



Universiteit
Leiden
The Netherlands

Onderzoek naar effectiviteit van straffen: een natuurlijk veldexperiment met de Bijleveld-Factor

Keijser, J.W. de; Elffers, H.; Rahim, D.; Rodermond, E.; Hendriks, J.; Slotboom, A.-M.; ... ; Huisman, W.

Citation

Keijser, J. W. de, Elffers, H., & Rahim, D. (2026). Onderzoek naar effectiviteit van straffen: een natuurlijk veldexperiment met de Bijleveld-Factor. In E. Rodermond, J. Hendriks, A. - M. Slotboom, C. van den Berg, S. van de Weijer, & W. Huisman (Eds.), *Kilometers maken, bergen verzetten* (pp. 121-134). Den Haag: Boom Juridisch. Retrieved from <https://hdl.handle.net/1887/4292703>

Version: Accepted Manuscript

License: [Leiden University Non-exclusive license](#)

Downloaded from: <https://hdl.handle.net/1887/4292703>

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

Onderzoek naar effectiviteit van straffen

Een natuurlijk veldexperiment met de Bijleveld-Factor

Jan de Keijser, Henk Elffers, Dania Rahim¹

Inleiding

Hoop op effectiviteit

In 2023 werden 807.000 misdrijven geregistreerd. In datzelfde jaar werden 88.000 straffen opgelegd, waaronder 34.000 vrijheidsstraffen en 26.000 taakstraffen (Moolenaar, 2023). Uiteraard staat de rechter een nog veel groter sanctie-arsenaal tot zijn beschikking met, onder andere, de geldboete, maatregelen (zoals ISD en TBS), bijkomende straffen (zoals ontzegging van de rijbevoegdheid) en tal van bijzondere voorwaarden bij voorwaardelijk opgelegde straffen (waaronder herstel van schade, elektronisch toezicht, behandelverplichting). Van al die strafrechtelijke interventies wordt gehoopt dat – naast overwegingen van vergelding – de sanctie een beoogd effect sorteert. Die hoop is zeer terecht, gezien het leed en de enorme maatschappelijke sociale en economische kosten die gepaard gaan met de getallen zojuist genoemd. Die beoogde, in de toekomst gelegen, effecten worden gezocht in de algemene en speciale preventie van criminaliteit. Strafrechtelijke interventies blijken echter niet zo goed te zijn in het voorkómen van nieuwe misdrijven.

Onderzoek naar effectiviteit van straf is moeilijk

Dat die effecten vaak tegen vallen, blijkt bijvoorbeeld uit de recidivecijfers na verschillende hoofdstraffen. Dan kijken we dus naar het speciaal preventieve effect van de interventies ten gevolge van resocialisatie/rehabilitatie of individuele afschrikking. Zo is het recidivepercentage drie jaar na een (kortdurende) gevangenisstraf 60 procent, na een taakstraf is dat ongeveer 40 procent, en na een geldboete recidiveert ruim 20 procent binnen drie jaren (Wermink et al., 2022). Daarmee is nog niet gezegd dat de ene straf effectiever is dan de andere. Immers de gestraften, hun achtergronden en de door hen gepleegde delicten variëren nogal, niet alleen binnen de categorieën van opgelegde straffen, maar vooral ook daartussen. Goed onderzoek naar de relatieve effectiviteit van strafrechtelijke interventies is dan ook niet gemakkelijk.

Om die relatieve effectiviteit te meten, zou men het liefst zuivere experimenten uitvoeren met een experimentele groep en een controlegroep, of in ieder geval minstens twee groepen. De proefpersonen worden in een zuiver experiment *at random* toegewezen aan de ene dan wel de andere groep. Op die manier zullen de te vergelijken groepen op allerlei andere, mogelijk met de uitkomstmaat samenhangende, aspecten goed vergelijkbaar zijn.

¹ Jan de Keijser is hoogleraar criminologie aan de Universiteit Leiden

Henk Elffers is emeritus hoogleraar empirische bestudering van de strafrechtpleging aan de Vrije Universiteit Amsterdam en gastonderzoeker bij het Nederlands Studiecentrum Criminaliteit en Rechtshandhaving NSCR te Amsterdam

Dania Rahim studeert criminologie en rechtsgeleerdheid aan de Vrije Universiteit Amsterdam

Als we zuiver experimenteel onderzoek willen doen naar de relatieve effectiviteit van sancties, zouden plegers van misdrijven die veroordeeld worden door de rechter vervolgens bijvoorbeeld *at random* een gevangenisstraf of een taakstraf moeten krijgen. Een bepaalde tijd na executie van de straf worden dan de recidivepercentages (of een andere uitkomstmaat) vergeleken.

Het lijden van Bijleveld

In haar oratie aan de VU “Mens, durf te meten” verzuchtte Catrien Bijleveld (2003) dat de Nederlandse straftoemettingspraktijk zo niet-*random* is als maar kan (p. 21). Sancties worden door de rechter niet willekeurig gekozen, maar, zoals de wet voorschrijft, juist zo veel mogelijk geïndividualiseerd. Kortom, in de praktijk zien we geen groepen gestraften die op alle mogelijke kenmerken nagenoeg gelijk zijn behalve op de sanctie die zij voor hun kiezen kregen. Onze kennis over de effectiviteit van straffen schiet dan ook tekort. Als de rechter bij het maken van een optimale keuze voor de ene straf of de andere zich wil laten leiden door de te verwachten verbetering die de straf teweeg brengt, dan tast hij goeddeels in het duister. “Het kan toch niet zo zijn dat we slechtzind zijn en opzettelijk slechtzind blijven als het om de genezende werking van straffen gaat?”, vroeg Bijleveld zich daarom niet zonder gevoel voor drama af in haar oratie (p. 29).

Gelukkig liet Bijleveld het niet louter bij een wanhoopskreet, maar zij toonde ons de methodologische route naar een betere (onderzoeks-)toekomst: “Het lijkt mij om al deze redenen van groot belang een begin te maken met het op wetenschappelijke, experimentele wijze onderzoeken van de relatieve effectiviteit van een aantal strafsoorten” (p. 21). Dat een zuiver experimentele studie met echte plegers van misdrijven en echte straffen in de praktijk geen haalbare (misschien zelfs geen wenselijke) kaart is, voegt Bijleveld daar zelf onmiddellijk aan toe. Maar zij wijst op de statistische mogelijkheden om groepen op zo veel mogelijk relevante kenmerken achteraf te matchen (bijvoorbeeld door *propensity score matching*). Ook noemt ze de mogelijkheid om veroordeelden die daarvoor toestemming hebben verleend door de rechter *at random* te laten toewijzen aan een van twee straffen die absoluut gelijkwaardig worden gezien (p.29). Uiteraard zijn er tal van andere mogelijkheden – die Bijleveld verder onbenoemd laat – om een experimentele onderzoeksmethode enigszins te benaderen. Waar het Bijleveld ongetwijfeld in hoofdzaak om ging, was haar experimentele hartenkreet luidkeels te doen horen: “We moeten ergens beginnen als we verder willen komen. Mens durf te meten.” (p. 30).

Een natuurlijk veldexperiment

Het Bijleveld-effect

De Amsterdamse oratie van Bijleveld ligt intussen meer dan 20 jaar achter ons. Heeft sindsdien haar niet mis te verstane roep om meer en methodologisch deugdelijker experimenteel onderzoek naar de relatieve effectiviteit van strafrechtelijke interventies weerklink gevonden?

Onbedoeld heeft Bijleveld met haar oratie de grondslag gelegd voor een natuurlijk veldexperiment. In een natuurlijk experiment maakt de onderzoeker gebruik van een gebeurtenis of situatie die zich in de werkelijkheid voordoet en waarvan een effect op een bepaalde uitkomstmaat wordt verwacht. Die gebeurtenis was dan niet bedoeld als experimentele interventie, maar de opportunistische wetenschapper ziet daar wel een kans in (zie bijvoorbeeld Van der Werff, 1979; Elffers et al., 2007). In een veldexperiment speelt de experimentele interventie zich af in een realistische, niet door de onderzoeker gecreëerde context: het veld. Welnu, in ons natuurlijk veldexperiment is de gebeurtenis waarvan wij enig effect verwachten, het verschijnen van de oratie van Bijleveld in het jaar 2003. Het veld betreft de publicerende academische gemeenschap. De onderzoeksvraag luidt: Heeft de oproep van Bijleveld in haar oratie in 2003 geleid tot meer experimenteel onderzoek naar effectiviteit van strafrechtelijke interventies?

Design

Ons experiment is een natuurlijk veldexperiment met de Bijleveld-factor. Nu zouden wij kunnen hypothetiseren dat de oratie van Bijleveld een wereldwijd effect heeft gehad op onderzoek naar relatieve effectiviteit van strafrechtelijke interventies. Hoewel wij Bijleveld die eer zonder meer zouden gunnen, lijkt het logischer om het experiment te beperken tot onderzoek dat betrekking heeft op effectiviteit van strafrechtelijke interventies in Nederland. De oratie was immers een Nederlandstalige publicatie en de ceremonie een Nederlandse aangelegenheid.² Omdat er inmiddels ruim 20 jaar verstreken is na verschijnen van die oratie, onderzoeken wij een even grote periode die daar aan vooraf ging. Het simpelweg vergelijken van het aantal gepubliceerde experimentele effect-studies in Nederland vóór 2003 met het aantal vanaf 2003 zou onze studie echter op zijn best de schamele score ‘2’ opleveren op de *Maryland Scientific Methods Scale* (Farrington et al., 2002). Om enige causaliteit neer te kunnen leggen bij de Bijleveld-factor, is dan ook een vergelijkbare controlegroep nodig. Die controlegroep hebben wij gevonden in onderzoekspublicaties in de dezelfde periode die betrekking hebben op de relatieve effectiviteit van strafrechtelijke interventies bij onze directe zuiderburen: België, er daarbij van uitgaand dat Bijlevelds oratie in België minder zal zijn opgemerkt dan in Nederland.³

Met het zojuist beschreven ontwerp, luidt onze verwachting – ten faveure van het Bijleveld-effect – dat er in Nederland sinds 2003 een toename is te constateren van experimentele studies naar de relatieve effectiviteit van interventies en dat die toename niet, althans niet in dezelfde mate, waarneembaar is in België.

² Een zoektocht in Web Of Science naar de vertaalde titel “*Human, dare to measure!*” leverde dan ook 0 hits op.

³ Google Scholar vermeldde op 22 oktober 2024 15 studies die de oratie aanhalen (exclusief de 3 keer dat Bijleveld dat zelf doet). Daaronder zijn er 10 in Nederlandse tijdschriften of boeken (waarvan 1 door een Belgische auteur), 1 in Belgische, en 4 in Engelse (door voornamelijk Nederlandse auteurs).

Afhankelijk variabele

De afhankelijke variabele is het aantal onderzoekspublicaties waarin op experimentele wijze de relatieve effectiviteit van interventies is onderzocht. Het adjectief ‘experimenteel’ hebben wij weinig restrictief opgevat. Uiteraard dient het om kwalitatief goede studies te gaan, maar ‘experimenteel’ in onze opzet is (noodzakelijkerwijs) ruimer dan zuivere experimenten en *randomized control trials*. Onder experimentele studies scharen wij ook: natuurlijke experimenten, studies waarin gebruik is gemaakt van vergelijkbare groepen of nog meer gesofisticeerd van post hoc *sample matching* zoals *propensity score matching*, en studies waarin gewerkt wordt met *instrumentele variabelen*.⁴

Populatie-onderzoek

Onze studie is een populatiestudie. De populatie betreft alle publicaties in Nederland en in België die aan onze inclusie-criteria (zie hieronder) voldoen in de periode 20 jaar vóór en 20 jaar ná de Bijleveld-interventie. Er heeft dan ook geen steekproeftrekking plaatsgevonden. Argumenten om inferentiële statistiek te bedrijven, die gelieerd zijn aan aspecten van ecologische validiteit, zijn hier dan ook niet van toepassing, en wij zullen ons beperken tot beschrijvende statistiek

Zoekstrategie

In de Nederlandstalige en in de Engelstalige wetenschappelijke literatuur is gezocht naar relevante publicaties die betrekking hebben op relatieve effectiviteit van strafrechtelijke interventies in Nederland of in België. ‘Interventie’ werd daarbij breed opgevat, en kon dus naast straffen of maatregelen, ook bijvoorbeeld een bepaald behandelprogramma in de gevangenis, bevatten. Bij het concept ‘effectiviteit’ is ook een brede opvatting gehanteerd, zolang het om een heldere uitkomstmaat ging die volgde uit het doel van een interventie. Zoals hierboven uitgelegd hebben we studies met tal van onderzoeksmethoden geïnccludeerd die een experimentele vergelijking van interventies benaderen.

Er is gezocht binnen de tijdvakken van 20 jaar voor 2003 (01-01-1984 t/m 31-12-2003) en 20 jaar na 2003 (01-01-2004 t/m 31-12-2023). Voor de digitale zoektocht in Web Of Science is onderstaande zoekleutel geconstrueerd. Daarnaast is voor diezelfde periode handmatig gezocht in de inhoudsopgaven van Tijdschrift voor Criminologie, Panopticon, en Recht der Werkelijkheid, en in het Repository van het WODC. In de verkregen resultaten is ook gekeken naar de literatuurverwijzingen (sneeuwbalmethode) en er is in een tweetal overzichtsstudies gekeken of er nog aanvullende relevante verwijzingen te vinden waren (Boschman et al., 2023; Wartna et al., 2013).

⁴ Voor een handzame beschrijving van deze benaderingen, zie Boschman et al. (2023).

Zoeksleutel Nederlandstalige literatuur

(Straf OR (strafrecht* AND interventie*))*

AND

(maatregel OR sanctie OR gevangenis OR taakstraf OR “strafrechtelijke interventies” OR straftoemeting OR boete* OR strafrecht* OR recidive*)*

AND

(effect OR impact OR relatie*)*

AND

(Experiment OR propensity OR psm OR “instrumentele variabele” OR experimenteel OR (quasi experiment OR quasi-experiment OR quasi) OR (“randomized controlled trial” OR rct) OR veld OR (controlegroep OR controle groep) OR gerandomiseerd OR randomisering)*

AND

(Nederland OR België)

Zoeksleutel Engelstalige literatuur

(punishment OR sanctions OR prison OR imprisonment OR criminal justice interventions OR criminal law interventions OR recidivism*)*

AND

(effect OR impact OR relation*)*

AND

(Experiment OR propensity OR psm OR “instrumental variable” OR experimental OR (quasi experiment OR quasi-experiment OR quasi) OR (“randomized controlled trial” OR rct) OR random* OR “control group”)*

AND

(Netherlands OR Dutch OR Belgium)

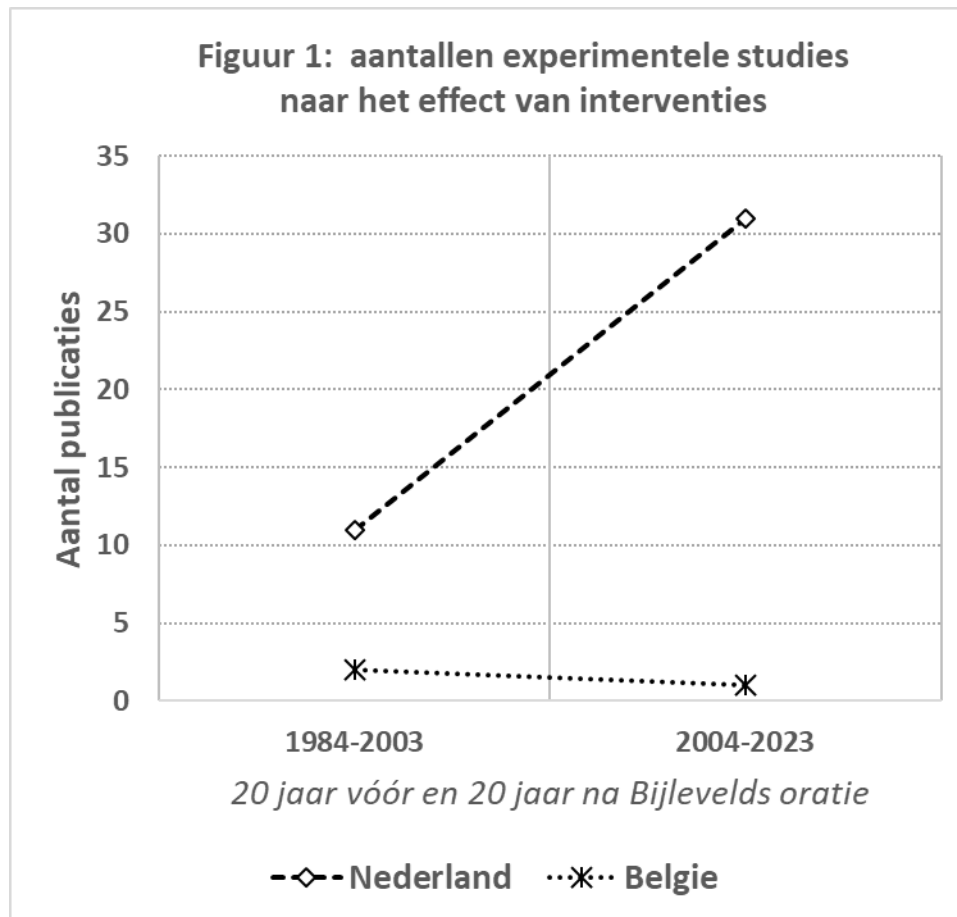
Resultaten

Selectie studies

De zo gevonden 415 artikelen (372 via Web of Science, 43 handmatig) zijn door de derde auteur op grond van titel en abstract, en indien dat onduidelijk bleef, door het hele artikel te bekijken, gescreend, of ze inderdaad voldeden aan het inclusiecriteria dat ze als onderwerp de effectiviteit van sancties – in bovengenoemde ruime zin – vergelijken. Daarbij is ook gecheckt of het artikel inderdaad over Belgische of Nederlandse straftoedeling gaat. Dat leidde tot het verwerpen van 344 artikelen. Vervolgens hebben de eerste en tweede auteur onafhankelijk van elkaar de 71 wel geselecteerde artikelen, op basis van titel en abstract en zo nodig op basis van de gehele tekst, gescreend of er inderdaad sprake is - in bovengenoemde ruime zin - van een experiment. Over 10 artikelen waren beiden het initieel niet eens, en werd er gediscussieerd tot er consensus was. Over de overige 61 waren beiden het wel eens. Dit resulteerde in het verwerpen van 26 artikelen, terwijl er 45 geïnccludeerd werden. De door ons geïnccludeerde studies staan in Appendix 1 vermeld. Ze staan daar per land op chronologische volgorde weergegeven.

Effect van de Bijleveld-Factor

In onderstaande Figuur 1 staat het resultaat gevisualiseerd van ons onderzoek naar de Bijleveld-Factor. Afgebeeld staan de aantallen publicaties voor Nederland en voor België, in de periode van 20 jaar vóór de oratie van Bijleveld en een even lange periode daarna. De lijntjes zijn bedoeld om het oog te geleiden.



We kunnen onomwonden concluderen dat Bijlevelds oratie een stevige invloed heeft gehad. In Nederland is het aantal publicaties dat wij als experimenteel onderzoek naar het effect van strafrechtelijke interventies klasseerden, gestegen van 11 naar 31, dus bijna verdrievoudigd, terwijl er in de Belgische controlegroep geen sprake is geweest van een vergelijkbare stijging, daar viel zelfs een afname waar te nemen. Wij schrijven die Nederlandse toename uiteraard volledig toe aan het Bijleveld-effect.

Kan Bijleveld nu dan met een gerust hart met pensioen? *Mission accomplished!*, zou ze wellicht geneigd zijn te roepen na het lezen van onze analyse. Dat komt ons helaas wat prematuur over. Er is nog veel werk aan de winkel. Zo is er weliswaar verhoudingsgewijs een flinke toename van het aantal experimentele studies (+182%), maar in absolute zin gaat het om een betrekkelijk gering aantal: 20 meer in twee decennia, waarbij ook nog meer dan eens dezelfde onderzoekers betrokken zijn. Bovendien, veel van de studies die we hebben gevonden zijn methodologisch niet de

mooist denkbare. Tabel 1 laat zien welk type experimenten naar het effect van interventies in Nederland voorkwamen.

Tabel 1. Soorten experimenten (Nederland)

Periode	Quasi-experiment of natuurlijk experiment	Sophisticated experiment (post hoc matching, instrumentele variabelen, propensity score matching)	Zuiver experiment (RCT)
1984-2003	11	0	0
2004-2023	14	15	2

Echte zuivere experimenten (RCTs) zijn zeldzaam, vaak zien we quasi-experimenten, soms zelfs zonder enige matching van experimentele en controlegroep. Het houdt dus niet over, zij het dat er beslist sprake is van vooruitgang. In die zin is Bijlevelds oproep uit 2003 nog steeds actueel.

Nabrand: nog twee noten te kraken

Ook al zouden we met de winst in aantallen en typen experimenten tevreden zijn, dan nog zijn we niet in veilige haven. Twee methodologische noten willen wij nog kraken. Het zijn noten die vooralsnog tamelijk onderbelicht zijn in onderzoek naar de relatieve effectiviteit van interventies. De eerste noot is er een van externe validiteit. Het betreft de verarming van vergelijkingsmateriaal ten gevolge van inkrimping van de vergelijkingsbasis, op voorhand (zoals bij *zuivere experimenten*) of achteraf (zoals bij *propensity score matching*). De tweede noot betreft de conclusies die wel en juist niet verbonden kunnen worden aan een experimentele vergelijking van twee interventies. In hoeverre betekent een verschil in recidive dat de ene sanctie *effectiever* is dan de andere?

Noot 1: Verarming van de vergelijkingsbasis

Als we een realistisch zuiver experiment naar straftoemeting doen, vragen we rechters om bij de straftoemeting het lot te laten bepalen welk van twee in het onderzoeksdesign betrokken straffen wordt uitgesproken. Dat zijn dan straffen die als beide passend in het geval van de verdachte worden beschouwd. Rechters houden daar niet van, zij vinden dat tot de straftoemeting onscheidbaar behoort de noodzaak na te gaan welke straf voor juist deze verdachte het meest aangewezen is. Anders gezegd, zij wensen zich eerst te overtuigen of de beide mogelijke straffen wel equivalent zijn voor juist deze verdachte. Dat is in de ogen van veel rechters echt niet altijd voor elke verdachte het geval. Als voorbeeld: een binnen twee jaar ten tweede male voor autodiefstal voor de rechter staande verdachte kan tot een taakstraf van 180 uur of een gevangenisstraf van drie

maanden worden veroordeeld.⁵ Veel rechters vinden echter dat sommige daders minder in aanmerking komen voor gevangenisstraf (bijvoorbeeld omdat ze dan een net begonnen opleiding niet kunnen afmaken). Die daders willen ze dan geen gevangenisstraf opleggen. Voor anderen vinden ze wellicht dat die beter geen taakstraf kunnen krijgen (bijvoorbeeld omdat een eerdere taakstraf niet correct werd uitgevoerd, of vanwege een antagonistische houding tegenover de reclassering). Dan willen ze deze laatste daders géén taakstraf opleggen. Eigenlijk zijn er drie categorieën daders in de ogen van de rechter: degenen voor wie gevangenisstraf niet in aanmerking komt, degenen voor wie een taakstraf niet in aanmerking komt, en een derde groep van degenen voor wie beide wèl in aanmerking komen, en als equivalent worden beschouwd. Als we vervolgens een zuiver experiment (RCT) doen dat *de facto* alleen op die derde groep gericht is (expliciet of impliciet), dan moeten we ons realiseren dat de uitkomsten van zo'n experiment ook alleen voor die groep gelden. De externe validiteit van de meting is niet vanzelfsprekend. Formulering van een resultaat in de vorm van "gevangenisstraf is minder effectief dan taakstraf" is dan te kort door de bocht, die uitspraak moet gekwalificeerd worden: dat geldt alleen voor de groep mensen voor wie, in de ogen van de oordelende rechter, beide straffen in aanmerking komen. Nu is dit probleem niet zo erg als de twee groepen die niet voor een van beide straffen in aanmerking komen, beperkt in omvang zijn. Dan is de formulering van het resultaat zoals boven verwoord strikt genomen wel onjuist, maar het scheelt niet zo veel. Maar als een van die of beide groepen omvangrijk zijn, is het andere koek. Wij pleiten er daarom voor bij elk experiment expliciet te verantwoorden op welke groep het experiment betrekking heeft en te kwantificeren hoe groot de groepen zijn die buiten boord worden gehouden. Niet alleen de grootte van de groepen is van belang, ook de uitkomstmaat, in vergelijking met de groep die wel aan loting onderhevig was. Zijn de mensen die een taakstraf krijgen, omdat ze niet in aanmerking kwamen voor gevangenisstraf anders dan veroordeelden uit de RCT-groep die een taakstraf krijgen? En *mutatis mutandis* voor de naar de gevangenis gestuurden die niet voor taakstraf in aanmerking kwamen. Zo ja, dan werpt dat wel een schril licht op de betekenis van de bevindingen die op de RCT zijn gebaseerd. Het is dus zaak te rapporteren over de stand van zaken in de uitvallergroepen. We moeten constateren dat dat nagenoeg nooit gebeurt! OK, fijn dat Bijleveld ons tot zuivere experimenten aanzet, prachtig, maar je kunt de externe validiteit van de resultaten pas goed interpreteren in het licht van rapportage over uitvallende groepen.

Dit probleem van de verschraling van de vergelijkingsbasis doet zich niet alleen voor bij zuivere experimenten. Het is ook aan de orde als we de vergelijking baseren op achteraf matchen van experimentele en controlegroep, bijvoorbeeld door *propensity score matching*: dan zijn er in het algemeen ook proefpersonen die niet gematcht

⁵ Die straffen gelden in de strafvorderingsrichtlijnen van het openbaar ministerie als equivalent voor dit delict.

kunnen worden. In het jargon heet dat dat ze buiten het *common support* vallen. De literatuur waarschuwt dat het er niet teveel mogen zijn (zonder standaarden te formuleren), maar analoog aan het bovenstaande dienen we ook te kijken naar de scores van deze uitvallers in vergelijking met hen die wel in het *common support* vallen. Ook dat gebeurt nagenoeg nooit, en verdient beter! (Elffers, 2017).

Noot 2: Betekent een verschil in recidive dat de ene sanctie effectiever is dan de andere? Studies waarin verschillende strafrechtelijke interventies op een methodologisch verstandige manier worden vergeleken op een bepaalde uitkomstmaat, laten nog niet *per se* zien dat de ene interventie ook *effectiever* is dan de andere. Zo laten Wermink en collega's in hun prachtige instrumentele variabele studie zien dat korte gevangenisstraffen tot meer recidive leiden dan niet-vrijheidsbenemende straffen (Wermink et al., 2022, p. 164). Maar betekent dat dat niet-vrijheidsbenemende straffen – denk bijvoorbeeld aan taakstraffen – effectiever zijn in het terugdringen van recidive dan gevangenisstraf? Die conclusie (die Wermink et al. zelf overigens niet trekken) lijkt voor de hand te liggen, maar is ons een brug te ver. Het gevonden verschil kan immers net zo goed het gevolg zijn van verschillen in de mate waarin gestraften onbedoelde negatieve effecten van interventies ondervinden. De gevangenisstraf heeft veel onbedoelde en indirecte negatieve effecten, onder andere op huisvesting, werk en (familie-)relaties die samenhangen met een verhoogde kans op recidive (zie bijv. Dirkzwager et al. 2009; Doekhie, 2019; Ramakers, 2014). Het niet opleggen van gevangenisstraf maar wel een andere, niet-vrijheidsbenemende interventie, zorgt er dan voor dat dergelijke (indirecte) negatieve onbedoelde effecten goeddeels uitblijven met als gevolg minder recidive na de niet vrijheidsbenemende interventie. Zo kan het bestaan dat vergelijking op de uitkomstmaat 'recidive' net zozeer iets zegt over de relatieve mate waarin negatieve gevolgen optreden of uitblijven als over de positieve (beoogde) effectiviteit in termen van terugdringen van recidive. Zonder controlegroep 'geen interventie' is dat pleit niet te beslechten.

De volgende 20 jaar?

Laten we hopen dat de criminologische en strafrechtelijke gemeenschap ook de komende jaren de wijze woorden van Catrien Bijleveld tot zich laat doordringen. Mogen we de volgende 20 jaar weer een verdrievoudiging van het aantal experimentele evaluaties van strafrechtelijke interventies tegemoet zien? Of krijgen we in haar afscheidsrede een heel andere aansporing om ons op te bezinnen?

Literatuur

- Bijleveld, C.C.J.H. (2003). *Mens, durf te meten!* Den Haag: Boom Juridische Uitgevers.
- Boschman, S., Verweij, S., Weijters, G. (2023). *Korte vrijheidsstraffen: Een literatuuronderzoek naar het bereiken van strafdoelen met korte vrijheidsstraffen ten opzichte van andere straffen*. WODC.
- Dirkzwager, A., Nieuwbeerta, P., & Fiselier, J. P. S. (2009). Onbedoelde gevolgen van vrijheidsstraffen; een literatuurstudie. *Tijdschrift voor Criminologie*, 51, 21-41.
- Doekhie, J. V. O. R. (2019). *Dimensions of desistance: A qualitative longitudinal analysis of different dimensions of the desistance process among long-term prisoners in the Netherlands*. Leiden Universiteit Leiden.
- Elffers, H. (2017) Quasi-experimentele criminologische effectstudies met propensity score matching. *Tijdschrift voor Criminologie* 59/1-2, 52-66.
- Elffers, H., Keijser J.W. de, Koppen P.J. van & Haeringen L. van (2007), Newspaper juries: A field experiment concerning the effect of information on attitudes towards the criminal justice system, *Journal of Experimental Criminology* 3, 163-182.
- Farrington, D., Gottfredson, D., Sherman, L.W., & Welsh, B. (2002). Maryland Scientific Methods Scale. In L.W. Sherman et al. (Eds.), *Evidence-based crime prevention* (13-21). Routledge.
- Moolenaar, D.E.G. (red.)(2024). *Criminaliteit en rechtshandhaving 2023: Ontwikkelingen en samenhangen*. WODC/CBS/Raad voor de rechtspraak.
- Ramakers, A. (2014). *Barred from employment? A study of labor market prospects before and after imprisonment*. Universiteit Leiden.
- Van der Werf (1979). *Speciale preventie: Verslag van een onderzoek naar het speciaal-preventieve effect van het strafproces en van het ondergaan van een korte vrijheidsstraf bij vermogensdelinquenten, agressieve delinquenten en verkeersdelinquenten*. Staatsuitgeverij.
- Wartna, B., Alberda, D., & Verweij, S. (2013). Een meta-analyse van Nederlands recidiveonderzoek naar de effecten van strafrechtelijke interventies. *Tijdschrift voor Criminologie* 55, 3-23.
- Wermink H.T., Blokland A.A.J., Tollenaar N., Been J., Schuyt P.M. & Apel R. (2022), Een oude vraag opnieuw onderzocht: effecten van korte gevangenisstraffen ten opzichte van niet-vrijheidsbenemende straffen op recidive op basis van een instrumentele variabele benadering, *Tijdschrift voor Criminologie* 64, 147-169.

Appendix 1: Geïnccludeerde studies

Deze bijlage geeft de titelbeschrijvingen van de geïnccludeerde studies, in vier blokken: Nederland 1984-2003, Nederland 2004-2023, België 1984-2003, België 2004-2023. Binnen elk blok staan de publicaties gerangschikt naar jaartal van publicatie, en daarbinnen op auteursnaam.

Nederlandse studies 1984-2003

- Werff, C. van der (1986). *Recidive 1977: Recidive van in 1977 wegens misdrijf veroordeelden en niet vervolgd*. Wetenschappelijk Onderzoek- en Documentatiecentrum. <http://hdl.handle.net/20.500.12832/1012>
- Kommer, M. M. (1987). *De analyse van recidivepatronen; een methodisch-technische verkenning*. Sociologisch Instituut Leiden en Wetenschappelijk Onderzoek- en Documentatiecentrum.
- Laan, P. H. van der & Essers, A. A. M. (1990). *De Kwartaalcursus en recidive*. Wetenschappelijk Onderzoek- en Documentatiecentrum. <http://hdl.handle.net/20.500.12832/1044>
- Kruissink, M. & Verwers, C. (1991). *Afhandeling van winkeldiefstal via de Halt-procedure: Evaluatie van een Rotterdams experiment*. Wetenschappelijk Onderzoek- en Documentatiecentrum. <http://hdl.handle.net/20.500.12832/836>
- Spaans, E. C., & Doornhein, L. (1991). *Evaluatie-onderzoek jeugdreclassering: de effectmeting*. WODC Reeks Onderzoek en Beleid 112. Arnhem: Gouda Quint. https://repository.wodc.nl/bitstream/handle/20.500.12832/1057/ob112-volledige-tekst_tcm28-78563.pdf?sequence=1&isAllowed=y
- Laan, P. H. van der (1992). Experimenteren met alternatieve sancties voor jeugdigen. *Kind en adolescent*, 13(1), Artikel 36. <https://doi.org/10.1007/BF03060446>
- Spaans, E. C. (1994). *Appels en peren: een onderzoek naar de recidive van dienstverleners en kortgestraften*. Arnhem: Gouda Quint.
- Wartna, B. S. J. (1995). *Dagdetentie en recidive. Een vervolgstudie naar de resultaten van het experiment in Rotterdam*. Den Haag: WODC. https://repository.wodc.nl/bitstream/handle/20.500.12832/2908/ov-1995-16-volledige-tekst_tcm28-78045.pdf?sequence=1&isAllowed=y
- Bosker, O. J. (1997). *Snelrecht: de generaal en speciaal preventieve effecten van sneller straffen*. Thesis fully internal (DIV), Rijksuniversiteit Groningen. [Dissertation. s.n.]. <https://research.rug.nl/en/publications/snelrecht-de-generaal-en-speciaal-preventieve-effecten-van-snelle>
- Spaans, E. (1997). *Bootcamps and electric monitoring: the Dutch experience*. Den Haag: WODC. https://repository.wodc.nl/bitstream/handle/20.500.12832/2930/ov-1997-07-full-text_tcm28-78111.pdf?sequence=1

Schippers, G. M., Hurk, A. A. van den, Breteler, M. H. & Meerkerk, G. J. (1998). Effectiveness of a drug-free detention treatment program in a Dutch prison. *Substance use & misuse*, 33(4), 1027–1046.
<https://doi.org/10.3109/10826089809056253>

Nederlandse studies 2004-2023

Beijersbergen, K. A., & Wartna, B. S. J. (2007). *Recidivemeting Glen Mills*. Den Haag: WODC. https://repository.wodc.nl/bitstream/handle/20.500.12832/329/fs07-7-voleldige-tekst_tcm28-69043.pdf?sequence=1&isAllowed=y

Boom, A. ten, Ferwerda, H. & Leiden, I. van (2007). Een pioniersstudie in justitieland: De evaluatie van halt in een experimentele setting. *Tijdschrift voor Criminologie*, 44(1), 33-44.

Nieuwbeerta, P., Blokland, A. A. J. & Nagin, D. (2009). Assessing the Impact of First-Time Imprisonment on Offenders' Subsequent Criminal Career Development: A Matched Samples Comparison. *Journal of Quantitative Criminology*, 25(3), 227-257.
<https://doi.org/10.1007/s10940-009-9069-7>

Wermink, H., Blokland, A., Nieuwbeerta, P. & Tollenaar, N. (2009). Recidive na werkstraffen en na gevangenisstraffen: een gematchte vergelijking. *Tijdschrift voor Criminologie*, 51(3), 211-227.

Wermink, H., Blokland, A., Nieuwbeerta, P., Nagin, D. & Tollenaar, N. (2010) Comparing the effects of community service and short-term imprisonment on recidivism: a matched samples approach. *Journal of Experimental Criminology*, 6(3), 325–349.
<https://doi.org/10.1007/s11292-010-9097-1>

Snodgrass, G. M., Blokland, A. A., Haviland, A., Nieuwbeerta, P. & Nagin, D. S. (2011). Does the time cause the crime? An examination of the relationship between time served and reoffending in the Netherlands. *Criminology*, 49(4), 1149-1194.
<https://doi.org/10.1111/j.1745-9125.2011.00254.x>

Tollenaar, N., & Van der Laan, A. M. (2012). *Effecten van de ISD-maatregel Technische rapportage*. Den Haag: WODC. Memorandum 2012-2.
https://repository.wodc.nl/bitstream/handle/20.500.12832/898/memorandum-2012-2-volledige-tekst-okt2012_tcm28-71993.pdf?sequence=1&isAllowed=y

Aarten, P. G. M., Denkers, A. J. M., Borgers, M. J. & Laan, P. H. van der (2014). Suspending reoffending? Comparing the effects of suspended prison sentences and short-term imprisonment on recidivism in the Netherlands. *European Journal of Criminology*, 11(6), 702- 722. <https://doi.org/10.1177/1477370814523402>

Asscher, J. J., Deković, M., Manders, W., van der Laan, P. H., Prins, P. J., van Arum, S. & Dutch MST Cost-Effectiveness Study Group. (2014). Sustainability of the effects of multisystemic therapy for juvenile delinquents in The Netherlands: Effects on delinquency and recidivism. *Journal of Experimental Criminology*, 10, 227-243. DOI 10.1007/s11292-013-9198-8

- Hoogsteder, L. M., Kuijpers, N., Stams, G. J. J. M., Horn, J. E. van, Hendriks, J. & Wissink, I. B. (2014). Study on the Effectiveness of Responsive Aggression Regulation Therapy (Re-ART). *International Journal of Forensic Mental Health*, 13(1), 25–35.
<https://doi.org/10.1080/14999013.2014.893711>
- Tollenaar, N., Van der Laan, A. M., & Beijersbergen, K. A. (2014). *Korte-en langetermijneffecten van de ISD-maatregel. Technisch rapport*. Den Haag: WODC. Cahier 2014-10.
https://repository.wodc.nl/bitstream/handle/20.500.12832/141/cahier-2014-10-volledige-tekst-nw_tcm28-72940.pdf?sequence=2&isAllowed=y
- Aarten, P. G., Denkers, A., Borgers, M. J. & Laan, P. H. van der (2015). Recidivism rates after suspended sentences: comparison of the effects of different types of suspended sentences on reconviction in the Netherlands. *International journal of offender therapy and comparative criminology*, 59(2), 143–158.
<https://doi.org/10.1177/0306624X13508929>
- Smid, W. J., Kamphuis, J. H., Wever, E. C., & Van Beek, D. J. (2016). A Quasi-Experimental Evaluation of High-Intensity Inpatient Sex Offender Treatment in the Netherlands. *Sexual Abuse*, 28(5), 469–485.
<https://doi.org/10.1177/1079063214535817>
- Stouwe, T. van der, Asscher, J.J., Hoeve, M., Laan, P. H. van der, Stams, G. J. M. (2016). Social skills training for juvenile delinquents: post-treatment changes. *Journal of Experimental Criminology*, 12(4), 515–536.
<https://doi.org/10.1007/s11292-016-9262-2>
- Wermink, H., Ramakers, A., Nieuwbeerta, P., Keijser, J. W. de & Dirkzwager, A. (2017). Onderzoeksnotitie: Recidive na een korte of langere periode in detentie. *Tijdschrift voor Criminologie*, 57(1-2), 30-51.
<https://doi.org/10.5553/TvC/0165182X2017059102003>
- Schaftenaar, P., Outheusden, I. van, Stams, G. J. & Baart, A. (2018). Relational caring and contact after treatment. An evaluation study on criminal recidivism. *International Journal of Law and Psychiatry*, 60, 45–50. <https://doi.org/10.1016/j.ijlp.2018.07.011>
- Wermink, H., Nieuwbeerta, P., Ramakers, A. A. T., Keijser, J. W. de, & Dirkzwager, A. J. E. (2018). Short-Term Effects of Imprisonment Length on Recidivism in the Netherlands. *Crime and delinquency*, 64(8), 1057–1093.
<https://doi.org/10.1177/0011128716687290>
- Tollenaar, N., Beerthuizen, M. G. C. J., Drieschner, K. H., & Van der Laan, A. M. (2019). *Effectiviteit van de ISD-maatregel-2e replicatie*. WODC Cahier 2019-19 Den Haag : WODC.
https://repository.wodc.nl/bitstream/handle/20.500.12832/237/Cahier_2019-19_Volledige_tekst_tcm28-425782.pdf?sequence=2&isAllowed=y
- Van Der Stouwe, T., Asscher, J. J., Hoeve, M., Van Der Laan, P. H., & Stams, G. J. J. (2019). Follow-up of a social skills training (SST) for juvenile delinquents: effects on

- reoffending. *Journal of Experimental Criminology*, 15, 243-252.
<https://doi.org/10.1007/s11292-018-9340-8>
- Bosma, A. Q., Kunst, M. J. J., Dirkzwager, A. J. E. & Nieuwbeerta, P. (2020). Recidivism After a Prison Based Treatment Program: a Comparison Between a Treatment and Control Group Using Proportional Weighting Within Strata. *Journal of Developmental and Life-Course Criminology*, 6(1), 95–126.
<https://doi.org/10.1007/s40865-020-00139-3>
- Blom, M., & Blokdijk, D. (2021) Long-term effectiveness of the alcohol ignition interlock programme: A retrospective cohort study in the Netherlands. *Accident Analysis & Prevention*, Volume 151, 105888. <https://doi.org/10.1016/j.aap.2020.105888>
- Duindam, H. M., Creemers, H. E., Hoeve, M. & Asscher, J. J. (2021). A Quasi-Experimental Study on the Effectiveness of Dutch Cell Dogs in Incarcerated Youth. *International Journal of Offender Therapy and Comparative Criminology*, 65(5), 644–666. <https://doi.org/10.1177/0306624X20983748>
- Duindam, H. M., Creemers, H. E., Hoeve, M. & Asscher, J. J. (2021). Who Lets the Dog In? Differential Effects of a Dog-Training Program for Incarcerated Adults. *Anthrozoös*, 34(6), 839–861.
<https://doi.org/10.1080/08927936.2021.1938405>
- Duindam, H. M., Creemers, H. E., Hoeve, M. & Asscher, J. J. (2021). Breaking the chains? The effects of training a shelter dog in prison on criminal behavior and recidivism. *Applied Developmental Science*, 26(4), 813–826.
<https://doi.org/10.1080/10888691.2021.2007768>
- Wermink, H., Blokland, A., Tollenaar, N., Been, J., Schuyt, P. & Apel, R. (2022). Een oude vraag opnieuw onderzocht: effecten van korte gevangenisstraffen ten opzichte van niet-vrijheidsbenemende straffen op recidive op basis van een instrumentele variabele benadering. *Tijdschrift voor Criminologie* 64(2), 147-169.
<https://doi.org/10.5553/TvC/0165182X2022064002001>
- Koops-Geuze, G. J., Wermink, H. T. & Weerman, F. M. (2023). A Quasi-Experimental Study on the Effects of Community versus Custodial Sanctions in Youth Justice. *Youth Violence and Juvenile Justice*, 21(2), 106-129. <https://doi.org/10.1177/15412040221133094>
- Prop, L.J.C., Laan, A.M. van der, Beerthuizen, M.G.C.J., Barendregt, C. S. & Nieuwenhuizen, C. (2023). Sentencing young adults with juvenile sanctions in The Netherlands: increasing risk or no differences on the chance of recidivism. *Journal of Experimental Criminology*. <https://doi.org/10.1007/s11292-023-09593-7>
- Stam, M. T. C., Wermink, H. T., Blokland, A. A. J. & Been, J. (2023). The effects of imprisonment length on recidivism: a judge stringency instrumental variable approach. *Journal of Experimental Criminology*, 20(3), 973–1004.
<https://doi.org/10.1007/s11292-023-09566-w>
- Vooren, M., Rud, I., Cornelisz, I., Klaveren, C. van, Groot, W. & Brink, H. Maassen van der (2023). The effects of a restorative justice programme (Halt) on educational

outcomes and recidivism of young people. *Journal of Experimental Criminology*, 19(3), 691–711. <https://doi.org/10.1007/s11292-022-09502-4>

Wermink, H., Been, J., Schuyt, P., Wijck, P. van & Blokland, A. (2023). The price of retribution: evidence from the willingness to pay for short-term prison sentences compared to community service orders. *Journal of Experimental Criminology*. <https://doi.org/10.1007/s11292-023-09583-9>

Wermink, H. T., Blokland, A. A. J., Been, J., Schuyt, P. M., Tollenaar, N. & Apel, R. (2023). Estimating effects of short-term imprisonment on crime using random judge assignments. *Justice Quarterly*, 41(3), 317–346. <https://doi.org/10.1080/07418825.2023.2193618>

Belgische studies 1984-2003

Geudens, H. (1998). The recidivism of community service as a restitutive judicial sanction in comparison with the traditional juvenile justice measures. *Restorative justice for juveniles. Potentialities, risks and problems for research*, 335-350.

Goossens, F., Kluppels, L. & Vanlaar, W. (2003). Leiden sensibilisatiecursussen voor bestuurders onder invloed van alcohol tot een lager recidivegehalte dan klassieke straffen? Een empirische evaluatie van de Belgische sensibilisatiecursussen op basis van de survival time tot de eerste recidive. *Panopticon*, 24(4), 365-390. <https://www.maklu-online.eu/nl/tijdschrift/panopticon/jaargang-volume-24/4-juli-augustus-july-august-2003/leiden-sensibilisatiecursussen-voor-bestuurders-on/>

Belgische studies 2004-2023

Blokland, A., Wermink, H., Robert, L. & Maes, E. (2015). Wederopsluiting na elektronische detentie en reguliere detentie in België. *Tijdschrift voor Criminologie*, 57(1), 31-58. <https://doi.org/10.5553/TvC/0165182X2015057001002>