op elke euro die het programma kost.

Een multicriteria-analyse is een extensie van een KBA. Ze houdt rekening met het feit dat een daderprogramma tegelijk meerdere doelen dient en dat niet alle effecten van de interventie te kwantificeren zijn. Daderprogramma's zijn gericht op recidivereductie. Dat is het hoofddoel. Maar er zijn ook neveneffecten. Een voorbeeld van een belangrijke, maar vaak vergeten bijwerking is het incapacitatie-effect. Bij sommige sancties is dit het primaire strafdoel (bijvoorbeeld ISD), maar bij andere interventies wordt dit effect nogal eens over het hoofd gezien. Men is geneigd alleen te kijken naar het recidivebeeld na afloop van de behandeling, terwijl ook *tijdens* de interventie vaak een hoeveelheid criminaliteit wordt voorkomen. Personen die gedetineerd zijn of anderszins onder toezicht staan, zijn niet of in elk geval minder in de gelegenheid om delicten te plegen. Meer in het algemeen is het zo dat verschillende soorten 'baten' en 'kosten' een rol kunnen spelen in de besluitvorming rond een programma. Een multicriteria-analyse neemt alle verwachte hoofd- en neveneffecten

⁶⁵ Bij deze prognose mogen de percentages natuurlijk niet zomaar uit de onderzoeksliteratuur worden overgenomen. Men moet rekening houden met de wijze waarop de recidive is berekend en ook rekening houden met de *base rate*, het recidiveniveau dat men normaal gesproken in de doelgroep aantreft.

in ogenschouw, bedoeld en onbedoeld en of zij nu in (geld)waarden uit te drukken zijn of niet. Een brede kijk voorkomt dat de aandacht zich te sterk richt op hetgeen gemakkelijk te meten is.

5.2.2 *Procesevaluatie: hoe werkt de interventie?*

De tweede schakel van de evaluatieketen is de procesevaluatie. In deze fase van onderzoek gaat het niet langer om de plannen, maar om de uitvoering daarvan. Centraal staan nu de processen die zich tussen de betrokken partijen afspelen. Niet alleen op het vlak van de bedrijfsvoering en de samenwerking tussen de betrokken instanties, maar ook (en vooral) op het niveau van de feitelijke beïnvloeding van de deelnemers. De procesevaluatie betreft niet alleen de organisatie, er vindt ook onderzoek plaats naar de gedragsmechanismen die met de interventie op individueel niveau in werking worden gezet. Het is het instrument om te bepalen of het daderprogramma in de praktijk werkt zoals het bedoeld is. Het is overigens denkbaar dat de evaluatie deze fase niet haalt. Als uit een gedegen planevaluatie is gebleken dat de interventie is gebaseerd op een onjuiste theorie of een verkeerde uitwerking daarvan, dan rijst de vraag of het daderprogramma niet beter kan worden stopgezet.

Een procesevaluatie is een intensieve vorm van onderzoek. Door observatie, vraagesprekken en analyse van administratieve gegevens en andere bronnen vormen de onderzoekers zich een oordeel over de wijze waarop het project in de praktijk is gebracht. Centrale vraag is of de plannen worden uitgevoerd zoals ze bedoeld zijn. Is dat het geval – zijn de beschreven randvoorwaarden gerealiseerd en verlopen de processen tussen alle betrokken partijen als beoogd –, dan is er sprake van ‘programma-integriteit’. Wanneer dit punt is bereikt, kan men overgaan tot de fase van de productevaluatie. Wordt een interventie nog niet consistent volgens plan toegepast, dan is het in feite te vroeg om een betrouwbaar beeld te krijgen van de eindresultaten. Eerst moet via een (serie van) procesevaluatie(s) duidelijk worden welke bijstellingen nodig zijn om het programma optimaal te laten draaien.

Een procesevaluatie is niet alleen bedoeld om de integriteit van een daderprogramma vast te stellen. Het geeft ook de gelegenheid om de veronderstelde werking van het programma in detail te bestuderen. Zijn de gedragsmechanismen werkzaam die op grond van de programma-theorie werden verondersteld? Zijn deze mechanismen bij alle deelnemers van kracht, of komen er onvoorziene reacties voor?⁶⁶ De bevindingen van de procesevaluatie vormen een toets voor de veronderstellingen die in

66 Bij de evaluatie van daderprogramma's mag niet alleen worden gekeken naar de deelnemers die het programma volledig hebben afgemaakt. Ook de resultaten van hen die vroegtijdig zijn afgehaakt tellen mee, want anders ontstaat er een vertekend beeld. Over de uitgevallen deelnemers moeten dus gegevens worden bijgehouden. Voor de procesevaluatie zijn deze van speciaal belang. De ervaringen van degenen die afvielen zijn heel instructief. Uit de informatie kan worden afgeleid welk deel van de doelgroep ongevoelig is voor de interventie.

de fase van de planevaluatie zijn geëxpliciteerd. Vooral bij nieuwe interventies is het nodig alle betrokken partijen uitvoerig aan het woord te laten en op te tekenen hoe zij het programma ervaren en waarderen. Bij programma's die sterk lijken op programma's die elders zijn uitgevoerd, is de behoefte aan nieuwe onderzoeksgegevens waarschijnlijk minder groot.

Box 5.3 Aandachtspunten voor de procesevaluatie

Het is onmogelijk om een volledige opsomming te geven van de zaken die bij een procesevaluatie aan de orde dienen te komen. De specifieke onderzoeksvragen die men in deze fase stelt, hangen sterk af van de inhoud, opzet en omvang van het kaderprogramma, de omgeving waarin het programma wordt uitgevoerd en het stadium waarin het zich bevindt. Punten die in elk geval aan de orde dienen te komen, zijn:

- In welke varianten komt de interventie voor? In welke context wordt het uitgevoerd? Maakt het programma deel uit van een traject of een ander geheel, of staat het op zichzelf?
 - Welke financiële, personele en organisatorische middelen worden ingezet bij de uitvoering van het programma?
 - Hoe is de informatievoorziening rond het programma vormgegeven? Worden de juiste gegevens vastgelegd om de voortgang van het project te kunnen volgen?
 - Wordt het programma volgens plan uitgevoerd? In welke opzichten verschilt de uitvoering in de praktijk van de plannen zoals beschreven in de projectopzet?
 - Hoe werkt de interventie in de praktijk? Welke veranderingsmechanismen zijn precies van kracht? Is dit conform de veronderstellingen van de 'programmatheorie'? Wat is volgens de betrokkenen het werkzame bestanddeel in het 'pakket' van interventies die de deelnemers ondergaan?
 - Welke knelpunten doen zich voor bij het screenen, de selectie, behandeling en nazorg van de deelnemers? Welk deel van de doelgroep wordt bereikt? Welke onvermoede selectie-effecten doen zich in de praktijk voor?
 - Met welke uitval van deelnemers heeft men te maken, in welke omstandigheden en om welke redenen haken deelnemers voortijdig af?
 - Zijn er vergelijkbare projecten of op onderdelen vergelijkbare projecten? Hoe verhoudt de uitvoering van het programma zich tot deze projecten?
 - Welke onvoorziene effecten tekenen zich af in de directe en indirecte omgeving van het project?
 - Hoe waarderen de medewerkers en deelnemers het programma? Onder welke omstandigheden werkt de interventie volgens de betrokkenen in de praktijk het best?
 - Welke aanpassingen moeten in de opzet van het programma worden gedaan? Wat hoopt men met deze bijstellingen te bereiken?
-

Ook het tijdstip waarop de procesevaluatie begint hangt af van de levensfase van het daderprogramma. Gaat het om een nieuw programma dat zich nog in een experimenteel stadium bevindt, dan bestaat er doorgaans veel onzekerheid over de toepassing van de interventie en is het goed om de implementatie van het programma vanaf de start intensief te volgen. Bestaat er echter al de nodige ervaringskennis, bijvoorbeeld omdat het programma eerder al elders is beproefd, dan is het wellicht verstandiger om nog even te wachten totdat de nieuwe praktijk zich heeft gevormd. Na enige tijd kan deze dan met een kortdurende procesevaluatie in kaart worden gebracht om vast te stellen of er nog aanpassingen moeten worden gedaan. In de Verenigde Staten is een lijst ontwikkeld waarmee de programma-integriteit kan worden vastgesteld (Van Voorhis & Brown, 1996). De vragenlijst kan worden gebruikt om tekortkomingen in de implementatie van daderprogramma's op te sporen. Lowenkamp en Latessa (2004) laten zien dat de scores op deze *correctional program assessment inventory* (CPAI) samenhangen met de effectiviteit van het programma. In Nederland is er geen vaste methode om een procesevaluatie uit te voeren. Een goed voorbeeld van een dergelijke analyse is het onderzoek van Van Poppel en Tackoen (2004) naar de testfase van ETS-COVA.

5.2.3 *Productevaluatie: werkt de interventie?*

De laatste schakel van de evaluatieketen is de productevaluatie. In deze fase van het onderzoek worden de uitkomsten van de daderprogramma's gemeten. De productevaluatie vindt plaats enige tijd na invoering van het programma, wanneer het min of meer vaste grond onder de voeten heeft gekregen en men in staat is gebleken om de interventie op consistente wijze ten uitvoer te brengen. Is er nog geen sprake van 'programma-integriteit', dan kunnen de resultaten niet goed worden vastgesteld. Zolang het programma nog van doelgroep, inhoud, opzet of setting verandert, meet men immers de effecten van iets wat niet meer bestaat. Overigens geldt voor de productevaluatie hetzelfde als voor de procesevaluatie. Als uit de eerdere fasen van onderzoek blijkt dat het programma geen kans van slagen heeft, kan het programma beter worden stopgezet en is er dus geen aanleiding meer om verder onderzoek te doen.

De uitkomsten van daderprogramma's kunnen worden ingedeeld in drie categorieën: output, impact en outcome. De *output* van een daderprogramma wordt gemeten in termen van de diensten die men met het uitvoeren van de interventie aflevert. Bij een behandeltraject voor verslaafde criminelen is dat bijvoorbeeld het aantal plaatsingen in een speciale inrichting of het aantal aangeboden therapieessies.⁶⁷ De *impact* van een interventie is het effect dat deze diensten – direct of indirect – hebben gehad op het functioneren van de deelnemers. Bij het project voor

67 Bij sommige auteurs behoort het meten van de output nog tot de fase van de procesevaluatie.

verslaafde criminelen telt men dus niet alleen het aantal keren dat de behandelmethodiek is toegepast (output), maar ook het aantal keren dat de behandeling heeft geleid tot regulering van het drugsgebruik (impact). Bij een arbeidstoeleidingsproject voor gedetineerden telt men naast het aantal aangeboden sollicitatietrainingen ook het aantal succesvolle bemiddelingen naar een baan of opleiding. De *outcome* van een interventie is het uiteindelijke resultaat, op langere termijn en in een breder verband. Bij daderprogramma's wordt de outcome vastgesteld via een recidivemeting. Daarnaast kan ook ander onderzoek worden verricht. Het projectplan van een daderprogramma stelt welke criminogene factoren men met het programma wil aanpakken. Beschreven wordt op welke leefgebieden de interventie zich richt, welke concrete doelstellingen er gelden en hoe men deze denkt te bereiken. Onderdeel van de productevaluatie is dat wordt nagegaan of de behandeldoelen worden gehaald. Men inventariseert de diensten die men heeft weten te realiseren (output) en men meet de vorderingen die de deelnemers op de leefgebieden hebben gemaakt (impact). Per deelnemer wordt een voor- en een nameting gehouden: bij de start van het programma wordt de uitgangssituatie in kaart gebracht en bij vertrek uit het project worden de criminogene factoren opnieuw gemeten. Als het goed is, dat wil zeggen als de programma-theorie klopt, vormen de vorderingen van de deelnemers een indicatie van de recidivereductie die op termijn zal worden gerealiseerd (outcome). Omdat men werkt aan 'problemen en kenmerken die het probleemgedrag veroorzaken, bevorderen of in stand houden', zou na een verbetering op één of meer van de behandelde criminogene factoren de kans op recidive moeten zijn gedaald.

Om de impact van het programma valide en betrouwbaar te meten, kan men het beste gebruikmaken van gestandaardiseerde meetinstrumenten die geschikt zijn voor herhaalde afname. De Commissie Testaangelegenheden Nederland (COTAN) geeft een opsomming van de vragenlijsten en psychometrische tests op die op het Nederlandse taalgebied worden gebruikt (Evers et al., 2000). Per instrument wordt beschreven welk concept het instrument meet, op welke populatie het zich richt, hoe het instrument moet worden afgenomen en welke validiteit en betrouwbaarheid het heeft. De instrumenten hebben betrekking op uiteenlopende onderwerpen. Er zijn vragenlijsten over arbeidsbeleving, over cognitieve vaardigheden en over attitudes ten aanzien van het eigen lichaam. Er zijn algemene persoonlijkheids- en intelligentietests en er zijn screeningsinstrumenten voor psychiatrische problematiek. De beperking van veel vragenlijsten is dat zij alleen door getrainde interviewers kunnen worden afgenomen. Bovendien zijn de meeste vragenlijsten nog niet op justitiabelen getest. We weten dus niet hoe goed zij de concepten in deze specifieke populatie meten.

Met de ontwikkeling van de RISC en andere instrumenten voor screening en risicotaxatie komt daar echter verandering in. Zoals we al zagen in

hoofdstuk 2 wordt in verschillende sectoren van ons strafrechtstelsel gewerkt aan de totstandkoming van standaardvragenlijsten, instrumenten waarmee de criminogene factoren van verschillende typen daders kunnen worden gescoord. De meeste instrumenten zijn nog niet in extenso getest. Het onderzoek naar de betrouwbaarheid en de validiteit van de vragenlijsten is in volle gang. Ook de voorspellende waarde van de gemeten scores wordt onderzocht. De meeste instrumenten zijn in aanleg geschikt voor herhaalde afname. Ze kunnen dus worden gebruikt om de vorderingen van de deelnemers tijdens de behandeling bij te houden. Overigens moet de impact van het programma niet alleen bij vertrek worden vastgesteld. Het is tevens van belang te weten of de effecten na enige tijd nog zichtbaar zijn. Idealiter beschikken de projecten daarom over informatie die betrekking heeft op de periode na afloop van het programma. Ook deze follow-up-gegevens moeten systematisch worden vastgelegd.

Het monitoren van de realisatie van de individuele behandeldoelen levert indicaties op van het vermoedelijke succes van het programma. De echte test vindt plaats wanneer wordt nagegaan welk deel van ex-deelnemers opnieuw in de fout is gegaan. Met de Recidivemonitor heeft het WODC een methodiek ontwikkeld om de terugval van justitiabelen vast te stellen. In hoofdstuk 3 werd de werkwijze van de Recidivemonitor beschreven. Juist omdat alle daderprogramma's mikken op speciale preventie is het van belang dat het uiteindelijke resultaat op eenduidige wijze wordt gemeten. Verschillen in uitkomsten kunnen dan niet worden toegeschreven aan verschillen in de gehanteerde methode van onderzoek. In de volgende paragraaf zal uitvoerig worden ingegaan op de vraag wat een goede recidivemeting tot een goede effectmeting maakt. Hier wordt nog gewezen op de mogelijkheid om ook in de fase van de productevaluatie een kosten-batenanalyse te doen. In een planevaluatie worden de financiële analyses gebaseerd op de verwachte kosten en de verwachte baten. Bij de productevaluatie gaat het om het vaststellen van het gerealiseerde rendement, dat is de verhouding tussen de werkelijk gemaakte kosten en de feitelijke reductie van de recidive in de onderzoeksgroep. Voor het overige zijn bij *ex post* kosten-batenanalyses dezelfde onderscheidingen van kracht. Net als bij de analyses vooraf dient bij de productevaluatie rekening gehouden te worden met neveneffecten, ook als deze niet in financiële waarden kunnen worden uitgedrukt. Een daderprogramma dat buiten de muren van de gevangenis wordt uitgevoerd, levert misschien minder recidive op, maar kan tegelijkertijd leiden tot een toename van het gevoel van onveiligheid bij omwonenden van het project. Beide aspecten zijn van belang in de uiteindelijke beoordeling van het programma.

5.3 Basisvoorwaarden voor een goede effectmeting

Ook al wordt ze voorafgegaan door ander onderzoek, de recidivemeting blijft – gelet op het hoofddoel van de interventies waar het hier om gaat – het culminatiepunt van de evaluatieketen. Met de recidivemeting doen daderprogramma's als het ware hun eindexamen. Programma's die goed zijn onderbouwd en waarvan in eerdere fasen van de evaluatie is komen vast te staan dat ze goed zijn uitgewerkt, houden een zekere belofte in. De recidivemeting maakt duidelijk of de verwachtingen worden waargemaakt. De eindtoets zelf moet wel aan een aantal voorwaarden voldoen. In deze paragraaf gaan we nader in op de vraag onder welke omstandigheden recidivemetingen kunnen bijdragen aan het onderzoek naar de effectiviteit van daderprogramma's. Natuurlijk moet de recidive op zorgvuldige wijze worden gemeten, maar aan de technische aspecten van de meting zelf werd al in hoofdstuk 3 aandacht besteed. Hier gaat het om de vraag hoe het onderzoek meer in het algemeen moet zijn vormgegeven om op grond van recidivecijfers uitspraken te doen over de effecten van daderprogramma's. Drie thema's zijn daarbij van belang: afstand, reliëf en tijd.

Afstand

Het spreekt vanzelf dat het recidiveonderzoek aan alle gangbare methodologische eisen moet voldoen. Dit betekent dat de gevolgde aanpak controleerbaar is en dat het onderzoek geldige en betrouwbare uitspraken oplevert. De wijze van gegevensverzameling en -verwerking moet duidelijk zijn omschreven. Uit de beschrijving van de opzet van het onderzoek blijkt dat men meet wat men behoort te meten en bij de bespreking van de resultaten is navolgbaar hoe de conclusies van het onderzoek tot stand zijn gekomen. Herhaling van het onderzoek zou hetzelfde resultaat moeten opleveren. Zorgvuldigheid betekent ook dat men de beperkingen van de gehanteerde methode kent en aangeeft welke alternatieven er eventueel zijn. In hoofdstuk 3 hebben we kunnen zien dat de gekozen operationalisering (van het domein, het bereik, de tel-eenheid en de termijn van de verzamelde gegevens) bij het meten van recidive bepalend zijn voor de hoogte van de cijfers. Het is daarom van groot belang dat de werkwijze waarvoor men kiest transparant is en consistent wordt toegepast. Binnen het onderzoek moeten steeds dezelfde operationalisering worden gebruikt, anders zijn de uitkomsten niet betrouwbaar. Een zorgvuldige recidivemeting vereist expertise, maar ook een zekere distantie. Daderprogramma's maken onderdeel uit van de strafrechtelijke keten. De uitvoerders zijn veelal politie-, justitie- of reclasseringsmedewerkers die uit hoofde van hun functie in de gelegenheid zijn om van individuele personen justitiegegevens op te vragen. In principe is de projectleiding van een daderprogramma in staat om zelf een recidivemeting uit te voeren. Toch is het niet verstandig om dit deel van het

onderzoek te laten doen door een partij die betrokken is bij de uitvoering. Men kan beter een externe instantie inschakelen. Het gevaar van belangenverstrengeling is te groot. Het voorbeeld in hoofdstuk 1 van de recidivemeting onder ex-pupillen van jeugdinstelling Den Engh illustreert wat er kan gebeuren wanneer het onderzoek in eigen beheer wordt uitgevoerd. Beleidsdirecties, methodiekontwikkelaars, leidinggevend en medewerkers van uitvoeringsorganisaties en toeleveringsbedrijven hebben allen baat bij lage recidivecijfers. Zelfs indien men gebruikmaakt van de meest zorgvuldige, best gedocumenteerde meetmethode, zal men de schijn van partijdigheid niet geheel kunnen vermijden indien men de eindresultaten van het daderprogramma zelf meet. Dit deel van de productevaluatie kan dus beter in handen worden gegeven van een onafhankelijk onderzoeksbureau, zeker als het om de definitieve meting gaat.

Ook voor het meten van de tussenliggende programmadoelen is het aan te bevelen externe waarnemers in te schakelen, hoewel de vraag is of dit altijd haalbaar zal zijn. Zelfs als men gebruikmaakt van gevalideerde meetinstrumenten, kan niet worden uitgesloten dat de medewerkers de vorderingen van de deelnemers te optimistisch inschalen. Ze beoordelen dan immers hun eigen werk. Men zou kunnen stellen dat dit geen bezwaar is, omdat uiteindelijk de recidivecijfers wel zullen uitwijzen of de interventie succesvol is geweest of niet. Toch is het van belang dat de impact van het programma betrouwbaar wordt gemeten. Niet alleen omdat er anders een bedrieglijk beeld zou ontstaan van de merites van een project, maar ook om te voorkomen dat bij het recidiveonderzoek valse verbanden aan het licht komen tussen de gemeten criminogene factoren en de recidive. Onjuiste voortgangsscores kunnen er toe leiden dat de onderliggende criminologische theorie ten onrechte geen steun krijgt. Uiteindelijk heeft iedereen er dus baat bij dat de realisatie van de programmadoelen nauwkeurig wordt gemeten. Ook hiervoor is overzicht nodig en is een zekere distantie op zijn plaats. Het ideaal is dus dat de vorderingen van de deelnemers worden vastgesteld door anderen dan de behandelaars.

Reliëf

De uitkomsten van een daderprogramma kunnen nog zo goed worden gemeten, de waarde van het programma kan pas worden bepaald indien zij wordt afgezet tegen de resultaten in een controlegroep. Alleen door een vergelijking met daders die de bewuste interventie *niet* hebben ondergaan, kan worden zekergesteld dat de interventie effect heeft gehad. De eis dat er sprake moet zijn van een controlegroep, is onderdeel van de experimentele wetenschapstraditie. Binnen deze traditie bestaan strikte opvattingen over oorzaak en gevolg-relaties. Een interventie kan alleen een oorzaak zijn van een verschijnsel (hier: de reductie van recidive) als zij er in de tijd aan voorafgaat en er sprake is van een logisch én empirisch verband waar geen alternatieve verklaringen voor te vinden zijn. In hoofdstuk 2 werd het begrip 'interne validiteit' geïntroduceerd. Dit

is de zekerheid waarmee een geobserveerd effect aan de experimentele conditie – in ons geval deelname aan een daderprogramma – kan worden toegeschreven. In hun standaardwerk over quasi-experimentele designs sommen Cook en Campbell (1979) factoren op die de interne validiteit van causale uitspraken in de weg kunnen staan. Twee bekende hinderpalen zijn ‘rijping’ en ‘selectie’. Als een groep na een interventie de gewenste reactie vertoont (geen recidive), hoeft dat niet per se door de interventie te komen. Het kan zijn dat de deelnemers intussen wijzer zijn geworden (rijping) en vanzelf zijn gestopt. Veel mensen die in aanraking komen met justitie, houden het bij één of twee contacten. Het kan ook zijn dat het programma alleen diegenen aantrekt die hun best doen om recidivevrij te blijven (selectie). Misschien waren de deelnemers daar ook zonder de interventie wel in geslaagd. Door te werken met een controlegroep kunnen deze en andere ‘bedreigingen van de interne validiteit’ worden tegengegaan en zijn er hardere uitspraken mogelijk over de effecten van een programma.

In de *what works*-literatuur worden de resultaten van afzonderlijke productevaluaties door middel van meta-analyses op een rij gezet. In hoofdstuk 2 hebben we uitvoerig bij deze methode stilgestaan. Niet alle evaluaties tellen even zwaar. De betere studies krijgen een groter gewicht. Sherman (1997) ontwierp een schaal waarmee de kwaliteit van de onderzoeksopzet van productevaluaties kan worden bepaald. Deze veelgebruikte Maryland *Scientific Methods Scale* (SMS) kent normaal gesproken vijf niveaus (zie box 5.4). Hoe hoger het niveau, des te beter is het design en des te ‘harder’ zijn de uitspraken die op basis van de onderzoeksresultaten kunnen worden gedaan. Veel meta-analisten kiezen ervoor om alleen studies van niveau 3 of hoger te bespreken. Het design dat hoort bij de eerste twee niveaus wordt niet sterk genoeg geacht, omdat een vergelijkingsgroep ontbreekt en dus kan niet worden bepaald of een effect toe te schrijven is aan de interventie.

Box 5.4 SMS-criteria voor recidivestudies

De Maryland Scientific Methods Scale (SMS) deelt wetenschappelijke studies in aan de hand van de kwaliteit van het onderzoeksdesign. Hoe hoger de score op de schaal, hoe hoger de interne validiteit van het onderzoek (Sherman, 1997; Farrington et al., 2002.) De schaal is ontworpen voor evaluatiestudies van verschillende soorten preventieprojecten. De oorspronkelijke criteria luiden als volgt:

- 1 Studies waarin de aanwezigheid van een programma wordt gecorrigeerd aan de totale omvang van de criminaliteit in een groep.
- 2 Studies waarin per deelnemer voor en na het programma metingen worden gedaan van de hoeveelheid criminaliteit.

- 3 Studies waarbij per deelnemer voor en na het programma metingen worden gedaan en de hoeveelheid criminaliteit wordt vergeleken met die van vergelijkbare justitiabelen (de controlegroep).
- 4 Studies waarbij per deelnemer voor en na het programma metingen worden gedaan en de hoeveelheid criminaliteit wordt vergeleken met die van vergelijkbare justitiabelen waarbij wordt gecontroleerd voor andere variabelen die van invloed zijn op de uitkomstmaat.
- 5 Studies waarbij de hoeveelheid criminaliteit per deelnemer voor en na het programma wordt vergeleken met die van een controlegroep en waarbij de groepen door random toewijzing tot stand zijn gekomen.

Bij recidiveonderzoek gaat de indeling niet helemaal op, omdat het niet per se nodig is eerst een voormeting te doen. Dit leidt tot een aanpassing van de criteria. Wartna en Alberda scoorden het design van de Nederlandse recidivestudies en onderscheidde de volgende vier niveaus. Een vergelijkbare indeling is te vinden bij Van Noije en Wittebrood (2008).

- 1 Recidivestudies zonder controlegroep, studies met een controlegroep die bestaat uit de afvallers van het programma of studies waarbij de recidive in de controlegroep op een andere wijze werd gemeten.
 - 2 Studies waarbij de recidive in de onderzoeksgroep wordt vergeleken met die van vergelijkbare justitiabelen die het programma niet hebben gevolgd.
 - 3 Studies waarbij de recidive in de onderzoeksgroep wordt vergeleken met die van een andere groep en waarbij sprake is van statistische controle van de invloed van eventuele instroomverschillen.
 - 4 Studies waarbij de recidive in de onderzoeksgroep wordt vergeleken met die van een andere groep en waarbij de kandidaten willekeurig aan een van beide groepen zijn toegewezen.
-

Vanaf niveau 3 van de SMS is er wel sprake van een controlegroep. Het recidivepercentage in de groep die deelnam aan het daderprogramma wordt afgezet tegen het percentage in een normgroep. Dat kan de totale populatie zijn – mits daar cijfers van zijn, natuurlijk – maar het kan ook gaan om een meer specifieke groep zoals ‘gedetineerden’ of ‘langgestrafte gedetineerden’. De vergelijking met een base rate geeft reliëf aan het resultaat in de onderzoeksgroep. Het zwakke punt van studies op niveau 3 is dat men niet weet hoe goed de vergelijking is. Het is niet uitgesloten dat men selectie-effecten meet. Een arbeidstoeleidingsproject voor langgestrafte trekt misschien deelnemers aan die een gunstige selectie vormen uit de totale groep van langgestrafte gedetineerden. Als het recidivepercentage in de deelnemersgroep lager uitvalt dan in de controlegroep, kan dat de oorzaak zijn.

Bij studies van SMS-niveau 4 tracht men de invloed van instroomverschillen te neutraliseren door de leden van de experimentele en de controlegroep zo goed mogelijk te *matchen*. De matching vindt plaats op

kenmerken waarvan bekend is dat ze samenhangen met de uitkomstmaat. In het geval van recidiveonderzoek zal men bijvoorbeeld nagaan hoe de leeftijdsopbouw van de deelnemersgroep is, de sekseverhouding en – niet onbelangrijk – de verdeling van het aantal eerdere politie- of justitie-contacten. De controlegroep moet op deze punten vergelijkbaar worden gemaakt. Dat kan op verschillende manieren: door middel van statistische controle (zie 2.2 en 3.5) of door tegenhangers te zoeken die gelijkwaardig zijn. Bij dit laatste is onderscheid mogelijk tussen matching op individueel en op groepsniveau. Bij matches op individueel niveau zoekt men voor elke aspirant-deelnemer een tegenhanger met precies dezelfde achtergronden. Hoe groter het aantal kenmerken waarop men matcht, hoe lastiger dit is. Bij matching op groepsniveau streeft men naar een gelijke verdeling op de kenmerken.⁶⁸ De beperking van niveau 4-studies is dat hoe goed men de twee groepen ook op elkaar afstemt, het onduidelijk blijft of *alle* instroomverschillen zijn geneutraliseerd. Ook studies van niveau 4 leveren dus geen zekerheid op. Maar als er veel bekend is over de factoren die buiten de interventie om van invloed zijn op de recidivekans van de doelgroep, komt men met dit type design echter heel ver. Studies van niveau 5 worden *randomized clinical trials* of *randomized controlled trials* (rct's) genoemd. De kandidaten voor een daderprogramma worden willekeurig toegewezen aan groepen die de interventie wel respectievelijk niet zullen ondergaan. Random toewijzing van kandidaten geeft de best mogelijke garantie op de vergelijkbaarheid van de onderzoeks- en controlegroep, omdat mag worden aangenomen dat de verschillen die er zijn gelijkelijk over de groepen zijn verdeeld. Het meetresultaat kan in zo'n geval probleemloos aan de interventie worden toegeschreven. In hoofdstuk 2 werd al gememoreerd dat een rct alom wordt beschouwd als de 'gouden standaard' van het evaluatieonderzoek. Niet alleen binnen de criminologie maar in alle disciplines waarin men de effecten van interventies wil blootleggen. In de praktijk blijkt dit type studie vaak onhaalbaar.⁶⁹ Willekeurige toewijzing van daders stuit vaak op ethische bezwaren of wordt in de weg gezeten door de bestaande wet- en regelgeving. Het is juridisch gezien onaanvaardbaar om zware misdadigers die normaal gesproken een zware straf krijgen opgelegd, omwille van een experiment op te nemen in een behandelprogramma met relatief veel vrijheden. En als dit al lukt, is het bovendien de vraag hoe de kandidaten die worden toegewezen aan de controleconditie zullen reageren op de gemiste kans om in een soepeler regime terecht te komen. Een *double blind*-studie zoals in de medische wetenschap is op het terrein van het

68 Een manier om dit te bereiken is propensity score matching (Rosenbaum & Rubin, 1983). Op basis van de achtergrondkenmerken wordt de kans berekend dat bepaalde individuen tot de experimentele conditie behoren. De tegenhangers zijn dan degenen die de dezelfde voorspelde kans hadden, maar die de sanctie niet hebben ondergaan.

69 Volgens Farrington & Welsh (2006) zijn er tussen 1957 en 2004 op het terrein van de strafrechttoepassing wereldwijd 122 rct's uitgevoerd met een onderzoeks-aantal van 100 personen of meer. De meeste vonden plaats in de VS.

strafrecht niet goed mogelijk. Behandelaars en justitiabelen weten doorgaans waaruit de interventie bestaat. Het kan zijn dat het eindresultaat daardoor wordt beïnvloed (Cook & Campbell, 1997).⁷⁰

Uiteindelijk blijft dus ook op niveau 5 van de SMS-schaal absolute zekerheid uit. Om deze reden stelt Weisburd (2004) dat rct geen gouden standaard is, maar een 'best practice' waar evengoed het een en ander aan mankeert. Overigens zijn bij de SMS-criteria zelf ook wel enige kanttekeningen te plaatsen. De schaal is niet erg fijn afgesteld. Studies zijn niet altijd even gemakkelijk in te delen en binnen één enkel niveau doen zich grote kwaliteitsverschillen voor. De vraag is ook of een hogere score altijd wijst op een beter design. Als er gebrekkig wordt gematcht, is men immers niet altijd beter af dan als men een goed gelijkende normgroep kiest. Het probleem van de SMS is dat zij studies rangschikt louter op grond van de wijze waarop de vergelijking met de controlegroep wordt gemaakt, terwijl ook van belang is *waarmee* de experimentele groep wordt vergeleken. Een goede studie op niveau 3 kan daarom beter zijn dan een 'slechte 4'.

Wartna en Alberda deden onderzoek naar de kwaliteit van de Nederlandse en Vlaamse effectstudies. In box 5.5 is een overzicht opgenomen van de SMS-scores van recidivemetingen van 103 justitiële interventieprojecten. Geen enkele recidivestudie maakte gebruik van random toewijzing en in meer dan 40% van de gevallen ontbrak het zelfs aan een vergelijkingsgroep. Duidelijk is dat het evaluatieonderzoek in de Lage Landen wel een kwaliteitsimpuls kan gebruiken. De vraag is echter of dit simpelweg betekent dat er meer niveau 5-studies moeten komen. Het is misschien praktischer eerst andere opties na te gaan. In Nederland zijn met de Recidivemonitor nieuwe mogelijkheden ontstaan om vergelijkend onderzoek te doen. Met dit project is er een enorm reservoir aan vergelijkingsmateriaal gecreëerd. Een evaluatiestudie van niveau 2 – dus zonder controlegroep – kan alsnog naar niveau 3, mits de recidive op dezelfde wijze is gemeten als beschreven in hoofdstuk 3. In principe is het zelfs mogelijk dat een simpele effectstudie promoveert naar niveau 4 van de SMS. Daarvoor moet naast het geobserveerde recidivepercentage in de groep de verdeling bekend zijn op de achtergrondkenmerken die zijn opgenomen in de eerder besproken voorspellingsmodellen. Vooralnog hebben we het dan over de sekse, de (start)leeftijd en het geboorteland van de daders, het gepleegde delict en het aantal eerdere justitiecontacten (zie bijlage 6). Zijn deze moderatorvariabelen bekend, dan kan bij de vergelijking met een controlegroep het geobserveerde recidivepercentage voor een aantal belangrijke 'instroomverschillen' worden gecorrigeerd en realiseert men post hoc alsnog een 'quasi-experimenteel design met statistische controle'.

70 In het onderzoek van Zaalberg (2007) was wel sprake van een double blind gerandomiseerd design. Hij onderzocht het effect van het toedienen van voedingssupplementen op het voorkomen van regelschendend gedrag bij gedetineerden.

Box 5.5 De kwaliteit van Nederlandse effectstudies van justitiële interventieprojecten

Tabel 5.2 toont de voorlopige resultaten van een inventarisatie van de methodologische kwaliteit van 103 recidivestudies die tot in 2008 op het Nederlandse taalgebied zijn gepubliceerd. Het complete overzicht is opgenomen in bijlage 7. De oudste studie dateert van 1969. Om de kwaliteit van de gebruikte onderzoeksdesigns te bepalen is gebruikgemaakt van de Scientific Methods Scale (SMS), ontworpen door Sherman (1997) en besproken door Farrington et al. (2002). De SMS scoort de interne validiteit. Hoe hoger de score, des te meer mogelijkheden zijn er om op basis van het onderzoek uitspraken te doen over de effecten van onderzochte sancties. Die mogelijkheden zijn het grootst indien er bij het onderzoek gebruik is gemaakt van random toewijzing van kandidaten aan de behandel- en de controlegroep (niveau 5).

Tabel 5.2 Kwaliteit van het onderzoeksdesign; naar jaar van publicatie

Jaar	SMS-score				Totaal
	2	3	4	5	
≤ 1979	5	3	6	0	14
1980-1989	3	4	2	0	9
1990-1999	23	12	11	0	46
≥ 2000	13	14	7	0	34
Totaal	44	33	26	0	103

In Nederland en België zijn, als het gaat om de evaluatie van strafrechtelijke interventies met behulp van recidivecijfers, geen voorbeelden van randomized clinical trials (rct's) bekend. In ruim 40% van de studies werd geen controlegroep opgenomen. Deze studies hebben score 2. De scores 3 en 4 horen bij het quasi-experimentele onderzoek. Dit is onderzoek waarbij geen sprake is van random toewijzing, maar waarbij wel gewerkt wordt met controlegroepen. Ongeveer een kwart van de studies heeft een SMS-score van 4. Bij studies van niveau 4 wordt een vorm van statistische controle gebruikt om rekening te houden met instroomverschillen tussen deelnemers en leden van de controlegroep.

Tijd

Een goede effectmeting kost tijd. Er is geduld nodig om op grond van recidivecijfers uitspraken te doen over de effecten van een daderprogramma. Bij een te vroege start van het onderzoek krijgt men geen goed beeld van de resultaten van het project. Daar is een aantal redenen voor te geven. Op de eerste plaats kent elk daderprogramma aanloopproblemen. Het duurt even voordat het draait zoals het hoort. Eerder spraken we in dit verband van 'programma-integriteit'. Van programma-integriteit is sprake als aan

alle randvoorwaarden is voldaan en de interventie of reeks van interventies consistent en conform de projectbeschrijving wordt uitgevoerd (zie paragraaf 5.2.2). Het moment waarop dit punt wordt bereikt, zal per programma verschillen; dit vergt een kwalitatief oordeel. In de Verenigde Staten is zoals aangegeven een vragenlijst ontwikkeld die daarbij behulpzaam kan zijn (Van Voorhis & Brown, 1996). In Nederland zijn nog geen criteria ontwikkeld voor het vaststellen van de programma-integriteit. Toch is het voor alle betrokken partijen van belang het startmoment van de productevaluatie zorgvuldig te kiezen. Politici, beleidsmakers, ontwikkelaars en uitvoerders willen natuurlijk graag zo snel mogelijk weten hoe het programma het ervan afbrengt. Zolang de interventie nog geen vaste vormen heeft aangenomen, heeft het echter weinig zin en is het mogelijk zelfs contraproductief om een recidivemeting te starten. Als men te vroeg begint en de resultaten vallen tegen, dan wordt het programma misschien wel ten onrechte stopgezet.

De tweede reden om te wachten met de recidivemeting heeft te maken met statistiek. De *what works*-literatuur wijst uit dat zelfs de meest effectieve interventies slechts een bescheiden reductie in de recidive weten te bewerkstelligen. Bij succesvolle daderprogramma's valt het percentage recidivisten in de onderzoeksgroep normaal gesproken 5 tot 10 procentpunten lager uit dan in de controlegroep (zie o.m. McGuire, 2000). Om effecten van een dergelijke omvang te kunnen meten zijn veel respondenten nodig.⁷¹ Zijn de steekproeven klein, dan kan niet worden uitgesloten dat de verschillen die men tussen de experimentele en de controlegroep aantreft toevalsfluctuaties zijn. Een verschil van 5 à 10% is bij een gering aantal respondenten doorgaans niet statistisch significant.

Tabel 5.3 geeft aan welke aantallen respondenten zijn vereist om het bestaan van een absoluut verschil van een bepaalde omvang aan te kunnen tonen. Ze zijn berekend met behulp van het softwarepakket Stata (2007). De aantallen gelden zowel voor de onderzoeks- als voor de controlegroep. De omvang van de benodigde steekproeven is afhankelijk van een drietal parameters: de base rate, de power en het significantieniveau. De base rate is het recidivepercentage in de controlegroep. Het aantal benodigde respondenten is het grootst als de base rate in de buurt van 50% ligt.⁷² De power is de kans om een werkelijk verschil in recidive te vinden. Deze kans moet zo hoog mogelijk zijn. Een power van 80% is de standaard in medisch en psychologisch onderzoek. Gelet op de doorgaans kleine steekproeven zou voor de criminologie in Nederland 70% een meer pragmatische keuze zijn. Het significantieniveau is de kans dat men concludeert dat er een verschil is, terwijl het in feite om een toevalsfluc-

71 Cohen (1988) ontwikkelde een maat voor de effectgrootte. Bij een base rate van 40% en een absoluut verschil van 15%, bedraagt de effectgrootte w 0,24. Cohen noemt het effect 'klein' als w 0,1 bedraagt, 'medium' als w 0,3 bedraagt en 'groot' als w 0,5 is.

72 Als men het percentage niet kent, kan men dus het beste het aantal respondenten nemen dat hoort bij een base rate van 50%. Dat is de meest conservatieve keuze.

tuatie gaat. Deze ‘type 1-fout’ of α , is hier op 5% gesteld. De aantallen in de tabel zijn netto aantallen. Er is geen rekening gehouden met voortijdige uitval van deelnemers.

Tabel 5.3 Minimaal benodigde steekproefomvang per groep bij diverse effectgroottes ($\alpha=0,05$)

Abs. verschil met experimentele groep	Power = 70%					Power = 80%				
	verwachte recidivekans in controlegroep (base rate)					verwachte recidivekans in controlegroep (base rate)				
	40%	50%	60%	70%	80%	40%	50%	60%	70%	80%
15%	105	115	117	111	96	133	147	150	141	122
12%	163	177	179	167	142	208	227	229	214	181
10%	234	252	252	234	196	300	325	325	300	251
9%	288	310	309	284	237	371	400	398	366	304
8%	363	390	387	354	293	469	504	500	458	376
7%	473	506	500	456	373	613	655	648	590	481
6%	642	684	674	611	496	833	888	874	792	641
5%	922	978	959	865	696	1.198	1.273	1.248	1.124	901

De tabel wijst uit dat er voor het meten van kleine verschillen betrekkelijk grote aantallen nodig zijn. Bij een base rate van 50% moet zowel de deelnemers- als de controlegroep uit minimaal 252 personen bestaan om een absoluut verschil van 10 procentpunten betrouwbaar, dat wil zeggen met een power van 70%, te kunnen meten. In een land als Nederland heeft dit belangrijke implicaties. De beschikbare onderzoeksantallen zijn vaak klein en dat betekent dat de effecten van daderprogramma's moeilijk zichtbaar gemaakt kunnen worden.⁷³ Daderprogramma's die niet breed worden toegepast, bijvoorbeeld omdat zij zich op een specifieke doelgroep richten of omdat zij zich nog in een experimenteel stadium bevinden, zullen lang moeten draaien om in aanmerking te komen voor een productevaluatie volgens de regelen der kunst.

Het probleem van de wachttijd klemmt des te meer omdat er bij recidive-onderzoek ook een observatietermijn in acht moet worden genomen. De personen uit het onderzoek moeten een periode in de gelegenheid zijn om nieuwe delicten te plegen. Bij de Recidivemonitor geldt een minimum van twee jaar (zie 3.3.2). Een kortere periode zou leiden tot een onbetrouwbaar meetresultaat, omdat dan niet zeker is of alle nieuwe delicten al zijn opgenomen in de administratie van justitie. Door gebruik te maken van een techniek als survival analyse kan de wachttijd iets worden bekort. Maar

⁷³ De gemiddelde power van de 59 effectstudies uit box 5.5 met een sms-score van 3 of hoger bedroeg 43%. Bijlage 7 geeft per studie de onderzoeksantallen, de power en een kwalificatie van het gemeten effect. Dit laatste uitgedrukt in termen van de effectgrootte w (zie noot 71).

men zal het recidiveonderzoek toch pas kunnen starten als er voldoende mensen twee jaar of langer 'at risk' waren sinds zij uit het programma vertrokken. Tel daarbij de tijd die nodig is voor dataverzameling, -verwerking en rapportage en het zal duidelijk zijn dat men vaak jaren zal moeten wachten voordat men de uiteindelijke effecten van een daderprogramma kent.

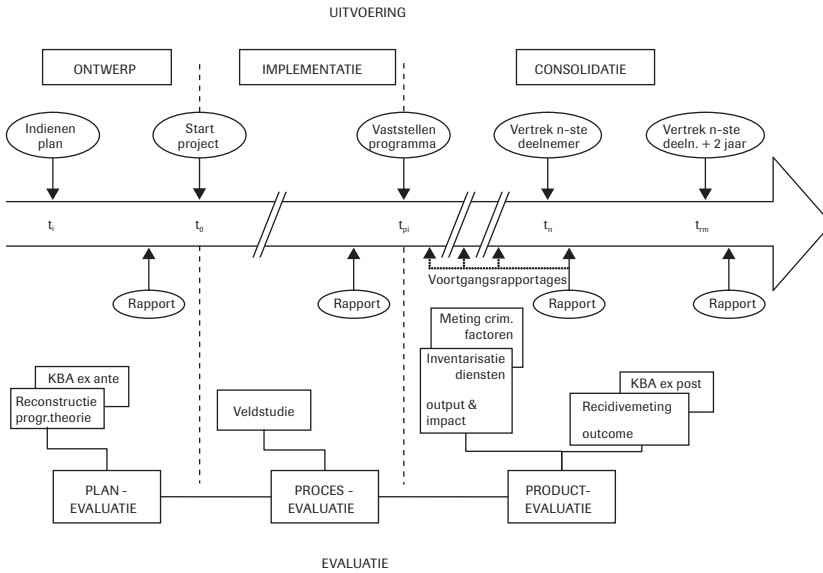
5.4 Slot: de evaluatieketen in beeld

Een recidivemeting is het geëigende middel om de uiteindelijke outcome van een daderprogramma te bepalen. Maar om teleurstellingen te voorkomen moet de meting wel aan een aantal voorwaarden voldoen. Bovendien dient ze vooraf te worden gegaan door een uitgebreide plan- en proces-evaluatie. Alleen dan krijgen we zicht op mechanismen die ervoor zorgen dat de interventie werkt zoals beoogd.

Figuur 5.2 maakt nog eens duidelijk hoe het onderzoek naar daderprogramma's kan worden gefaseerd. Onderin de figuur herkennen we de drie hoofdfasen van het onderzoek: de plan-, de proces- en de productevaluatie. Ze lopen gelijk op met de drie stadia die in de uitvoering van een daderprogramma te onderscheiden zijn: de fase waarin het programma wordt ontworpen, de fase waarin het programma in de praktijk wordt uitgetest en het stadium waarin het wordt uitgevoerd zoals het bedoeld is. Elke fase wordt afgesloten met een rapportage. In de fase van de productevaluatie kunnen de prestaties van het project in eerste instantie door middel van voortgangsrapportages worden bijgehouden. Op het moment dat er voldoende deelnemers zijn uitgestroomd om een vergelijkende analyse te doen die aan de technische kwalificaties voldoet, kunnen de verzamelde gegevens worden afgezet tegen de resultaten in een controlegroep.⁷⁴ Het bijhorende rapport brengt in beeld welke impact de interventie heeft gehad op de criminogene factoren van de deelnemers. Later volgt de recidivemeting en kan definitief worden vastgesteld of het daderprogramma aan de verwachtingen heeft voldaan. Het eindrapport kan worden opgesteld.

74 De n in tijdstip t_n staat voor het aantal observaties dat nodig is om het verwachte effect te kunnen meten (zie tabel 5.3).

Figuur 5.2 De PPP-evaluatieketen van daderprogramma's



Bron: Wartna (2005)

De aaneenschakeling van uiteenlopende typen onderzoek biedt de gelegenheid om de ‘kwalitatieve’ benadering die Pawson en Tilly voorstaan en de ‘kwantitatieve’ aanpak van de aanhangers van de Campbell Collaboration met elkaar te combineren. In hoofdstuk 2 zagen we dat Pawson en Tilly scherpe kritiek hebben op de experimentele traditie. De nadruk die deze traditie legt op de eis van ‘interne validiteit’ zou te zeer ten koste gaan van de vraag waar het werkelijk om draait, de vraag naar de context waarin de interventie werkzaam is en naar de onderliggende mechanismen die ermee in werking worden gezet.

De research syntheses van de ‘realistische’ school worden vaak gecontrasteerd met de experimentele benadering van de Campbell Collaboration en haar meta-analyses gericht op de decompositie van effect sizes. In feite echter vullen beide benaderingen elkaar aan.⁷⁵ Met Pawson en Tilly willen we begrijpen hoe een daderprogramma werkt en met Farrington willen we toetsen of het succesvol is. De standpunten van beide kampen zijn niet strijdig. Ze hebben betrekking op verschillende fasen in het onderzoeksproces. Een meta-analyse vindt plaats in de eindfase. Zij is een research synthese van een groep productevaluaties en onderzoekt de heterogeniteit

75 Knaap et al. (2008) wezen al eerder op de mogelijkheid de Campbell-standaard te combineren met de uitgangspunten van de realistische school van Pawson en Tilly.

in de outcome van de programma's die werden geëvalueerd. Elke afzonderlijke productevaluatie is – als het goed is – voorafgegaan door een plan- en een procesevaluatie waarin ruim aandacht is besteed aan context en mechanismen.

'Which methods work best for which types of offenders and under what conditions or in what types of settings?' (Palmer, 1975) – om deze vraag draait het uiteindelijk. Niet alleen een meta-analyse, maar elk type onderzoek levert een bijdrage aan het antwoord op deze vraag. Een kwantitatieve meta-analyse is pas aan de orde als de evaluatieketen van elk van de geselecteerde daderprogramma's is doorlopen. Tot die tijd mag evengoed geen enkele aanwijzing van de mogelijke werking van een daderprogramma worden veronachtzaamd. We hoeven niet te wachten op de uitkomsten van een rct om uitspraken te kunnen doen over de effectiviteit. Het doorlopen van de evaluatieketen levert immers 'voortschrijdend bewijs' op. Naarmate het onderzoek vordert, worden – als het goed is – de aanwijzingen van de werking van de interventie steeds sterker.⁷⁶ Indien een planevaluatie uitwijst dat een interventie goed is onderbouwd, als een procesevaluatie duidelijk maakt dat de interventie volgens plan wordt uitgevoerd en als uit tussentijdse metingen blijkt dat de gestelde behandeldoelen op overtuigende wijze worden gehaald, dan mogen we met toenemend vertrouwen *aannemen* dat het daderprogramma in zijn opzet slaagt en de recidive onder de deelnemers verlaagt. En mocht bij de uiteindelijke recidivemeting onverhoopt blijken dat we ons daarin hebben vergist, dan stellen we onze theorieën over effectieve interventiestrategieën bij en zoeken we verder.

76 Vergelijk dit met de 'effectiviteitsladder' van Veerman (2006), die vier treden onderscheidt waarop duidelijk wordt gemaakt of interventies effectief (kunnen) zijn.

6 Slotbeschouwing

Naar een samenhangend en uitvoerbaar programma van recidiveonderzoek

Hoe kunnen recidivemetingen worden vormgegeven en welke plaats nemen zij in bij de evaluatie van strafrechtelijke interventies? Zo luidde de probleemstelling van dit proefschrift. Het eerste gedeelte van deze vraag werd beantwoord in hoofdstuk 3. Met de procedures ontwikkeld in het kader van de Recidivemonitor is het mogelijk de verschillende facetten van de terugval van justitiabelen efficiënt en op eenduidige wijze in kaart te brengen. De belangrijkste beperking die daarbij geldt is dat het onderzoek alleen de strafrechtelijke recidive aan het licht brengt. De feitelijke recidive in een dadergroep zal in veel gevallen hoger zijn, omdat een onbekend deel van de gepleegde delicten aan de aandacht van de autoriteiten ontsnapt. Het is niet mogelijk deze beperking op te heffen. Het gebrek dient te worden geaccepteerd. Maar er valt mee te leven, want bij het vergelijken van de recidive van twee groepen zal steeds grofweg dezelfde 'meetfout' worden gemaakt. Omdat er bovendien controles zijn ingebouwd op het bestaan van registratie-effecten, is ook bij herhaalde metingen sprake van een constante onderschatting van de feitelijke recidive. De uitkomsten van het onderzoek zijn dus wel degelijk indicatief voor verschillen tussen groepen of in de tijd. Ze zijn in elk geval indicatief voor de mate waarin deze groepen in aanraking komen met justitie. Het tweede gedeelte van de probleemstelling – de vraag naar de rol van recidivemetingen bij de evaluatie van strafrechtelijke interventies – vergt een langer antwoord, een antwoord dat afhangt van het soort interventie dat men beschouwt. In dit proefschrift hebben we ons geconcentreerd op een specifieke groep van strafrechtelijke interventies. Het ging om programma's gericht op speciale preventie door de aanpak van criminogene factoren. Deze daderprogramma's komen in alle soorten en maten voor. Naast hun gemeenschappelijke doel delen zij echter nog een belangrijk kenmerk. Bij een daderprogramma wordt de interventie waarvan men de werking wil vaststellen, bij herhaling toegepast. Dit maakt het in principe mogelijk om willekeurige kandidaten de interventie niet te laten ondergaan. En dit laatste opent de weg naar een experimentele opzet van het evaluatieonderzoek.

Het sterkst denkbare bewijs van de effectiviteit van een daderprogramma is een verschil in recidive in een vergelijkende studie met een gerandomiseerd design. Met een experiment worden twee parallele werelden gecreëerd waarin alleen de interventie het verschil uitmaakt. Maar constateren dat een ingreep succesvol was, is niet voldoende. We willen ook weten waarom het ene daderprogramma werkt en het andere niet. Daarom dient zelfs een effectstudie met een gerandomiseerd design te worden voorafgegaan door een zorgvuldige plan- en procesevaluatie. Deze vormen van onderzoek leveren ons de kennis op die nodig is om de resultaten van de

effectmeting te interpreteren en te gebruiken bij de verdere invulling van het strafrechtelijke beleid.

Een experimenteel design of een quasi-experimenteel design (waarin randomisering is vervangen door matching) is niet altijd haalbaar. Bij complexe, eenmalige operaties zoals de invoering van een nieuwe wet of een algemene wijziging in het strafrechtelijke beleid is het opzetten van een experiment geen optie. Alleen als de interventie kan worden ‘aan- of uitgezet’ en zij plaatsvindt in een wereld die voor het overige onveranderd blijft, behoort willekeurige toewijzing tot de mogelijkheden. Als er geen experiment mogelijk is, moet het vraagstuk van de causaliteit – het ‘attributieprobleem’ – op een andere manier worden opgelost. Men dient dan goed op de hoogte te zijn van ‘wat er in de wereld speelt’, om de invloed van de interventie op de uitkomst die men meet te kunnen scheiden van het effect van gelijktijdige ontwikkelingen. Dat is niet eenvoudig en soms ronduit onmogelijk. De metingen die men doet, kunnen nog zo nauwkeurig zijn, maar de betekenis van de meetresultaten blijft diffuus. Een eenduidige conclusie levert het onderzoek niet op. De betekenis van de resultaten van het onderzoek wordt gevormd in het maatschappelijke veld waarin het onderzoek wordt opgemerkt en waar iedereen – politici, opinieleiders, professionals en wetenschappers voorop – zijn eigen interpretatie aan de uitkomsten geeft.

Ook recidivecijfers kunnen bijdragen aan de maatschappelijke discussie over wat ‘waar’ en ‘goed’ is. Daarbij geldt een eenvoudige wet: hoe globaler de onderzoeksvraag, hoe minder informatief de cijfers zijn. In dit slothoofdstuk wordt nog eens op een rij gezet op welke wijze het recidive-onderzoek onze kennis van de werking van straffen en maatregelen vergroot. Er zijn grofweg drie niveaus waarop recidivemetingen worden toegepast om uitspraken te doen over effecten van ‘strafrechtelijke interventies’: nationaal, regionaal en lokaal. Op elk niveau speelt het causaliteitsvraagstuk. Op elk niveau wil men weten hoe men de uitkomsten moet interpreteren en waaraan men ze kan toeschrijven.

6.1 Nationale overzichten en trendstudies

Het meest basale niveau waarop recidivecijfers behulpzaam zijn, is dat van de landelijke overzichten. Dit is het terrein van de vaste metingen van de Recidivemonitor. In hoofdstuk 3 hebben we gezien hoe de ontwikkeling van de OBJD het mogelijk heeft gemaakt om op nationale schaal onderzoek te doen. De Recidivemonitor brengt het Nederlandse recidiveland-schap in kaart. De jaarlijkse metingen hebben betrekking op zo goed als alle personen die in ons land met justitie in aanraking zijn gekomen. Voor de meeste groepen zijn data beschikbaar vanaf 1997, voor de tbs-sector zelfs vanaf 1974.

De cijfers van de Recidivemonitor zijn niet erg geprononceerd. De uitkomsten die in hoofdstuk 4 zijn gepresenteerd, geven een algemeen beeld van de globale ontwikkeling van de strafrechtelijke recidive; veel diepgang hebben ze niet. Bij de jeugdigen en de volwassen justitiabelen, de twee grootste onderzoekspopulaties, kunnen de resultaten slechts naar enkele dader- of zaakkenmerken worden uitgesplitst. Het is op dit moment bijvoorbeeld niet mogelijk om binnen de groep daders die een werk- of een leerstraf kregen, onderscheid te maken naar de aard van de opgelegde straf. We weten dus niet of een 'kale' werkstraf een hoger terugvalpercentage te zien geeft dan een werkstraf waarbij sprake is van reclasseringsbegeleiding. Om dergelijke hiaten op te vullen zullen de metingen van de monitor in de komende jaren worden uitgebreid. Naast de cliënten uit de drie sectoren van DJI kunnen bijvoorbeeld ook de cliënten van andere uitvoeringsorganisaties, zoals de reclassering en de raad voor de kindbescherming, in het onderzoek worden betrokken. Door gebruik te maken van de in- en doorstroomgegevens van deze organisaties kunnen meer onderscheidingen in het materiaal worden aangebracht en ontstaat er een genuanceerder beeld van de recidive die volgt op de sancties die in Nederland worden uitgevoerd.

Naast uitbreiding van het aantal dadergroepen moet er echter meer gebeuren. Want hoe gedifferentieerd het recidivebeeld ook moge zijn, de platte cijfers laten niet zien waarom de recidive is gestegen of gedaald. De trends die de Recidivemonitor blootlegt zijn voor meerdere uitleg vatbaar. Stijgen de landelijke recidivepercentages omdat er sprake is van toenemende criminaliteit, nemen de effecten van onze straffen en maatregelen misschien af, of is er toch sprake van een registratie-effect? De Recidivemonitor kan de discussie over de betekenis van de resultaten niet beslissen, maar wel verder op weg helpen. Om te begrijpen waarom de cijfers zijn gedaald of juist zijn toegenomen zijn 'verklarende analyses' nodig. De uitkomsten van de Recidivemonitor moeten in verband worden gebracht met ontwikkelingen die zich gelijktijdig hebben voorgedaan. Dat kan door aan het voorspellingsmodel dat gebruikt wordt voor de correctie van de ruwe cijfers factoren toe te voegen die deel uitmaken van een criminologische theorie. In paragraaf 3.4.1 bespraken we een aantal Amerikaanse meta-analyses waarin is nagegaan welke indicatoren consistent samenhangen met de in de geselecteerde studies gebruikte criminaliteitsmaat. Een dergelijke benadering is ook mogelijk in Nederland. De beschikbaarheid van digitale gegevens neemt toe en daar kan op worden geanticipeerd. Door de gegevens van de Recidivemonitor te koppelen aan informatie uit andere bronnen kan worden onderzocht welke factoren de hoogte van de recidive het best weten te voorspellen. Te denken valt hier bijvoorbeeld aan sociaaleconomische informatie zoals algemene werkloosheidscijfers en cijfers met betrekking tot het aantal voortijdige schoolverlaters. Maar ook aan informatie over buurten en aan aanvullende gegevens op persoonsniveau met betrekking tot opleiding bijvoorbeeld, of

eerdere hulpverleningscontacten. Door te werken vanuit causale modellen kan worden geanalyseerd hoe de recidive tot stand komt en onder welke omstandigheden die uitblijft. De ‘verdiepingsslag’ die hier wordt voorgesteld, zal het nut van de Recidivemonitor belangrijk doen toenemen. Naast kennis van de omvang van de terugval levert het onderzoek dan immers ook inzicht op in de achtergronden ervan.

Het is de ambitie van het kabinet-Balkenende IV de prevalentie van recidive onder ex-gedeteneerden en jeugdige justitiabelen met 10% te verlagen. Deze operationele doelstelling kan met de gecorrigeerde cijfers van de Recidivemonitor betrekkelijk nauwkeurig worden getoetst. De vraag in hoeverre de ‘netto-ontwikkeling’ van de recidive een direct gevolg is van het gevoerde beleid, zal de Recidivemonitor niet uitputtend kunnen beantwoorden. Maar met behulp van verklarende analyses kan het antwoord wel dichterbij worden gebracht. Door na te gaan hoe de trend in de recidivecijfers zich verhoudt met de ontwikkeling van de criminaliteit in het algemeen, door nog beter na te gaan of er misschien sprake kan zijn van registratie-effecten en door het bijsluiten van theoretisch relevante informatie die nieuw licht werpt op de instroomverschillen tussen de opeenvolgende cohorten van justitiabelen, kan de omvang van de recidive op nationaal niveau verder worden ontleed. Wie weet blijft na het wegnemen van de invloed van deze gelijktijdige ontwikkelingen nog een daling in de recidivecijfers over. Dat zou dan het effect van het strafrechtelijke beleid kunnen zijn, meer bepaald het effect van de maatregelen gericht op het voorkomen van herhalings-criminaliteit. Maar de recidive kan natuurlijk ook stijgen.

6.2 Uitsplitsingen naar regio, instelling of inrichting

De volgende toepassing van recidivecijfers doet zich voor op het niveau van organisaties. Het strafrechtelijke bedrijf kent een aantal schakels die zijn opgesplitst in regionaal werkende eenheden. De politie kent 25 districten, de rechterlijke macht 19 arrondissementen en de reclassering 10 regio’s. Daarnaast bestaan de drie sectoren van de DJI elk uit een aantal inrichtingen. Het gevangeniswezen telt 20 clusters met een totaal van meer dan 60 hvb’s en gevangnissen. Het JJI-veld gebruikt minstens 13 opvang- en behandelinrichtingen en de tbs-sector bestaat uit 12 klinieken en centra. Binnen elke schakel hebben de eenheden eenzelfde of in elk geval een soortgelijke taak. Dit maakt het mogelijk om de prestaties van deze organisatie-eenheden in een *benchmark* met elkaar te vergelijken. Camp (1989) omschrijft benchmarking als ‘the search and implementation of best practices’. Maar met minder neemt men ook genoegen; de bedoeling van een benchmark is dat de organisatieonderdelen door vergelijking en analyse van elkaar leren om zo tot betere prestaties te komen.

Door ze uit te splitsen naar regio of instelling kunnen recidivecijfers onderdeel uitmaken van een benchmark. Daar zit wel een gevaar aan vast.

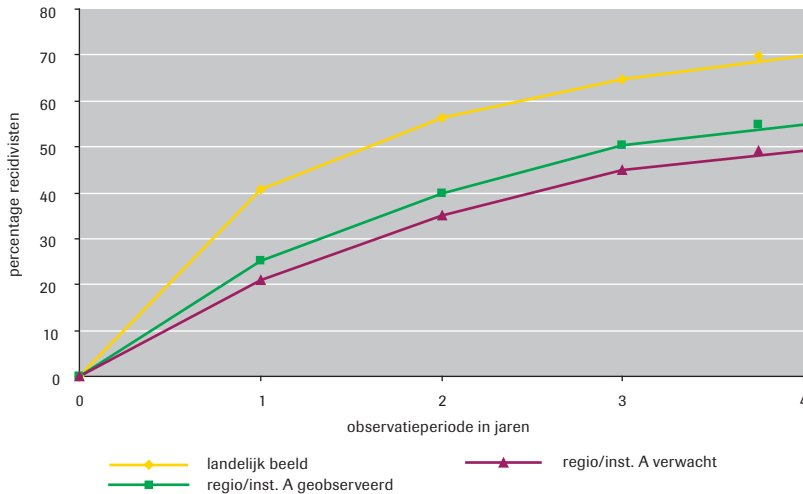
Recidive wordt snel gezien als prestatie-indicator. De hoogte van de recidive laat zien 'hoe goed men het heeft gedaan'. Voor een tbs-kliniek of een jeugdbehandelinrichting past dat wel in het beeld, maar andere organisatieonderdelen werken toch vooral aan andere taken. Voor de overheid is het terugdringen van recidive een centrale doelstelling, maar dat betekent niet dat alle onderdelen van de organisatie daar in gelijke mate verantwoordelijk voor zijn. De kans dat een justitiabele recidiveert is bovendien van vele factoren afhankelijk. Vaak ligt slechts een deel daarvan binnen de invloedssfeer van de medewerkers in de instellingen. Het is dus zeker niet zo dat elk organisatieonderdeel simpelweg kan worden 'afgerekend' op de hoogte van de recidive. Dat hangt af van de mate waarin de reductie van de terugval expliciet tot de opdracht behoort.

Wel heeft het uitsplitsen van de cijfers altijd informatieve waarde. Het is goed om te weten hoe omvangrijk de terugval is, al was het maar om in te kunnen schatten welk deel van de 'klanten' men terug kan verwachten. En er gaat van de cijfers toch een zekere signaalwerking uit. Als de recidive relatief hoog is, zal men zich toch afvragen hoe dit komt en of men er iets aan kan doen. Een kaal recidivepercentage fungeert dus wel degelijk als 'kengetal'. Het kan binnen de organisatie een discussie op gang brengen over de taakopvatting en de inrichting van de werkzaamheden. Als het recidivebeeld van een instelling sterk afwijkt van dat van haar soortgenoten uit andere regio's, zal men daarvoor een verklaring willen hebben. Het probleem is dat de instellingen te maken hebben met instroomverschillen. Het ene arrondissement krijgt andere verdachten te verwerken dan het andere en de ene inrichting herbergt zwaardere gedetineerden dan de volgende; niet alle justitiabelen hebben hetzelfde risicoprofiel. Dit probleem kan echter worden opgelost door gebruik te maken van een voorspellingsmodel. Per instelling wordt niet alleen gekeken naar de ruwe, geobserveerde recidive, maar ook naar het percentage recidivisten dat mocht worden verwacht afgaand op de achtergronden van de justitiabelen die de instelling te verwerken kreeg. Figuur 6.1 illustreert het een en ander. We zien de geobserveerde en verwachte recidivepercentages in de groep personen waar één bepaalde instelling mee te maken kreeg. Het zou hier kunnen gaan om een politiedistrict, een arrondissement of een penitentiaire inrichting. De geobserveerde recidive is lager dan het gemiddelde voor alle soortgelijke organisatie-eenheden, maar hoger dan recidive die kon worden verwacht gelet op de samenstelling van de groep. In dit voorbeeld heeft de eenheid uiteindelijk dus een ongunstige 'score'. Misschien ligt dit aan de groep die men te verwerken kreeg, maar misschien ook is het terug te voeren naar een factor binnen de eigen organisatie.

Men hoeft dit type analyse niet op de gehele populatie toe te passen. Desnoods beperkt men zich tot een deel van de cliënten of een deel van de instellingen binnen de sector. Zo zou de reclassering bijvoorbeeld speciaal geïnteresseerd kunnen zijn in regionale verschillen in de recidive van taakgestraften en de DJI in de verschillen tussen de inrichtingen die

uitvoering geven aan penitentiaire programma's voor langgestraften. Afhankelijk van de vraag die voorligt, past men de analyse aan. In de tbs-sector, meer algemeen de forensische zorg, ligt een dergelijke benchmark sterk voor de hand. Het vraagstuk van de effectiviteit van de behandeling van terbeschikkinggestelden kan vanwege de lage aantallen alleen via een *multicenter study* worden aangepakt. De klinische gegevens van de instellingen kunnen worden gekoppeld aan de justitiële gegevens van de Recidivemonitor. Het voorspellingsmodel kan dan worden uitgebreid, zodat de verwachte recidive nauwkeuriger kan worden vastgesteld en er nog beter rekening kan worden gehouden met instroomverschillen tussen de deelnemende instellingen.

Figuur 6.1 Verwachte en geobserveerde recidive onder ex-cliënten van een regio of organisatieonderdeel



Zoals aangegeven ligt bij het gebruik van recidivecijfers in een benchmark het gevaar van 'over-interpretatie' op de loer; men kan gemakkelijk te veel aan de cijfers aflezen. Een directeur van een organisatieonderdeel zal naar aanleiding van uitkomsten als in figuur 6.1 ter verantwoording worden geroepen. Daar is op zich niets mis mee, mits duidelijk is welke rol de cijfers spelen. Op voorhand moet duidelijk zijn hoezeer de hoogte van de recidive tot de verantwoordelijkheden behoort van de medewerkers in de verschillende schakels van het strafrechtelijke bedrijf. Belangrijk is bovendien dat de cijfers niet geïsoleerd worden beschouwd, maar worden ingebed in een systeem waarin meer en juist ook andersoortige output- en outcome-indicatoren zijn opgenomen. Ideaal zou zijn als men een dergelijke benchmark opbouwt vanuit een expliciet model waarmee de onderlinge verhouding tussen de verschillende indicatoren is vastgelegd

en wordt getoetst. Indien de recidivecijfers een goede plaats krijgen in het geheel, kan er sprake zijn van een zinvolle bijdrage. Essentieel is dat er binnen de organisatie helder wordt gecommuniceerd wat op dit punt de verwachtingen zijn.

6.3 Evaluatie van lokale, specifieke interventies

Naast de nationale overzichten en de rapportages per instelling of regio worden recidivecijfers ingezet bij specifiek evaluatieonderzoek. Op dit niveau is het object van onderzoek niet het strafrechtelijke beleid of het functioneren van een onderdeel van de organisatie, maar een op zichzelf staande interventie, een 'daderprogramma', zoals wij dat hier genoemd hebben. In hoofdstuk 5 werd beschreven hoe de evaluatie van een daderprogramma dient te verlopen. Benadrukt werd dat het meten van de recidive slechts een onderdeel van het onderzoek is. Evaluatie van daderprogramma's vindt plaats in een aaneenschakeling van uiteenlopende onderzoeksactiviteiten. Aan het vaststellen van de uiteindelijke uitkomsten gaat een plan- en een procesevaluatie vooraf.

Ook op dit niveau speelt uiteraard het attributieprobleem: hoe weten we dat de interventie effect heeft gehad? Omdat er bij de evaluatie van daderprogramma's gewerkt kan worden met vergelijkingsgroepen, zijn er nu meer mogelijkheden om causale uitspraken te doen. De beste optie is het opzetten van een *randomized controlled trial*. Kandidaten voor de interventie worden willekeurig toegewezen aan een onderzoeks- en een controleconditie. Randomisering zorgt ervoor dat de uitgangssituatie voor beide groepen gelijk is. Dat is vooral nuttig indien de omstandigheden die leiden tot een bepaalde uitkomst niet goed bekend zijn (Weisburd, 2004). Als we die omstandigheden wel goed kennen, hebben we het toeval minder hard nodig. Een quasi-experimenteel design waarin voor de belangrijkste overgebleven verschillen tussen de onderzoeks- en de controlegroep wordt gecorrigeerd, is dan acceptabel, vooral in situaties waar randomisering op juridische of zwaarwegende praktische bezwaren stuit. Het is dan wel van belang dat we weten welke kenmerken het beste kunnen fungeren als criteria voor het matchen van de personen in de onderzoeks- en de controlegroep.

Dit laatste geeft nog eens aan hoe belangrijk het is voorafgaand aan de recidivemeting kwalitatief onderzoek te doen. Het denken in termen van een keten van onderzoeken maakt dat een recidivemeting niet in het luchtledige plaatsvindt. De plan- en de procesevaluatie hebben – als het goed is – al de nodige kennis opgeleverd. De uitgangssituatie van de kandidaten voor de interventie is bekend. Er is een analyse gemaakt van de mechanismen die met de interventie in werking worden gezet en er is vastgesteld welke veranderingen er in de uitgangssituatie van de deelnemers zijn opgetreden. Met de plan- en de procesevaluatie is er

goed beschouwd gewerkt aan een basis om causale uitspraken te doen. Wanneer we goed op de hoogte zijn van de omstandigheden die normaal gesproken tot recidive leiden, is het niet per se nodig om in de slotfase van het onderzoek te randomiseren om de causale rol van de interventie vast te stellen. Er is tijdens de plan- en de procesevaluatie immers al *waargenomen* dat de interventie tot gewijzigde omstandigheden leidt. Alleen de vraag of dat in de onderzoeksgroep vaker gebeurt dan in een goed lijkende controlegroep, moet nog worden beantwoord. Indien het daderprogramma volgens de regels der kunst is opgezet en de eerdere evaluatiefasen zijn doorlopen, dan vormt de uitkomst van de recidivemeting – als alles goed is – niet meer dan een bevestiging van wat we al wisten.

6.4 Beter meten is beter weten

Er zijn aanwijzingen dat daderprogramma's kunnen werken, maar deze aanwijzingen komen voornamelijk uit het buitenland. In Nederland zijn we nog niet zo ver. Hier staan we in feite nog voor de vraag welk belang we willen hechten aan dergelijke programma's. De overheid bepaalt het centrale beleid, maar in de praktijk van de strafrechttoepassing worden de concrete strafdoelen gekozen. Houden we het de komende jaren bij vergelding en incapacitatie of geven we het idee van resocialisatie een nieuwe kans? Voor de inrichting van ons strafrechtstelsel is het van cruciaal belang dat er ook in Nederland bewijzen van goed lopende programma's boven water komen. Dit vraagt om veel en vooral passend onderzoek, maar ook om onderzoek dat efficiënt is ingericht en – niet onbelangrijk – praktisch haalbaar is. Wachten tot zich een schaarse mogelijkheid voordeet om een gerandomiseerde studie uit te voeren, is geen goede strategie. Doorgaan met het uitvoeren van kleinschalige studies zonder geschikt vergelijkingsmateriaal evenmin.

Passend onderzoek wil zeggen dat de evaluatie is afgestemd op de fase waarin de ontwikkeling van een daderprogramma zich bevindt. Efficiënt onderzoek wil zeggen dat er in elke schakel van de evaluatieketen sprake is van accumulatie van kennis. Met de komst van de Recidivemonitor zijn stappen gezet om de effectmetingen van een systematischer grondslag te voorzien. De toegang tot een belangrijke gegevensbron is verbeterd, de dataverzameling is geautomatiseerd en door de standaardisering van het meetproces ligt het vergelijkingsmateriaal inmiddels bijna letterlijk voor het oprapen.⁷⁷ Via het onderzoek in de vaste daderpopulaties houden we bij welke persoons- en zaakkenmerken over het algemeen geassocieerd zijn met hoge of juist lage terugvalkansen.

⁷⁷ In oktober 2008 werd op de site van het WODC een webapplicatie geplaatst die toegang geeft tot de uitkomsten van het onderzoek van de Recidivemonitor.

De voorspellingsmodellen die binnen de Recidivemonitor worden gebruikt spelen op alle toepassingsniveaus een rol. Op nationaal niveau leiden de modellen tot de correctie van de ruwe cijfers. Binnen de verschillende sectoren van het strafrechtelijke bedrijf worden ze aangevend om per organisatie-eenheid de verwachte recidive te berekenen, en bij het evaluatieonderzoek van lokaal geïmplementeerde daderprogramma's zorgen de modellen voor statistische controle van instroomverschillen. De modellen zijn op het moment niet bijzonder fijn van samenstelling, het is belangrijk ze verder te ontwikkelen. De predictie van recidive berust op een handvol kenmerken uit de OBJD, het bronbestand van de monitor. De sekse, de leeftijd en het geboorteland van de dader, het type delict dat werd gepleegd, het aantal eerdere justitiecontacten en de leeftijd bij het eerste contact; het zijn alle krachtige voorspellers, maar ze vertellen ons weinig over de achtergronden en oorzaken van de recidive. Om beter zicht te krijgen op de onderliggende processen zullen de databestanden van de Recidivemonitor worden verrijkt met informatie uit andere bronnen. Welke gegevens dat precies zijn hangt af van de dadergroep die met het onderzoek wordt gevolgd. Als zij passen in een theoretisch model, kunnen er verklarende analyses worden gedaan. De beschikbaarheid van relevante gegevens neemt toe, ook op persoonsniveau. In paragraaf 2.1 zagen we dat er in diverse sectoren gewerkt wordt aan de ontwikkeling van screeningsinstrumenten bedoeld om de criminogene factoren van individuele justitiabelen in kaart te brengen. De informatie die met behulp van deze instrumenten is vastgelegd, zal aan de justitiële gegevens van de OBJD worden gekoppeld. In eerste instantie om de instrumenten te valideren, maar in tweede instantie om algemeen onderzoek te kunnen doen naar de oorzaken en achtergronden van (strafrechtelijke) recidive. Als we meer weten van de determinanten van herhaald crimineel gedrag, kunnen we ook beter formuleren wat we er aan kunnen doen om het te voorkomen. De randvoorwaarden voor het uitvoeren van effectieve interventies worden in Nederland steeds verder ingevuld. Het wordt tijd om ons geloof in de mogelijkheden tot resocialisatie serieus op de proef te stellen.