

Een pioniersstudie in justitieland: de evaluatie van Halt in een experimentele setting

Annemarie ten Boom, Henk Ferwerda en Ilse van Leiden

Het vraagstuk van de effectiviteit van strafrechtelijke interventies, en dus ook het meten daarvan, staat de laatste jaren hoger op de justitieagenda dan voorheen. Zo beoogt het kabinet in de nota 'Naar een veiliger samenleving' uit oktober 2002 onder meer dat de justitiële reactie wordt ingericht op de doeltreffendheid van de afdoening, waarbij de reductie van recidive een belangrijke graadmeter is. Met betrekking tot de aanpak van jeugdcriminaliteit wordt in 2002 door het kabinet in de nota 'Vasthoudend en effectief' gemeend dat uitvoering van de vele programma's en projecten nog te veel een zaak van goede bedoelingen is. In het daarop volgende actieprogramma 'Jeugd Terecht' uit 2003 stelt het kabinet dat er in Nederland nog te weinig bekend is over de feitelijke uitvoering van de verschillende sanctiemodaliteiten en de uiteindelijke effecten daarvan. Aangekondigd wordt dat uiteindelijk alleen nog programma's die als effectief (in termen van voorkomen van recidive) of veelbelovend zijn beoordeeld, zullen worden gefinancierd. De sinds ruim een jaar opererende Erkenningcommissie Gedragsinterventies Justitie is hier een uitvloeisel van.

Tegelijkertijd bestaat er in het Nederlandse justitieonderzoek geen traditie van (gerandomiseerde) experimenten. Ook internationaal zijn relatief weinig van deze experimenten uitgevoerd (Farrington & Welsh, 2005); de redenen hiervoor zijn de vele ethische, juridische en praktische obstakels waar onderzoekers tegenaan lopen. In situaties waarin de vergelijking tussen een experimentele en een controlegroep noodzakelijk werd geacht, werd gezocht naar zo goed mogelijke alternatieven. Kruissink en Verwers (1989) werkten in hun evaluatie van Halt bijvoorbeeld met een controlegroep van jongeren uit een gemeente waar Halt op dat moment nog niet functioneerde: het betrof dus jongeren die voor soortgelijke feiten enkel een reactie van de politie kregen. Voor het daadwerkelijk willekeurig door de onderzoeker toewijzen van kandidaten aan een groep die de interventie wel en een groep die de interventie niet zou ondergaan, zijn de handen tot voor kort, voorzover wij konden nagaan, niet op elkaar gegaan. Het justitieveld staat hierin overigens niet alleen, ook bijvoorbeeld in het onderwijs worden experimenten soms belemmerd door ethische bezwaren of restricties in tijd en geld, en zoekt men naar 'natuurlijke experimenten': toevallige situaties in de werkelijkheid waardoor een aselechte controlegroep wordt gecreëerd (Webbink, 2005).

In dit artikel schetsen wij enkele drempels en problemen die we bij ons streven naar een gerandomiseerd experiment ten behoeve van een evaluatie van de Haltafdoening zijn tegengekomen. We beschrijven eerst het proces dat voorafging aan de beslissing bij de Halt-evaluatie te gaan werken met een experimentele opzet. Aan de orde komt welke alternatieve designs zijn overwogen en welke argumenten zijn gebruikt om behalve het justitiedepartement ook iedereen die daarna

beroepsmatig zou moeten meewerken, te overtuigen van de aanvaardbaarheid van het design. Vervolgens staan we stil bij het onderzoek als zodanig en beschrijven we – beknopt – de belangrijkste resultaten. We sluiten dit artikel af met een aantal discussie- en aandachtspunten met het oog op het verder verbeteren van de kwaliteit van (quasi-)experimenteel onderzoek.

Alternatieven voor experimentele opzet

Tegen de hiervoor beschreven achtergrond begint het Wetenschappelijk Onderzoek- en Documentatiecentrum (WODC) van het Ministerie van Justitie in 2002 met het voorbereiden van een evaluatie van de Halt-afdoening. In de onderzoeks-aanvraag die bij drie onderzoeksbureaus wordt uitgezet, vraagt het WODC hen de resultaten na te gaan, in termen van recidive en andere kenmerken, van de Halt-afdoening en de factoren die van invloed zijn op het effect van de maatregel. Daarbij wordt aangegeven dat naar het oordeel van het WODC een groep Halt-jongeren middels een voor- en nameting moet worden gevolgd *en* dat een vergelijkingsgroep nodig is. Aan de bureaus wordt expliciet gevraagd om in hun offerte aan te geven of en op welke wijze het volgens hen mogelijk is om een vergelijkingsgroep van voldoende omvang in het onderzoek te betrekken. Het WODC ontvangt de volgende vier voorstellen (afkomstig van drie bureaus¹):

- Stel de controlegroep samen uit jongeren die naar Halt zijn verwezen en een intakegesprek hebben gehad, maar die daarna nooit (meer) zijn komen opdagen op de werkplek of bij Halt. Deze groep is reeds door de politie en de Halt-medewerkers gescreend op 'Halt-waardigheid' en dus – volgens de indieners – qua gepleegde delicten en hoogte van het eventuele schadebedrag vergelijkbaar met de groep die wel deelneemt. Daarbij wordt opgemerkt dat deze groep 'weigerars' een ander sociaal profiel zou kunnen hebben dan de groep deelnemers; het feit dat ze weigeren, wijst wellicht op probleemgedrag. Maar, zo stelt men, de Halt-medewerker heeft van tevoren een gesprek gevoerd met die jeugdigen en daaruit is geen sociale indicatie naar voren gekomen. Bovendien beloven de indieners dat zal worden gecontroleerd of het sociale profiel overeenkomt.
- Vraag aan jongeren die zijn doorverwezen naar Halt om een of twee vriendjes te noemen die evenals zij strafbaar gedrag vertonen, maar niet in aanraking zijn geweest met de politie. Via de Halt-jongeren wordt met andere woorden een gericht steekproef getrokken uit het 'dark figure'.
- Stel de controlegroep samen uit jongeren die een politiewaarschuwing hebben gekregen voor gedrag dat vergelijkbaar is met een Halt-waardig delict. Deze keus berust op het vermoeden dat de grens tussen een waarschuwing en een Halt-waardig delict vaak niet zo scherp te trekken is. Via de administratieve gegevens in het bedrijfsprocessensysteem (BPS) van de politie zal een controlegroep moeten worden samengesteld die zoveel mogelijk overeenkomt met de groep jongeren die wel naar Halt is verwezen.

1 Met dank voor hun instemming met het geanonimiseerd weergeven van delen uit hun onderzoeksoffertes.

- Kies voor een onderzoeksmodel waarbij sprake is van een 'zuivere' controlegroep. Voorgesteld wordt om uit de groep jongeren, die door de politie naar Halt wordt doorverwezen, aselect een onderzoeksgroep en een controlegroep samen te stellen. De ene doorgestuurde jongere krijgt een Halt-afdoening en de andere niet. Beiden zullen deelnemen aan het onderzoek. Gesteld wordt dat dit eigenlijk de enige manier is om een goed effectonderzoek uit te voeren. Alternatieve oplossingen worden dan ook niet gesuggereerd in deze offerte. Wel wordt ingegaan op mogelijke tegenwerpingen en obstakels en deze worden van tegenargumentatie voorzien. We komen hier later op terug.

Doorslaggevend bij de keuze voor het onderzoeksdesign, dat het beste antwoord kan geven op de onderzoeksvragen, is vanzelfsprekend de geschiktheid van de voorgestelde controlegroep. De offertebeoordelaars, twee WODC-ers en een beleidsambtenaar van het Ministerie van justitie, achten de controlegroep in het eerste voorstel onvoldoende vergelijkbaar. Dit niet alleen vanwege het in hun ogen zeer waarschijnlijk andere sociale profiel, maar ook omdat hen een recent onderzoek bekend is, waaruit blijkt dat onder jongeren die hun deelname aan een jeugdpreventieproject voortijdig afbreken, een veel hogere recidive voorkomt dan onder jongeren die hun traject voltooien (Woldringh, 2001).

Tegen het tweede voorstel wordt ingebracht dat opgepakt worden door de politie een autonome invloed heeft op de jongeren. Dit maakt een jongere die wel met de politie in aanraking is gekomen onvergelijkbaar met een jongere die geen politiecontact heeft gehad. Als de controlegroep niet alleen verschilt van de experimentele groep wat betreft de Halt-afdoening, maar ook op het punt of men al dan niet met de politie in aanraking is gekomen, is het vaststellen van de effectiviteit van de Halt-afdoening als zodanig niet vast te stellen. Terwijl dat wel degelijk het doel is.

Wat betreft het derde voorstel twifelen de beoordelaars sterk of er voldoende van deze jongeren in het BPS te vinden zullen zijn, omdat deze werkwijze officieel niet meer is toegestaan. Men acht het niet uitgesloten dat tegen de instructies in politiewaarschuwingen (het oude 'politiesepot') worden uitgedeeld, maar het lijkt onwaarschijnlijk dat deze dan vervolgens allemaal netjes in het BPS worden geregistreerd. Los daarvan vraagt men zich af of deze jongeren toch niet gemiddeld een minder ernstig delict hebben gepleegd dan de jongeren die wel naar Halt zijn verwezen. Men is dus ook van de vergelijkbaarheid van deze groep niet overtuigd. De beoordelaars geven, gezien de geschetste alternatieven en de hen bekende methodologische voordelen (zie bijvoorbeeld Farrington & Welsh, 2005; Bijleveld, 2006), de voorkeur aan het vierde voorstel. Bij elke alternatieve samenstelling van de controlegroep zullen te veel twijfels blijven bestaan over de interne validiteit. Daarbij zijn zij zich bewust van het ontbreken van enige justitietraditie op dit vlak, maar ook van de hiervoor beschreven beleidspolitieke tijdgeest, waarin het inzichtelijk maken van de effectiviteit van interventies, althans op papier, belangrijk wordt gevonden. Tegen die achtergrond besluiten de beoordelaars een serieuze poging te ondernemen om het gewenste design mogelijk te maken, door hiervoor toestemming te vragen aan de ambtelijke leiding van het ministerie.

Argumenten voor en tegen het experiment

Op een aantal tegenwerpingen werd – in navolging van de indieners van het onderzoeksvorstel – in het advies aan de ambtelijke leiding geanticipeerd, door zowel algemene als Halt-specifieke tegenargumenten naar voren te brengen:

- Het juridische bezwaar van rechtsongelijkheid: voor hetzelfde delict krijgt de een de Halt-afdoening en de ander niet.

Hiertegen wordt aangevoerd dat de jongeren in de controlegroep weliswaar geen Halt-afdoening krijgen, maar wel moeten meewerken aan het onderzoek. Indien ze niet willen meewerken aan het onderzoek, krijgen ze alsnog de Halt-afdoening. Verder is bekend dat niet-succesvolle Halt-afdoeningen relatief vaak door het OM alsnog niet worden vervolgd, zodat genoemde rechtsongelijkheid zich in de praktijk wel vaker voordoet (zie ook Korf, 2003)

- Het ethische bezwaar dat een groep jongeren een interventie wordt onthouden, waarbij zij baat kunnen hebben.

Tegen dit bezwaar wordt ingebracht dat de groep die een interventie wordt onthouden relatief klein is en bovendien vele malen kleiner dan de groep die op termijn door dit onderzoek mogelijk een effectievere afdoening kan krijgen. Een ander tegenargument is dat de Halt-afdoening een relatief lichte pedagogische sanctie is, die maximaal twintig uur behelst. Onthouding zal naar verwachting geen ingrijpende gevolgen hebben voor de persoon zelf en voor de samenleving. Feder e.a. (2000) gebruiken een ander argument tegen het aangevoerde ethische bezwaar: het is juist onethisch om een interventie (counseling in dat geval) op te leggen, zonder eerst de positieve en negatieve gevolgen daarvan uitvoerig te testen. Dit argument is door ons niet gebruikt; de Halt-afdoening is immers in het verleden ook geëvalueerd, zodat dit punt geen hout zou snijden. Het argument is mogelijk wel bruikbaar in gevallen, waarin een nieuwe interventie moet worden geëvalueerd.

- Een praktisch-financieel bezwaar is de verwachting dat jongeren die in de controlegroep worden geplaatst en dus geen Halt-afdoening zullen ondergaan, voor de Halt-bureaus inkomstenderving tot gevolg hebben. De Halt-bureaus worden immers op output gefinancierd. Om constructieve medewerking van de bureaus te verkrijgen, kan het aanbieden van een financiële tegemoetkoming belangrijk zijn.

Daarom wordt ervoor gepleit om de Halt-bureaus ook voor de controlegroep-jongeren een vergoeding te geven. Gesteld is dat dit ook niet onredelijk is, omdat de bureaus tijd moeten steken in het onderzoek, zowel bij de experimentele als de controlegroep. Ook is erop gewezen dat de kosten van vergoeding slechts een fractie zijn van wat er jaarlijks in de Halt-sector omgaat (zo'n 9 miljoen euro in 2002; inmiddels zo'n 11 miljoen).

In het advies aan de ambtelijke leiding om de uitvoerbaar geachte experimentele opzet mogelijk te maken, wordt vanzelfsprekend ook verwezen naar de betreffende actuele beleidsnota's en naar de belangrijke plaats die de Halt-afdoening in het reactiearsenaal met betrekking tot risicojongeren inneemt: het gaat jaarlijks om ruim 20.000 jongeren. Op hoog niveau wordt uiteindelijk besloten, dat de

voordelen opwegen tegen de nadelen. Doorslaggevend wordt gevonden, dat een betekenisvolle stap voorwaarts wordt gemaakt in de kennis over de effectiviteit van de Haltafdoening. Er zijn op dat moment weliswaar indicaties dat met Halt de kans op recidive kleiner is dan bij een politiesepot, de empirisch-wetenschappelijke onderbouwing voor de stelling dat de recidive afneemt *door* Halt is nog altijd tamelijk zwak. Dat concludeert ook Korf (2003) in een artikel, waarin hij bestaande studies over recidive na Halt analyseert en dat ongeveer tegelijkertijd verschijnt.

De uitvoering van het onderzoek

Bij de uitvoering van het onderzoek² is, uitgaande van de richtlijnen van de Campbell Collaboration Crime and Justice Group, gewerkt met een onderzoeksopzet die minimaal vereist is om verantwoorde conclusies te kunnen trekken over de effectiviteit van – in dit geval – de Halt-afdoening. Zo zijn er voor en na de interventie metingen uitgevoerd en is gewerkt met een controlegroep die de afdoening niet ondergaat en een experimentele groep die de afdoening wel ondergaat (Farrington, 2003).

Met het oog op de praktische en financiële haalbaarheid is besloten het onderzoek te starten op het moment dat het eerste gesprek met de jongere bij Halt zal gaan plaatsvinden. Op de geselecteerde twaalf Halt-bureaus voert een beperkte groep Halt-medewerkers het onderzoek uit (hier komen we later op terug). Hiertoe hadden de vijf leden van het onderzoeksteam elk hun eigen bureaus ‘geadopteerd’. Voor aanvang van het onderzoek zijn de Halt-medewerkers geïnstrueerd en daarnaast is met betrekking tot iedere onderzoeksfase een onderzoeksprotocol voor hen opgesteld. Ook tijdens het onderzoek was er regelmatig contact tussen onderzoekers en Halt-bureaus, minstens over de voortgang, evaluatie, instructie voor de eerste meting en afronding van het onderzoek.

Gedurende een jaar zijn ruim 800 jongeren (en hun ouders) gevolgd, die vanwege het plegen van een strafbaar feit door de politie naar bureau Halt zijn verwezen. Nadat de jongere op uitnodiging van het Halt-bureau, al dan niet met zijn/haar ouder(s), op het Halt-bureau verscheen, is hen door de Halt-medewerker gevraagd deel te nemen aan een onderzoek. Verteld werd dat deelname inhield dat de jongere en zijn of haar ouder op twee momenten een vragenlijst zouden moeten invullen. Voorafgaand aan deze eerste meting heeft de Halt-medewerker van de onderzoekers gehoord, in welke onderzoeksgroep de betreffende jongere moet worden geplaatst, wanneer hij instemt met deelname. Over de indeling was geen discussie mogelijk; deze werd op basis van objectieve criteria door de onderzoeker uitgevoerd. (Ook hier komen we later op terug.)

Indien de jongere niet wenste deel te nemen, noteerde de Halt-medewerker enkele gegevens van hem of haar, vervolgde de gebruikelijke Halt-procedure en lichtte nadien de onderzoeker in. Bij instemming van de jongere werd de vragenlijst

2 Voor de complete verantwoording en alle onderzoeksresultaten verwijzen we naar het volledige onderzoeksrapport: Ferwerda, H., van Leiden, I., Arts, N. & Hauber, A. (2006). *Halt: het alternatief? De effecten van Halt beschreven*. WODC, Onderzoek en beleid, nr. 244, Den Haag: Boom Juridische uitgevers.

voorgelegd en na deze te hebben ingevuld, werd in het geval dat de jongere in de experimentele groep was geplaatst de Halt-procedure vervolgd. Indien de jongere aan de controlegroep was toegewezen, werd na het invullen van de vragenlijst verteld dat er geen Halt-afdoening zou volgen, mits hij of zij bereid was om over een half jaar nogmaals een vragenlijst in te vullen.

Na een half jaar hebben de jongeren en hun ouders uit beide onderzoeksgroepen de tweede vragenlijst ingevuld. Nog een half jaar later, dus een jaar na de eerste meting, is voor alle jongeren een recidiveonderzoek uitgevoerd.³ De uiteindelijke onderzoeksgroep bestond uit 824 jongeren, 450 in de controlegroep en 374 in de experimentele groep. De respons was zeer goed: 87 procent van de jongeren en 77 procent van de ouders deed aan beide metingen mee.

Halt vooral effectief voor oorspronkelijke doelgroep

Het onderzoek wijst uit dat de effectiviteit van Halt beperkt is. In het algemeen geldt dat jongeren die wel een Halt-afdoening hebben doorlopen, daarna niet meer of minder (ernstige) delicten plegen dan jongeren die zijn vrijgesteld van Halt. Recidive is gemeten aan de hand van zelfrapportage én door de politie geregistreerde criminaliteit. Hierbij is recidive geoperationaliseerd als het wederom als verdachte voorkomen in het bedrijfsprocessensysteem van de regiopolitie. Op basis van door de jongeren zelf gerapporteerde criminaliteit recidiveert in zowel de controlegroep als de experimentele groep 76,7 procent van de jongeren binnen een half jaar na het eerste gesprek bij Halt. De door de politie geregistreerde recidive bedraagt, een jaar na het eerste contact bij Halt, voor de controlegroep 29,1 procent en voor de experimentele groep 29,9 procent.

De jongeren in beide groepen hebben het opgepakt worden door de politie en de eerste gang naar Halt meegemaakt. De indruk die dit op de jongeren maakt, kan ertoe hebben bijgedragen, dat er verder geen verschillen in recidive zijn gevonden, maar dat is in het onderzoek niet onderzocht. Omdat het leeuwendeel van de Halt-waardige feiten als relatief licht te bestempelen is, heeft het opgepakt worden door de politie en mogelijke strafdreiging (de jongere is immers doorverwezen naar Halt) misschien een sterker effect dan de Halt-afdoening zelf.

Nadere analyse van de onderzoeksgroep leert dat de Halt-afdoening wel effectief is voor jongeren met bepaalde kenmerken. Zo zijn jongeren die als motivatie voor het plegen van het delict geven 'omdat iedereen het deed', de 'meelopers', wel gebaat bij een Halt-afdoening. Daarnaast vloeit er uit het onderzoek een profiel van de ideale Halt-jongeren voort, met een zestal kenmerken: first offenders, wonen buiten de Randstad, zich bewust zijn van de gevolgen van hun gedrag, geen of lichte problematiek hebben, sociaal aangepast gedrag vertonen en een positieve vrijetijdsbesteding hebben. Naarmate jongeren meer aan dit profiel voldoen, blijkt Halt een adequatere maatregel, al is het effect maar klein. Al met al lijken de resultaten erop te wijzen, dat de Halt-afdoening in haar huidige vorm vooral effectief is voor de oorspronkelijke doelgroep van Halt, jongeren die voor

3 Behalve recidive vormden ook 'moeilijkheden' en 'sociaal gedrag' – gemeten met de *Strengths and Difficulties Questionnaire* (SDQ) effectmaten in dit onderzoek. We verwijzen naar het onderzoeksrapport voor meer informatie daarover.

het eerst én voor een licht vergrijp met de politie in aanraking komen. Bij het bepalen van de effectiviteit van de Halt-afdoening is niet alleen gekeken naar de kenmerken van de jongeren, maar ook of er een samenhang is tussen de afzonderlijke elementen van de Halt-afdoening en recidive.

Het blijkt dat excuses aanbieden aan het slachtoffer een positief, dus recidiveverminderend effect heeft op toekomstig crimineel gedrag. Overige elementen van de Halt-afdoening, zoals het aantal uren dat de jongere taken moet uitvoeren, het aantal gevoerde gesprekken, het al dan niet treffen van een schaderegeling en de wijze van begeleiding van de jongere, zijn niet van invloed op recidive.

De onderzoeksresultaten indiceren dat de Halt-afdoening niet in dezelfde vorm voor elke jongere even effectief is wat betreft positieve gedragsbeïnvloeding. Screening van de jongeren door professionals, nadat ze zijn opgepakt door de politie, om beter te kunnen beoordelen of de jongere daadwerkelijk geschikt is voor doorverwijzing naar Halt (maatwerk), lijkt essentieel om met Halt te bereiken wat ermee beoogd wordt. Jongeren die negatief gedrag vertonen, bij wie sprake is van een bredere problematiek of van wie de recidivekans hoog wordt ingeschat, zouden met andere woorden intensiever begeleid dienen te worden dan binnen de huidige Halt-afdoening plaatsvindt. Bij jongeren zonder dergelijke kenmerken zou bijvoorbeeld kunnen worden volstaan met een gesprek van een professional met de jongere en zijn of haar ouder(s) én met het aanbieden van excuses aan het slachtoffer.

Eerdere studies naar de effectiviteit van Halt indiceerden een sterke samenhang tussen recidive en de persoonskenmerken (geslacht, opleidingsniveau en etniciteit) van de jongeren en het type delict waarvoor zij bij Halt terecht komen (Korf, 2003). Deze samenhangen zijn in het onderhavige onderzoek noch na bivariate, noch na multivariate analyses aangetroffen. Ook het eerder door Dijkman en Gunther Moor (1998) aangetroffen verband tussen intensiteit van de begeleiding en recidive is niet bevestigd.

Onze bevindingen kunnen worden gezien als een understreping van de voorzichtigheid waarmee conclusies op basis van minder krachtige onderzoeksdesigns getrokken moeten worden.

Discussie- en aandachtspunten

Op basis van het onderzoek zijn wij bij het uitvoeren van dit (quasi-)experimentele onderzoek in onze rollen als projectbegeleider vanuit het WODC en onderzoekers tegen een aantal zaken aangelopen. Deze discussie- en aandachtspunten willen we graag met collega-onderzoekers delen om de kwaliteit van dergelijk onderzoek in de toekomst verder te verbeteren.

Belemmeringen voor randomisatie in de praktijk

Idealiter zou uit de groep jongeren, die door de politie naar Halt wordt verwezen, aselekt een onderzoeksgroep en een controlegroep moeten worden samengesteld. Immers, in tegenstelling tot matching of statistische controle van variabelen, controleert randomisatie zowel voor onbekende en ongemeten variabelen als voor bekende en gemeten variabelen. In dit onderzoek heeft uiteindelijk geen

volledig aselechte toewijzing kunnen plaatsvinden, om de volgende redenen. De eerste en belangrijkste reden was de angst c.q. het risico dat met een aselechte steekproef, over twaalf Halt-bureaus verdeeld, twee onderzoeksgroepen zouden worden samengesteld, die op een aantal theoretisch interessante variabelen significant van elkaar zouden afwijken. Vanzelfsprekend zou dit probleem met een grotere steekproef op te lossen zijn geweest, maar dat was praktisch en financieel niet haalbaar. Bij de samenstelling van de twee onderzoeksgroepen is gestreefd naar een gelijke verdeling van onafhankelijke variabelen over de experimentele en de controlegroep: leeftijd, geslacht, etniciteit, strafbaar feit alleen of in groepsverband gepleegd, en type overtreding of misdrijf.

De variabele 'strafbaar feit alleen of in groepsverband gepleegd' was ook als zodanig een complicerende factor. Er worden regelmatig jongeren naar Halt verwezen, die het delict in groepsverband hebben gepleegd. Het werd niet wenselijk geacht uit dezelfde groep sommige jongeren in de experimentele en andere in de controlegroep te plaatsen. Verwacht werd dat jongeren en/of ouders dat niet zouden accepteren en wegens het belang om niet te veel ruchtbaarheid aan het design te geven (zie ook later in dit artikel), is daarvoor niet gekozen. Groepen zijn dus in hun geheel in de ene of in de andere onderzoeksgroep geplaatst en daarbij is rekening gehouden met de plaatsing van eerdere groepen; we wilden immers groepsplegers in zowel de experimentele als de controlegroep. In toekomstig onderzoek lijkt het aan te raden om de eenlingen en de groepen apart te randomiseren.

Tot slot werd niet alleen voor de totale onderzoeksgroep, maar ook voor de onderzoeksgroep per deelnemend Halt-bureau (twaalf in totaal) gestreefd naar een vergelijkbare experimentele en controlegroep. Dit had in de eerste plaats te maken met het feit dat we in beide groepen een gelijke verdeling over plattelandsjongeren en stadsjongeren wilden hebben. Ook waren we er vooraf niet zeker van, dat ieder Halt-bureau dezelfde werkwijze hanteert of mogelijk andere accenten in haar aanpak heeft.

Om te streven naar een steekproef van in totaal duizend jongeren is er per Halt-bureau uitgegaan van 14 à 15 procent van de jaarpopulatie, met een evenredige verdeling over de twee onderzoeksgroepen. De keuze voor de twaalf Halt-bureaus is gebaseerd op geografische spreiding en variatie in jaarpopulatie. Bij de kleinere Halt-bureaus betrof de steekproef een zo klein aantal jongeren (bijvoorbeeld Breda: 20, Maastricht: 26, Enschede: 34)⁴ dat bij randomisatie bovendien het risico aanwezig was dat het toeval (aselechte steekproef) tot te grote verschillen in de experimentele en controlegroep zou leiden. Farrington (in Farrington & Welsh, 2005) stelde in zijn studie van Engelstalige publicaties over gerandomiseerde experimenten tussen 1957 en 1981 als ondergrens, dat tenminste vijftig personen aan iedere conditie moesten zijn toegewezen. Bij kleinere aantallen zou het belangrijkste voordeel van randomisatie, zorgen voor gelijkwaardigheid van personen op alle gemeten en niet-gemeten externe variabelen, waarschijnlijk niet worden gerealiseerd.

4 De helft hiervan telkens in de experimentele groep.

Draagvlak van de Halt-bureaus en de medewerkers

Ten behoeve van een nauwkeurige samenstelling en controle van de onderzoeksgroepen geschiedde de toewijzing aan de experimentele dan wel controlegroep door de onderzoekers. Het belang daarvan is in theorie helder en werd ook in de praktijk regelmatig duidelijk. Meerdere keren hebben Halt-medewerkers het verzoek ingediend om een bepaalde jongere in de experimentele groep te plaatsen 'omdat deze gebaat zou zijn bij een Halt-afdoening'. Het onthouden van de Halt-afdoening aan jongeren druiste in tegen het 'hulpverlenershart' van de Halt-medewerkers. Juist een dergelijke, gekleurde toewijzing aan de onderzoeksgroepen is voorkomen door altijd alleen de onderzoekers te laten beslissen. Dit dilemma heeft er overigens wel toe geleid dat de onderzoekers bij de start, maar soms ook nog lopende het onderzoek veel tijd hebben moeten investeren in uitleggen en 'acceptabel' maken van het onderzoeksdesign.

Moment van scheiden van de onderzoeksgroepen

Een belangrijke vraag is ook die naar het beste moment van plaatsing in de onderzoeksgroepen. Om de effectiviteit van de Halt-afdoening optimaal te meten, is het wenselijk dat de controlegroep geheel niet in aanraking komt met de Halt-organisatie. Het zuiverst zou gemeten worden als de jongeren uit de controlegroep reeds op het politiebureau zouden worden heengezonden. Dit zou betekenen dat bij een landelijk onderzoek als dit, gezien het belang van toewijzing door de onderzoekers, op vele politiebureaus in Nederland (alleen Amsterdam telt al 30 wijkteams) onderzoeksmedewerkers zouden moeten worden gestationeerd om ter plaatse opgepakte jongeren toe te wijzen aan een van de onderzoeksgroepen en vragenlijsten af te nemen. Deze opzet is door zowel onderzoekers als begeleidingscommissie als organisatorisch en financieel niet haalbaar beoordeeld. Er is gekozen voor het haalbare *second best* design: de jongeren worden bij aankomst op het Halt-bureau toegewezen aan een van de groepen. We verliezen hierbij wel het inzicht in een deel van de werking van de Halt-afdoening:

- Nadat de jongeren op het politiebureau hebben 'gekozen' voor de Halt-afdoening, leven zij een aantal dagen tot weken in de veronderstelling dat zij een Halt-afdoening zullen ondergaan.
- Zij krijgen een brief thuis waarin zij worden uitgenodigd om met hun ouders naar het Halt-bureau te komen.
- Zij gaan (vaak samen met hun ouders) daadwerkelijk naar het Halt-bureau toe.

Kortom, zij maken mee dat de Halt-afdoening hen 'boven het hoofd hangt'. Dit is in feite een concrete strafdreiging, die effect kan hebben. Ook hun ouders maken mee dat hun zoon of dochter een afdoening boven het hoofd hangt, en reageren daar mogelijk op. We missen dus inzicht in de werking van de strafdreiging en van de actie van ouders. Wat we overhouden en dus meten, is het effect van de Halt-afdoening als zodanig (een of meer gesprekken, het uitvoeren van werk- of leerelementen, het aanbieden van excuses en het vergoeden van de schade aan het slachtoffer).

Het aankondigen van de tweede meting

Een punt van discussie is ook het aankondigen van de tweede meting (na zes maanden) onder de controlegroep tijdens de eerste meting. Daarbij werd verteld dat niet-invullen van de tweede lijst zou betekenen dat de Halt-afdoening alsnog zou plaatsvinden. Welk effect gaat hier mogelijk van uit? Daar staat tegenover, dat de respons vrijwel zeker minder hoog zou zijn geweest als de tweede meting niet vooraf al was aangekondigd. Om zowel een hoge respons te realiseren als een eventueel effect te achterhalen van de wetenschap dat men terug zou moeten komen, zou in de toekomst overwogen kunnen worden de controlegroep te splitsen in een groep die wel, en een groep die niet op de hoogte wordt gebracht van de tweede meting.

Discrepancie tussen (intern) communiceren en (extern) stilhouden

Omdat bij een goede uitvoering van het onderzoek veel partijen betrokken zijn, is de communicatie een belangrijk aspect geweest. De kern daarvan betrof het feit dat de helft van de jongeren uit het onderzoek zou worden vrijgesteld van de Haltafdoening. Terwijl de informatie hierover naar buiten toe zo minimaal mogelijk gehouden moest worden, moest de interne communicatie helder zijn, om geen enkele betrokken partij te passeren.

Wat betreft de externe partijen werden alleen de jongeren die in de controlegroep werden geplaatst en hun ouder(s) geïnformeerd over de vrijstelling in het kader van het onderzoek. De jongeren die in de experimentele groep werden geplaatst en hun ouder(s) zijn niet geïnformeerd over het bestaan van een tweede onderzoeksgroep.

Indien voortijdig publiekelijk bekend zou worden, dat jongeren die voor het plegen van een strafbaar feit naar Halt worden verwezen er zonder Halt-afdoening 'vanaf komen', zou dit het onderzoek negatief kunnen beïnvloeden én het imago van Halt kunnen beschadigen. Ondanks de lange looptijd van het onderzoek is geen berichtgeving over het onderzoek in de media gekomen.

Waar een minimale openbaarheid van het onderzoeksdesign naar externen van belang was, was juist een brede communicatie naar de interne partijen essentieel. Tot de interne partijen rekenen we de Halt-bureaus, de betrokken politieregio's en het Openbaar Ministerie. Nadat al deze partijen door een brief van het ministerie van Justitie waren geïnformeerd over achtergrond en inhoud van het onderzoek, zijn de onderzoekers in nauw contact getreden met de Halt-bureaus. Deze hadden op hun beurt de contacten met de politie. De politie is in het communicatietraject een belangrijke partij, omdat de politie de jongeren naar Bureau Halt verwijst en daarmee ook verwacht, dat de jongeren een Halt-afdoening doorlopen. Ouders zouden de politie kunnen vragen, waarom hun zoon of dochter geen Halt-afdoening had gekregen en een ander wel. De Halt-medewerkers gaven aan te vrezen dat het onderzoek, met name door het onthouden van de Halt-afdoening aan de controlegroep, de goed lopende contacten met politie en Openbaar Ministerie zou verstoren, indien er onvoldoende gecommuniceerd zou worden over het onderzoeksdesign. Hier zijn – voorzover wij hebben kunnen overzien – geen problemen geweest.

Op het juiste moment de juiste personen informeren is met andere woorden van het grootste belang. De eerlijkheid gebiedt te zeggen dat het bij een dergelijk landelijk onderzoek lastig is om er zeker van te zijn, dat werkelijk iedereen in het Halt-netwerk, die het zou moeten weten ook geheel en al op de hoogte is, denk aan jeugdofficieren van justitie, parketleiding, jeugdrechercheurs, korpsleiding.

Datamanagement: de veelheid aan onderzoeksdata

Gedurende het onderzoek is bij verschillende bronnen een grote hoeveelheid informatie verzameld. Zo is er op twee momenten, bij de eerste en de tweede meting, een vragenlijst voor de Halt-medewerker gebruikt, voor de jongeren en voor de ouders uit beide onderzoeksgroepen en met recidivegegevens van de politie. Dit leverde vanzelfsprekend een enorm databestand op. Twee zaken zijn hierbij van belang en leveren wellicht voor anderen bruikbare ervaringen op. Allereerst vraagt een dergelijk groot onderzoek om personen binnen een onderzoeksteam, die vaste taken uitvoeren. Zowel de algehele organisatie als de analyse van het enorme databestand is in beide gevallen door slechts één persoon uitgevoerd. Vanaf het moment van het verzamelen van de data tot en met de analyse moet sprake zijn van eenduidige coderingen voor de verschillende typen vragenlijsten en dient ook de analist geïnformeerd te zijn over de opzet en uitvoering van het onderzoek.

Wat betreft de data-analyse is ons gebleken dat door gevalideerde schalen dan wel beproefde vragensets binnen onze meetinstrumenten de veelheid aan data eenvoudiger te verwerken en te interpreteren was. Wanneer sprake was van grote aantallen losse vragen binnen thema's, hebben we via factoranalyse in veel gevallen ook gewerkt met schaalscores. Ondanks alle clustering om te komen tot reductie van de beschikbare data, is uiteindelijk van alle uitgevoerde analyses maximaal 50 procent in het uiteindelijke rapport terecht gekomen; analyseren is weggooien.

Tot slot

Feder e.a. (2000) beschrijven uitgebreid welke weerstanden zij hebben ontmoet bij een gerandomiseerd experiment naar de effecten van counseling bij huiselijk geweld in Florida. Deze counseling werd door de rechter bovenop de gebruikelijke afdoening opgelegd aan mannen, die schuldig waren bevonden aan het plegen van huiselijk geweld. De tegenwerking die men ondervond, liep zelfs uit op het aanspannen van een geding door de plaatselijke aanklager. In vergelijking hiermee is ons onderzoek zeer voorspoedig verlopen. Slechts één Halt-bureau heeft niet willen meewerken, om redenen van rechtsongelijkheid en omdat men op dit bureau vreesde, dat al snel door de stad rond zou gaan dat jongeren bij Halt vrijgesteld worden van sancties.

Het onderzoek '*Halt: het alternatief?*' geeft inzicht in de effecten van een relatief jonge strafrechtelijke interventie: de Halt-afdoening. Van veel andere, veelal oudere strafrechtelijke interventies kennen we dergelijke effecten in het geheel niet of veel minder goed. Met de resultaten van dit onderzoek staat Halt in de etalage. Het zou interessant zijn om de resultaten van de Halt-afdoening te vergelij-

ken met de resultaten c.q. effecten van andere strafrechtelijke interventies om te kunnen beoordelen of de effecten van Halt goed of slecht zijn. In ons onderzoek is een jaar na het eerste gesprek bij Halt gekeken naar recidive. Echter, we weten uit de literatuur (Van der Heiden-Attema & Wartna, 2000a; 2000b; Van Espen, 2002) dat pas na drie jaar de effecten van een interventie op langere termijn zichtbaar worden. We pleiten er dan ook voor om over twee jaar een vervolgstudie naar de recidive van de onderzochte groep uit te voeren.

We hopen tot slot dat op het justitieterrein vaker dan voorheen experimenteel onderzoek zal kunnen worden uitgevoerd, en dat onderzoekers dan hun voordeel kunnen doen met onze ervaringen.

Literatuur

- Bijleveld, C.C.J.H. (2006). Methoden en technieken van onderzoek in de criminologie. Den Haag: Boom Juridische uitgevers.
- Dijkman T. & Gunther Moor, L. (1998). Effectiviteit en kwaliteit van de Halt-afdoening. Nijmegen: ITS.
- Espen, M. van (2002) *Recidivepatronen Haltafdoening 1997*. Amsterdam: Bureau Halt Amsterdam-Amstelland.
- Farrington, D. (2003). Methodological Quality Standards for Evaluation Research. *Annals of the American Academy of Political and Social Science*, 587: 49-68.
- Farrington, D. & Welsh, B. (2005). Randomized experiments in criminology: What have we learned in the last two decades? *Journal of Experimental Criminology*, 1, 9-38.
- Ferwerda, H., Leiden, I. van, Arts, N. & Hauber, A. (2006). *Halt: het alternatief? De effecten van Halt beschreven*, Den Haag/Arnhem: WODC /Boom Juridische uitgevers/Advies- en Onderzoeksgroep Beke, Onderzoek en beleid, nr. 244.
- Feder, L., Jolin, A. & Feyerherm, W. (2000). Lessons from two randomized experiments in criminal justice settings. *Crime & Delinquency*, 46(3), 380-400.
- Heiden-Attema, N. van der & B. Wartna (2000a) *Recidive na verblijf in een JBI. Een vervolgstudie naar de geregistreerde criminaliteit onder jongeren uit een justitiële behandelinstelling*, Den Haag: WODC, Onderzoeksnotitie 2000/9.
- Heiden-Attema, N. van der & Wartna, B. (2000b). *Recidive na een strafrechtelijke maatregel. Een studie naar de geregistreerde criminaliteit onder jongeren met een pibb, een jeugd-tbr of een pij*, Den Haag: WODC, Onderzoeksnotitie 2000/9.
- Korf, D. (2003). Hoe succesvol is Halt? Een beschouwing over recidive. *Tijdschrift voor Criminologie*, 45(1), 17-34.
- Kruissink, M. & Verwers, C. (1989). *Halt: een alternatieve aanpak van vandalisme. Eindrapport van een evaluatie-onderzoek naar Haltprojecten*. Den Haag: WODC.
- Ministerie van Justitie (2002). *Vasthoudend en effectief. Versterking van de aanpak van jeugdcriminaliteit*. Kamerstuk 2001-2002, 28292, nr. 2, Tweede Kamer.
- Ministerie van Justitie en Ministerie van Binnenlandse Zaken en Koninkrijksrelaties (2002). *Naar een veiliger samenleving*. Den Haag: Ministerie van Justitie en Ministerie van Binnenlandse Zaken en Koninkrijksrelaties.
- Ministerie van Justitie (2003). *Jeugd Terecht. Actieprogramma aanpak jeugdcriminaliteit 2003 – 2006*. Den Haag: Ministerie van Justitie, directie Jeugd en Criminaliteitspreventie.
- Webbink, H. (2005). Causale effecten van beleid; over gecontroleerde en natuurlijke experimenten. *Justitiële verkenningen*, 31(8), 72-84.
- Woldringh, C. (2001). *Boefjes in het land van Cuijk. Korte- en lange-termijneffecten van een preventieproject voor jongeren van 12 tot 25 jaar*. Nijmegen: ITS.