

LUC ROBERT^a

Wat werkt in Nederland en wat niet?

Een meta-analyse van Nederlands recidiveonderzoek naar de effecten van strafrechtelijke interventies

WARTNA, B.S.J., ALBERDA, D.L. & VERWEIJ, S. (2013). *Wat werkt in Nederland en wat niet? Een meta-analyse van Nederlands recidiveonderzoek naar de effecten van strafrechtelijke interventies*. Boom: Lemma Uitgevers, 196 p.



Panopticon, 35 (5), 480-485

© 2014 MAKLU | ISSN 0771-1409 | SEPTEMBER 2014

^a Onderzoeker, Nationaal Instituut voor Criminalistiek en Criminologie (NICC), Brussel (Corresp.: Luc.Robert@just.fgov.be).

Dit boek is de neerslag van een eerste deel van een onderzoek aan het Wetenschappelijk Onderzoek- en Documentatiecentrum (het WODC) in Nederland waarbij nagegaan wordt of strafrechtelijke interventies in Nederland en Vlaanderen recidive terugdringen, in welke mate welke programma's daar al dan niet in slagen en hoe en waarom ze (in)effectief zijn.

De auteurs van deze studie beperken zich in dit werk tot het beantwoorden van drie onderzoeksvragen (p. 27): welke empirische studies bestaan er in Nederland en Vlaanderen die het speciaal preventieve effect van strafrechtelijke interventies nagaan? Welke methoden van onderzoek zijn daarbij gebruikt? Wat zijn de uitkomsten in termen van recidivereductie en welke types interventies tonen (on)gunstige resultaten? Waarom een bepaalde interventie een specifiek resultaat heeft, komt in dit onderzoek niet aan bod.

De structuur van dit werk zit logisch in elkaar. Eerst geven de auteurs een samenvatting van het onderzoek, met inbegrip van de conclusies en discussie van de bevindingen. Het eigenlijke werk bestaat uit drie grote onderdelen. In de inleiding schetsen de onderzoekers de context voor dit werk, vermelden ze de onderzoeksvragen en hoe ze daarop een antwoord willen geven. In hoofdstuk 2 beschrijven ze nauwgezet de methode van het onderzoek, met inbegrip van de inclusiecriteria voor het opnemen van studies, de manier waarop ze gezocht hebben, hoe de studies verwerkt zijn en de analyses zijn uitgevoerd. Hoofdstuk 3 bevat de voornaamste resultaten, met inbegrip van een algemeen overzicht van de effectgroottes (*effect sizes*) overheen alle opgenomen onderzoeken, een beschrijving van publicatievariabelen, methodologische aspecten, daderkarakteristieken en interventiekenmerken, en afsluitend een meta-regressie en de bespreking daarvan. De hoofdtekst van het boek eindigt al op p. 76. Er volgen 9 zeer uitgebreide en gedetailleerde bijlagen, in totaal meer dan 100 pagina's, met onder meer een begrippenlijst, een overzicht van de recidivestudies die ze in hun literatuurzoektocht hebben gevonden, een omschrijving van de in de meta-analyse

opgenomen interventies, de codeerlijst, technische aspecten van de meta-analyse en methodologische aspecten van de effectstudies.

De lezer wordt van bij het begin van dit boek bij de hand genomen en stapsgewijs vertrouwd gemaakt met keuzes, begrippen, methode, rekenwijzen, interpretaties en diets meer. Zelf ben ik allerm minst een expert als het op meta-analyses aankomt, maar door de manier waarop de auteurs van dit rapport de lezer bij de hand nemen, heb ik dat niet als een handicap ervaren. De behapbare en glasheldere manier waarop ze deze soms complexe materie uiteenzetten, verdient dan ook een dikke pluim. Zelf raad ik geïnteresseerden aan om het werk bij de hand te nemen. Ik geef hier een relatief uitgebreide bespreking van hun onderzoek en werp op het einde enkele vragen en reflecties op.

De auteurs beogen met dit onderzoek een eerste systematische review van alle Nederlandse en Vlaamse onderzoeken die (minstens) de prevalentie van recidive na een strafrechtelijke interventie gemeten hebben (effectstudie). Een *systematic review* definiëren ze zelf als ‘een samenvatting van de onderzoeksliteratuur op een bepaald gebied. Er wordt gebruikgemaakt van transparante, repliceerbare methoden om relevante studies te verzamelen en te analyseren’ (p. 100). Aan die omschrijving komen de auteurs zelf met verve tegemoet. WARTNA en collega’s willen op basis van kwaliteitsvolle onderzoeken een meta-analyse uitvoeren. Een meta-analyse, zo omschrijven ze zelf, is ‘het onderdeel van een systematic review waarin een kwantitatieve samenvatting wordt gegeven van de uitkomsten van de verzamelde evaluatiestudies’, zodat die onderling vergelijkbaar zijn (p. 25, ook p. 95), wat in de ogen van de auteurs een specifieke invulling geeft aan een systematic review. Om dat te kunnen doen, dienen ze eerst uitvoerig na te gaan of studies beantwoorden aan een aantal vooraf bepaalde kwaliteiten en kenmerken. Ze omschrijven haarfijn de inclusiecriteria om een studie op te nemen en voegen een codeerlijst toe in bijlage, zodat lezers een beeld krijgen van de uiteenlopende kenmerken die twee codeurs onafhankelijk van elkaar hebben gecodeerd.

Alle onderzoek dat een meting van de recidiveprevalentie bevat, is opgenomen tot 1 juli 2011, ongeacht hoe oud de studie is en of de studie al of niet gepubliceerd is. Ook ‘grijze literatuur’, niet-gepubliceerde rapporten, dissertaties en andere teksten, komen in aanmerking om in de meta-analyse te belanden (p. 30, noot 9, omschrijft wat de onderzoekers onder ‘grijze literatuur’ verstaan: het gaat om een studie die ‘niet is gepubliceerd via een erkende uitgever en niet eenvoudig te verkrijgen valt’). Om opgenomen te worden in het onderzoek dient het wel te gaan om een ‘primaire studie’, een studie die gebruik maakt van ‘originele empirische gegevens’ (p. 30). Andere inclusiecriteria betreffen het soort interventie (dit is breed gehouden: ‘alle soorten strafrechtelijke interventies’ (p. 30)); het soort onderzoeksgroep: de onderzoeksgroep dient te bestaan uit daders of verdachten (p. 30); de uitkomstmaat (prevalentie van recidive als een uitkomstvariabele) (p. 31); en het soort onderzoekdesign, te weten enkel studies waarbij een controlegroep is opgenomen (p. 31-32). Studies kunnen gepubliceerd zijn in het Nederlands of het Engels.

De onderzoekers zetten uiteen hoe ze op basis van alle inclusie- en exclusiecriteria uiteindelijk bij een lijst evaluatiestudies zijn terechtgekomen (p. 47-50). De selectie is treffend gevat in figuur 3.1. (p. 48). Maar liefst 350 publicaties leken bij aanvang een kandidaat om opgenomen te worden in de meta-analyse. 209 studies zijn uitgefilterd omdat er geen recidive gemeten is, het niet ging om een ‘primaire studie’ of om een effectstudie, of omdat het geen strafrechtelijke interventie betreft. In nog eens 60 andere studies ontbreekt een controlegroep. In 1 studie ontbreken gegevens over de

grootte van de experimentele en controlegroepen. 12 studies zijn weggelaten omdat ze geen informatie bevatten over de recidiveprevalentie. Uiteindelijk blijven nog 68 effectstudies over, waarin in totaal 83 vergelijkingen zitten die in de meta-analyse kunnen worden opgenomen.

Voor elke geselecteerde studie is de vergelijking gemaakt tussen de uitkomsten van de groep die een interventie heeft ondergaan en de resultaten van de controlegroep. De effectgroottes zijn berekend op basis van de *odds ratio* (verder: OR). De *odds ratio* is 'de verhouding tussen de kans dat een gebeurtenis plaatsvindt en de kans dat een gebeurtenis uitblijft' (p. 43). De auteurs vermelden de formule om dat te doen (p. 44). Het komt er op neer dat de *odds* op recidive (te weten: de proportie recidive gedeeld door de proportie niet-recidive) in de controlegroep gedeeld wordt door de *odds* op recidive in de experimentele groep. Het resultaat daarvan is een cijfer tussen 0 en oneindig. Ligt de OR tussen 0 en 1, dan heeft de controlegroep minder recidivisten dan de experimentele groep; ligt de OR op 1, dan is er geen verschil; ligt de OR boven de 1, dan is er minder recidive bij de experimentele groep dan bij de controlegroep en is de interventie dus positief (p. 44). Bij de OR is een betrouwbaarheidsinterval van 95% berekend, wat een schatting is voor de nauwkeurigheid van het gemeten effect (p. 44). Voor het geheel van de geselecteerde studies kunnen de OR onderling vergeleken worden, wat een indruk geeft van welke interventies effectief blijken, welke interventies niet tot een verschil leiden tussen de experimentele en controlegroep en welke interventies leiden tot meer recidive voor de personen die de interventie hebben ondergaan.

Daarnaast is ook gekeken naar de invloed van mogelijke moderatoren. De auteurs onderzoeken de invloed van vier groepen moderatoren: publicatievariabelen, methodologische aspecten van het effectonderzoek, daderkarakteristieken en interventiekenmerken. Ze hebben dit eerst beschrijvend nagegaan, maar nadien hebben ze de invloed van de moderatoren onderzocht aan de hand van een meta-regressie, 'een multivariate analysetechniek die berekent wat de onafhankelijke bijdrage is van een moderatorvariabele bij de voorspelling van een uitkomstmaat' (p. 45).

WARTNA, ALBERDA en VERWEIJ stellen vast dat de OR voor alle onderzoeken samen 1,09 bedraagt, met een betrouwbaarheidsinterval dat ook de waarde 1 bevat. Statistisch gezien is er dus overheen het geheel van alle studies in de meta-analyse geen significant verschil tussen de effecten van strafrechtelijke interventies en de controlegroepen (p. 63). Dat cijfer zegt echter bijzonder weinig, want de effectstudies zijn veel te uiteenlopend van aard; de onderzoekers spreken zelf van een 'bijzonder sterke heterogeniteit' (p. 63, noot 18) (indien het een evaluatie van eenzelfde type interventie betreft, zou dergelijke algemene OR wel zeer zinvol zijn).

Een globaal overzicht van de individuele *odds ratio* van de in de meta-analyse opgenomen interventies is weergegeven in figuur 3.3. (p. 66). Deze figuur (in het jargon spreekt men van een *caterpillar plot*) geeft visueel zeer goed weer hoe de verschillende interventies kunnen gesitueerd worden ten opzichte van elkaar in termen van de OR (p. 66). 53 van de 83 interventies hebben een OR die hoger is dan 1. Voor 26 interventies (of 31,3%) ligt het 95% betrouwbaarheidsinterval boven de 1. Dat wil zeggen dat de recidiveprevalentie van de groep die de interventie onderging, significant lager ligt dan die van de controlegroep. Volgens een statistische conventie wijst een significante OR hoger dan 2,47 op een middelmatig positief effect, een significante OR boven de waarde 4,25 wijst op een sterk effect (p. 44). Slechts 10 studies hebben een significante OR boven de 2,47. Voor 18 interventies (toch meer dan één vijfde van alle opgenomen interventies – 21,7%) blijkt een significant negatief effect. Een OR kleiner dan 1 wijst erop dat de prevalentie van recidive bij personen die de interventie ondergingen,

hoger ligt dan bij een controlegroep; met andere woorden: van diegenen waarop in deze interventies gefocust is, recidiveerden er achteraf meer dan in een controlegroep (p. 64-65). 9 studies hebben zelfs een significante OR van minder dan 0,4, dus een middelmatig tot sterk negatief effect. In de geneeskunde zou men hier van een *nocebo*-effect durven spreken.

Daarna richten de onderzoekers hun pijlen op eventuele moderatoreffecten bij de meta-analyse. Eerst geven ze een beschrijvend beeld van de vier types moderator-effecten; ze gaan de gemiddelde OR na (dus per categorie of groep) en staan stil bij significante invloeden. Daarna bekijken ze in een meta-regressie of die significante invloeden hier bevestigd blijven en wat hun eventuele individuele impact is. Er wordt hier gekozen voor het 10% significantieniveau (dus niet het gebruikelijke 5%-niveau) vanuit de overweging dat het gaat om een explorerende studie (zie p. 68). Uit de meta-regressie komen vijf kenmerken naar voor die een significante samenhang vertonen met de *odds ratios* (p. 73-76). Grijs literatuur heeft een lagere OR dan gepubliceerde studies, wat WARTNA en collega's zien als een indicatie voor het *file drawer* probleem: studies met minder spectaculaire resultaten verdwijnen iets vaker in een schuif, studies met sterke resultaten geraken gemakkelijker gepubliceerd. Twee modererende variabelen over methodologische kenmerken van de opgenomen effectstudies tonen ook een significante bijdrage. Wanneer een effectstudie een goede uitkomstmaat heeft gebruikt om het effect van een interventie te meten, dan ligt de OR lager dan bij studies waarbij de uitkomstmaat minder goed geoperationaliseerd is: 'Een slordige operationalisering van de uitkomstmaat veroorzaakt extra 'ruis' en dat kan leiden tot een overschatting van het ware effect van de interventie' (p. 75). De andere modererende variabele betreft het al of niet meenemen van de groep die uitvalt bij een interventie. Het resultaat van de meta-regressie bevestigt wat internationaal onderzoek daarover meldt. Wanneer een studie de uitvallers meegenomen heeft bij de effectevaluatie, dan ligt de OR voor de interventie een stuk lager dan wanneer de afhakers zijn weggelaten. Met andere woorden: 'Het uitsluiten van de afhakers leidt ertoe dat de recidivecijfers van het project gunstiger afsteken, vooral wanneer de controlegroep minder uitvallers kent' (p. 75). Het is dus belangrijk om uitvallers ook mee te nemen in effectonderzoek, zodat deze overschatting vermeden kan worden. Bij de dadergroepenkenmerken heeft de leeftijd van de deelnemers aan de interventie een impact op de gemeten OR. Interventies voor volwassenen geven een hogere OR dan interventies die gericht zijn op jeugdigen of jongvolwassenen. Met andere woorden: de effectonderzoeken in deze meta-analyse tonen dat de onderzochte interventies voor volwassenen minder recidiveprevalentie tot gevolg hebben in vergelijking met de interventies voor jongeren. De onderzoekers weten niet hoe dat komt en formuleren enkele mogelijke verklaringen (p. 75-76). Een laatste modererende variabele betreft een kenmerk van de geëvalueerde interventie: interventies gericht op resocialisatie hebben een hogere OR dan interventies gericht op individuele afschrikking. Dit vormt voor WARTNA en collega's 'een sterke aanwijzing [...] dat met pogingen tot resocialisatie van justitiabelen wel degelijk successen te behalen zijn' (p. 76).

In hun conclusies overlopen WARTNA, ALBERDA en VERWEIJ de belangrijkste bevindingen en plaatsen dan drie belangrijke kanttekeningen bij hun onderzoek. Ten eerste wijzen ze er zelf op dat aan de hand van de gebruikte methode niet kan worden geïdentificeerd 'welke processen met de interventies zijn in gang gezet en hoe deze hebben aangegrepen op de houding en het gedrag van de individuele deelnemers' (p. 13). In dat licht refereren ze aan bijkomend onderzoek dat onderweg is en waar wel antwoorden op dergelijke vragen zullen gegeven worden. Ook betwijfelen ze zelf of

hun databank met studies volledig en representatief is en dat de meta-analyse ook maar gericht is op één van de strafdoelen, naast meerdere andere (p. 13-14). Een derde kanttekening die ze formuleren, heeft te maken met de kwaliteit van evaluatieonderzoeken die ze in hun studie opgesnord hebben (p. 14-15). 'Alles overziend, kunnen we stellen dat de kwaliteit van de Nederlandse en Vlaamse effectstudies op het terrein van de strafrechttoepassing uiteenloopt, maar dat gemiddeld gesproken het niveau wel het een en ander te wensen overlaat' (p. 61).

Dit onderzoek is aan de hand van strikte methodologische criteria op een nauwgezette manier uitgevoerd en werpt aan de hand van recidiveprevalentie een eerste licht op de effectiviteit van strafrechtelijke interventies in Nederland en Vlaanderen. De auteurs vermelden dat bijkomend onderzoek in de pijplijn zit. Het is uitkijken naar wat bijkomende analyses aan nieuwe inzichten opleveren.

Afsluitend vermeld ik hier enkele reflecties en bedenkingen die bij het lezen van de meta-analyse en nadien kwamen bovendrijven.

Waarom nemen de auteurs informatie op afkomstig uit Nederland en Vlaanderen? Dit betreft de keuze tussen het taalgebied en het rechtsgebied. Nederland en Vlaanderen verschillen misschien meer van elkaar dan op het eerste gezicht blijkt. De taal is voor opname in de meta-analyse op zich geen criterium, want studies mogen in het Nederlands of in het Engels verschijnen (weliswaar slechts 3 Engelstalige publicaties, zie bijlage 3). De kans is groot dat daarmee qua taalvereiste zo goed als alle Nederlandse studies opgenomen worden. Dat daarmee selectie binnenkruipt voor wat aan Belgische zijde kan bestaan, wordt niet vermeld. Franstalige effectstudies binnen de Belgische context worden daardoor een blinde vlek voor dit onderzoek. Overigens zou de keuze voor een Nederlandse meta-analyse met uitsluitend Nederlands onderzoek niet tot veel uitval van effectstudies leiden. In de uiteindelijke meta-analyse is amper één Vlaamse studie opgenomen; in de ruimere literatuurstudie zitten drie Nederlandstalige Vlaamse studies. Dit weglaten zou weinig veranderen aan de globale resultaten, maar zou het geheel wel meer consistent maken qua rechtsgebied. Vlaanderen maakt (tot nader orde toch) immers deel uit van België; strafrechtelijke interventies zijn vaak het gevolg van rechtspraak op basis van rechtsregels die hetzelfde zijn in het noorden en het zuiden van het land, terwijl het Nederlandse rechtssysteem en strafrechtelijk beleid op dat vlak toch wel wat verschillen vertoont. De beleidskeuze voor de principes van *what works* in de sfeer van strafrechtelijke interventies biedt op dat vlak een sterke indicator. Daar waar in Nederland de beleidskeuze genomen is om volgens de principes van deze benadering interventies te organiseren, heeft de *what works* benadering in België nog niet iedereen in het beleid en de praktijk bereikt, laat staan overtuigd. De idee alleen al van een *evidence-based* beleid in de strafrechtsbedeling roept bij sommigen grote praktische bezwaren op en botst niet zelden op ideologische tegenargumenten. Wat dus vanzelfsprekend is in het ene land, is dat helemaal niet in het andere, zelfs al gaat het om buurlanden waarvan een deel van de bevolking dezelfde taal deelt. Maar laat dit geen kritiek zijn op de auteurs van de meta-analyse. Hun studie is veeleer een gelegenheid om onszelf als Vlaams en Belgisch criminoloog af te vragen of we zelf niet beter de hand aan de ploeg kunnen slaan en al het bestaand effectonderzoek in de Belgische context kunnen verzamelen. Dat zou ons dan weer snel leiden tot de vaststelling dat in België verbijsterend weinig effectonderzoeken over strafrechtelijke interventies bestaan, en dat is helemaal het geval voor onderzoek op basis van recidiveprevalentie.

Een bedenking die eveneens bij de inclusiecriteria voor het onderzoek aansluit, heeft te maken met het verkennende karakter van de meta-analyse en de daarmee samenhangende grote heterogeniteit van de opgenomen interventies in de meta-

analyse. Gegeven de heterogeniteit op het vlak van de onderzochte interventies (van pre-justitiële afhandeling tot reclasseringsinterventies van ex-gedetineerden), de publicatiedatum en de onderzochte populatie (volwassen daders, jongeren) dreigt de meta-analyse appels met peren te vergelijken. Dit gaat allemaal terug tot de vraag of het zinvol is om een overkoepelende vraag te stellen naar wat er in Nederland en Vlaanderen ‘werkt’, zonder op een scherpere manier te vertrekken van inclusiecriteria (bijvoorbeeld enkel interventies waarover na een bepaald jaar is gepubliceerd, enkel interventies voor een specifieke doelgroep of een bepaald type interventie, bv. enkel interventies aan de hand van de *what works* principes bij gedetineerden). Het bij elkaar brengen van effectstudies met een meting van de recidiveprevalentie is een sterk punt. Maar hoe deze werkende oefening inhoudelijk bijdraagt tot het identificeren van betekenisvolle verkennende interventies, is mij daarentegen niet helemaal duidelijk (per interventie is er immers al een bestaand effectonderzoek dat in principe in de meta-analyse is opgenomen; is er geen effectstudie, dan zal dit ook in de meta-analyse immers een blinde vlek blijven). Hoe uit deze meta-analyse op basis van zoveel heterogene interventies lessen kunnen worden getrokken naar de toekomst toe, blijft een vraag die niet echt duidelijk beantwoord is in de studie. We weten nu dat bepaalde interventies op basis van het gebruikte criterium ‘werken’ of in het verleden bij een evaluatie gewerkt hebben, maar wat moeten we hiermee wetenschappelijk en beleidsmatig naar de toekomst toe? Het vervolgonderzoek zal hierop wellicht antwoorden bieden.

Bovenstaande bedenkingen zijn echter details. Ik heb met veel plezier deze studie gelezen en heb er veel uit geleerd. Tevens kijk ik met een beetje wetenschappelijke afgunst naar wat in Nederland allemaal gebeurt in de sfeer van de strafrechtsbedeling (zonder dat te willen idealiseren: realisme en nuchterheid zijn geboden, zeker in een emotioneel dooraderd terrein als het strafrecht). De sterke aandacht voor degelijk empirisch onderzoek lijkt in Nederland een normaal onderdeel te vormen van het beleidsproces. Onderzoek naar effecten van interventies is breed ingeburgerd en zelfs geïstitutioniseerd (de Erkenningscommissie Gedragsinterventies Justitie). De voedingsbodem voor een op *knowledge* en *evidence* gebaseerd strafrechtelijk beleid is bij onze noorderburen aanwezig. Dat ligt in België anders. Nog steeds ontbreken criminografische basisgegevens zoals nationale recidivecijfers, blijft de koppeling tussen databanken binnen Justitie, laat staan tussen Justitie en andere ministeries, een zeer moeilijke opgave, en is er geen traditie van effectonderzoek bij strafrechtelijke interventies, met soms zelfs verzet tegen evaluatieonderzoek. Wil dat dan zeggen dat we in België niet goed bezig zijn? Dat er hier niets ‘werkt’? Nee, natuurlijk niet. We weten het eenvoudigweg voor de overgrote meerderheid van justitiële interventies niet. Middelen inzetten zonder dat we weten of het opbrengt wat we ermee beogen, roept veel vragen op. De studie van WARTNA en collega’s herinnert ons als betrokkenen (wetenschappers, praktijkwerkers, beleidsmedewerkers) in de sfeer van het strafrecht eraan dat het anders kan. Dat het wel degelijk kan om na te gaan of de doelen van interventies bereikt worden en dat op die manier de effectiviteit van strafrechtelijke ingrepen kan gecapteerd worden. Aan ons om daar lessen uit te trekken.