

Naar zinvolle evaluatiemethoden voor community-projecten

Van een focus op interventies naar een focus op jeugdnetswerken

Chris Baerveldt

PED 37 (2): 173–185

DOI: 10.5117/PED2017.2.BAER

Abstract

The paper discusses the evaluation of two community-based intervention projects that were evaluated using a methodology that appears to differ strongly from traditional quasi-experimental designs. The main reason for this deviation is that traditional designs do not match the way community projects develop over time. Moreover, these designs do not provide enough information to improve these intervention projects in the future. It is argued that traditional designs can be partially tailored, but not sufficiently to address these needs. The evaluation methods that were actually used, also have drawbacks. It is argued that both the traditional and the here described evaluation methods do not provide enough information about what communities offer to children. Moreover, they do not account sufficiently for the democratic dynamics of community-based projects. A different direction is proposed, suggesting that not projects and interventions should be evaluated, but the children's experience of (change of the) quality of the entire network of people and organizations in their community.

Keywords: evaluation, methods, community, networks, youth

Inleiding

Een van de belangrijkste uitgangspunten van Micha de Winter is geweest dat kinderen onderdeel uitmaken van de samenleving. Het welzijn van kinderen, en ook hun opvoeding, is volgens hem een maatschappelijke opgave (De Winter, 2011). De afgelopen tien jaar heeft De Winter mede

aan de wieg gestaan van twee community-projecten, namelijk 'Allemaal opvoeders' (ALOP) en 'Vreedzame wijk' (VW). Bij Vreedzame wijk gaat het om het verspreiden en toepassen van een aantal principes die zijn ontwikkeld op basisscholen in de Utrechtse wijk Overvecht. Het gaat met name om een methode van conflicthantering door kinderen zelf (peer mediation) en om een grotere deelname van kinderen aan besluitvorming. Het project 'Allemaal opvoeders' gaat minder expliciet uit van een inhoudelijke methodiek, maar is meer gericht op het samenbrengen van zoveel mogelijk partijen voor een gedeelde pedagogische aanpak in buurten. (De Winter, 2011, bijv. pp. 88-89; Horjus & Van Dijken 2014; Pauw, 2013; Kesselring, 2016).

Nu de projecten enige jaren onderweg zijn, worden ze ook op wetenschappelijke wijze geëvalueerd. Wat leveren deze projecten op voor kinderen? De gebruikte evaluatiestrategieën (Kesselring, 2016; Horjus & Van Dijken, 2014) lijken nogal sterk te verschillen van het type evaluatie dat de sociaalwetenschappelijke wereld de afgelopen twintig jaar domineerde: het quasi-experimentele effectonderzoek. In dit essay beschrijf ik de beide methodologische strategieën en hun onderlinge verschillen en laat ik zien welke redenen er zijn voor die verschillen. Ik beschrijf ook enkele mogelijkheden voor onderzoekmatige oplossingen. Maar uiteindelijk blijven ook die oplossingen onbevredigend. De voornaamste tekortkoming van het evaluatieonderzoek tot nu toe, is namelijk dat dat onderzoek, ongeacht de gebruikte strategie, nog steeds te weinig zegt over de opbrengst van communities voor kinderen. Ik stel een radicaal andere wijze van evalueren voor die ook meer recht doet aan de democratische doelstellingen van community-projecten.

De Roethof-projecten en hun evaluatie

ALOP en VW zijn zeker niet de eerste projecten waarbij opvoeding en community gecombineerd werden. Al in 1818 startte generaal Johannes van den Bosch een grootschalig project in Fredriksoord en later in andere gemeenten waarbij kinderen een gezinnen in 'kolonies' werden geplaatst. Deelname aan dit project was gedwongen en de aanpak was bijzonder autoritair, maar doet in zijn omvattendheid soms verrassend hedendaags aan.¹

Gezien de aard en inhoud en uiteraard ook de historische context van de projecten staan de zogenaamde Roethof-projecten van rond 1990 een stuk dichterbij ALOP en VW. Deze projecten, vernoemd naar de PvdA-er

Hein Roethof, waren, overigens in tegenstelling tot ALOP en VW, in de eerste plaats bedoeld om criminaliteit te bestrijden via een buurtaanpak. De 104 projecten die metterdaad zijn uitgevoerd (Polder & Van Vlaardingen 1992, pp. 22 en 32) gebruikten sterk uiteenlopende interventies, zoals weerbaarheidscursussen voor vrouwen, groepsactiviteiten voor jongeren, de inrichting van schooltuinen, een opknopbeurt voor de buurt, bewaakte fietsstallingen, de aanstelling van huismeesters, arbeidsgewinningsprojecten voor jongeren, spijbelcontrole en begeleiding van voetbalsupporters (Polder & Van Vlaardingen, 1992, pp. 35-43; zie bijlage 4, pp. 100-116 voor een volledig overzicht).

De projecten waren in die tijd politiek omstreden; ze hadden een 'soft' imago vergeleken met de toen meer gebruikelijke op sancties gerichte aanpak van criminaliteit. Het WODC, het wetenschappelijk bureau van het Ministerie van justitie, liet daarom een meta-evaluatie uitvoeren, gebaseerd op de rapportages van afzonderlijke projecten (Polder & Van Vlaardingen, 1992). Deze meta-evaluatie verliep uiterst moeizaam. Er waren vooraf geen duidelijke criteria gesteld voor de evaluatie van de projecten. Mede daardoor werd bij slechts 63% van de projecten effecten gemeten, 30% was min of meer bruikbaar en bij 16% "(...) is deze informatie 'betrouwbaar' te noemen" (p. 3). Ondergetekende ontwierp een systematiseringsmethodiek om de grote inhoudelijke diversiteit van de 104 projecten te kanaliseren (Baerveldt, 1993). Dat leidde tot een systematisch overzicht van de strategische keuzes achter de interventies bij de projecten, maar de verwachte effecten van die keuzes op criminaliteit bleken moeilijk aantoonbaar. Polder en Van Vlaardingen (1992) geven wel enkele post hoc conclusies: zo leken projecten met meer inzet en middelen effectiever.

De rapportage riep veel vragen op over de kwaliteit van de evaluaties. Sommige projecten waren intern geëvalueerd (door hun managers), andere extern (door onafhankelijke onderzoekers). Bij de externe evaluaties werden gemiddeld driemaal zoveel knelpunten gemeld als bij de interne evaluaties. (p. 52). De auteurs eindigen dan ook met aanbevelingen voor verbetering van evaluaties (p. 81). Na de publicatie in 1992 werden vanuit het ministerie van Justitie steeds hogere eisen gesteld aan evaluaties. Veel van die eisen kwamen overeen met voorwaarden voor quasi-experimenteel onderzoek.

Het quasi-experimentele design in de Nederlandse sociale wetenschappen

Cook en Campbell (1979) gaven in hun klassieke boek een overzicht van problemen en oplossingen van effectonderzoek in het veld. Een van de belangrijkste inspiratiebronnen is voor hen het klassieke experimentele natuurwetenschappelijke onderzoek zoals het klassieke experiment waarmee Toricelli de effecten van hoogte op luchtdruk aantoonde in 1648 (pp. 3-4). Het idee was dat onderzoek naar effecten van interventies in de werkelijke wereld gebruik kon maken van de methoden die in laboratoria gebruikt werden. Het grootste probleem was daarbij dat veldonderzoek niet kon voldoen aan de strenge voorwaarden van experimenteel onderzoek. Om die redenen gebruikten Cook en Campbell de term 'quasi-experimenteel'.

Het basisidee van het quasi-experimentele onderzoek is dat een interventie X, zoals het toedienen van een medicijn, een bepaalde uitkomst Y verandert, bijvoorbeeld de bloeddruk verhoogt. In symbolen: $X \rightarrow Y$. Omdat de Y (bloeddruk) ook hoger kan worden door andere omstandigheden (bijvoorbeeld door stress of praten), is het mogelijk dat onderzoekers schijneffecten vinden. Cook en Campbell beschrijven dit soort 'validity problems' uitgebreid en geven er, voor zover mogelijk, oplossingen voor. Een van de problemen die we vaak in jeugdprojecten tegenkomen is het 'maturation' probleem (pp. 52). Men neme een interventie X die winkeldiefstal (Y) van 15 tot 18 jarige jongeren moet verminderen. Als de winkeldiefstal afneemt kan dat het gevolg zijn van de interventie, maar ook van het feit dat jongeren ouder worden en zich uit zichzelf volwassener gaan gedragen. Deze effecten komen vermoedelijk vaak voor. Een van de oplossingen is om het gedrag van de jongeren in de interventie door de tijd te vergelijken met een groep leeftijdsgenoten. Een ander probleem is (een vorm van) regressie naar het midden. Winkeldiefstal gaat, ook per jongere, vaak op en neer in golven. De kans dat jongeren in een interventie terecht komen is vaak groter als ze net even op een golf van veel plegen zitten. Tijdens het project gaat de golf naar beneden en dat kan ten onrechte worden gezien als een effect van de interventie. Tenslotte kan de inhoud van de interventie, de X, veranderen tijdens het onderzoek. Dat is vrij gebruikelijk in community-projecten, waar interventies vaak 'on the run' uitgebreid en ontwikkeld worden. Het gevolg is dat onduidelijk is waar een eventueel effect op winkeldiefstal (Y) vandaan komt; van de interventie in het begin, die aan het eind of die ergens daartussenin.

Cook en Campbell (1979) besteden uitgebreid aandacht aan de wijze

waarop validiteitsproblemen kunnen worden gerepareerd of moeten worden geaccepteerd. Veel, maar zeker niet alle, problemen kunnen volgens hen worden voorkomen door winkeldiefstal van de betrokken jongeren voor en na de interventie te meten en die naast metingen te leggen van gelijksoortige jongeren. Het is lastig te bepalen wanneer jongeren 'gelijksoortig' zijn. Daarom worden jongeren indien mogelijk bij toeval (at random) geloot voor deelname aan de interventie. Ook randomisatie kan overigens niet alle problemen voorkomen (pp. 56-58). De methodologische criteria voor de evaluatie van jeugdinterventies hebben zich de afgelopen decennia steeds meer bewogen in de richting van de normen uit de quasi-experimentele traditie. De norm in de academische wereld werd daarbij dat de onderzoeksopzet zelf al zoveel mogelijk alle mogelijke soorten van validiteitsproblemen moest voorkomen.

Het evaluatieonderzoek van community-projecten

Zonder direct een oordeel te vellen, ga ik nu in op de verschillen tussen de bij ALOP en VW gebruikte methoden met het quasi-experimentele design. Ik gebruik ALOP en VW daarbij als voorbeelden van community-projecten, wel wetende dat er een omvangrijke internationale traditie bestaat van soortgelijke projecten en soortgelijke evaluatiemethoden (zie bijvoorbeeld Robson, 1993). De inhoudelijke beschrijving van VW en ALOP komt al uitgebreid aan de orde in dit boek en veronderstel ik bekend.

Het doel (Y) van VW is om de sociale samenhang in de buurt te bevorderen. De ingezette interventie is verspreiding en brede toepassing van een op basisscholen ontwikkelde methodiek voor verbetering van conflicthantering en participatie door kinderen. De meeste concrete activiteiten zijn gericht op methodiekverspreiding (cursussen) en bevordering van samenwerking tussen professionals van verschillende instellingen. Het doel (Y) van ALOP is om te bevorderen dat burgers (meer) bereid zijn om de opvoeding van kinderen te delen met anderen. Het ingezette middel bestond uit het stimuleren daarvan door contacten te bevorderen tussen ouders en anderen met een (informele) opvoedersrol. Er waren geen vooraf bepaalde inhoudelijke interventies: de elf deelnemende gemeenten bepaalden zelf inhoud en aard van de concrete activiteiten.

Beide projecten werden en worden geëvalueerd. De gebruikte evaluatiemethoden wijken daarbij op drie punten af van de academische traditie van voorgaande decennia. In de eerste plaats wordt er door de onderzoekers geen quasi-experimenteel design gebruikt. Dat blijkt bijvoorbeeld uit

het feit dat de evaluatie niet primair gericht is op het aantonen van effecten van de projecten op de uitkomstvariabele (Y). Weliswaar worden in beide projecten metingen van uitkomsten verricht, bijvoorbeeld van de mate van samenwerking tussen professionals, maar die metingen spelen een veel minder grote rol dan bij quasi-experimenteel onderzoek. Daarentegen spelen andere 'variabelen' een veel grotere rol. Zo worden bij ALOP en VW ook opvattingen onderzocht over het nut van VW-methoden of opvoeden in de openbare ruimte. Een flink deel van het ALOP-onderzoek gaat met name over hinderpalen bij gezamenlijk opvoeden. Al deze kennis staat in verband met gedachten over werkzaamheid van de projecten, maar het gaat niet om metingen van de X of Y zoals in quasi-experimenteel onderzoek. Tenslotte ontbreekt ook de in quasi-experimenteel onderzoek gebruikelijke controlegroep.

In de tweede plaats verschillen de doelen van de evaluatie. Quasi-experimenteel onderzoek is over het algemeen summatief: de bedoeling is om achteraf een oordeel te kunnen vellen over een project. Het idee daarachter is dat een positief oordeel (effect) een politieke voorwaarde of beleidsvoorwaarde wordt voor het starten van soortgelijke projecten in de toekomst. De evaluaties van ALOP en VW zijn in de eerste plaats formatief: ze zijn bedoeld om kennis aan te leveren die de lopende projecten kan verbeteren. De kwalificaties summatief en formatief sluiten elkaar overigens niet helemaal uit. Resultaten van summatief (effect)onderzoek zijn ook bruikbaar voor verbetering van al lopende projecten. Het grootste nadeel van summatief onderzoek is traagheid, omdat het project over het algemeen start met een voormeting en de resultaten pas beschikbaar komen na de name-ting. Formatief onderzoek is meestal nuttiger als resultaten real-time binnenkomen.

In de derde plaats gebruiken de onderzoekers gemengde methoden van dataverzameling, met een sterke nadruk op kwalitatieve methoden. Het quasi-experimentele onderzoek is over het algemeen kwantitatief: de uitkomstvariabele is gebaseerd op de response van grote aantallen deelnemers, antwoordcategorieën zijn vooraf vastgelegd en de analyse gebeurt met behulp van statistische methoden. Kwalitatieve methoden, feitelijk een verzamelnaam voor alle methoden die niet kwantitatief zijn, maken gebruik van bijvoorbeeld (half) open interviews, focusgroepen, observaties en de inhoudsanalyse van websites of papieren documenten.²

Waarom geen quasi-experimenteel effectonderzoek?

Het voorafgaande laat zien dat effectonderzoek uit de gratie is bij de evaluatie van de twee community-projecten ALOP en VW. Voor een groot deel is dat het gevolg van problemen bij het toepassen van quasi-experimenteel onderzoek. Deze problemen betreffen de statistische methoden, de modellering, het onderzoeksontwerp en de interpretatie van resultaten.

In de eerste plaats zijn er statistische problemen. Om even 'exotisch' te beginnen: de gebruikelijke statistische testen zijn niet geschikt voor de evaluatie van community-projecten en kunnen zowel de kans op een effect als de grootte van dat effect fout inschatten. Deze testen veronderstellen namelijk dat het gedrag van de mensen in het onderzoek niet direct met elkaar samenhangt. In community-projecten is het echter juist de bedoeling dat mensen elkaars gedrag beïnvloeden. Methoden van netwerkanalyse zijn wel geschikt voor toetsing en schatting van effecten, maar die worden tot nu toe (bijna) niet toegepast. Belangrijker is ook dat de verwachte korte termijn-effecten van community-projecten klein zijn, net als bij (andere) preventieprojecten. De statistische power van het effectonderzoek van deze projecten is daardoor meestal te klein: het aantal deelnemers aan het onderzoek moet vele malen groter zijn dan in onderzoek naar, bijvoorbeeld, individugerichte curatieve interventies.

In de tweede plaats zijn er modelleringsproblemen. Het quasi-experimentele onderzoek gaat uit van invloed van een interventievariabele X op uitkomstvariabele Y : $X \rightarrow Y$. Het probleem is dat de X niet constant is in community-projecten. De inhoud van deze projecten staat namelijk niet bij voorbaat vast, maar verandert voortdurend, onder meer door de inbreng van de betrokkenen in het project. Er zijn nog wat mogelijkheden om dit door Cook en Campbell (1979) ook al gesignaleerde probleem aan te pakken. De meest voor de hand liggende is om zeer regelmatig, bijvoorbeeld wekelijks, zowel de inhoud van het project (X) als de uitkomsten (Y) te meten. Opnieuw lijken hoge kosten en nieuwe validiteitsproblemen (zoals onderzoeksmoeheid van deelnemers) het onderzoek in de weg te staan.

In de derde plaats is het bij community-projecten praktisch niet mogelijk om te voldoen aan alle criteria voor quasi-experimenteel effectonderzoek. Wanneer het effect van een project in een buurt geëvalueerd wordt, zou deze buurt bijvoorbeeld vergeleken moeten worden met een andere buurt waar dat project niet draait, maar die verder in alle opzichten gelijk zou moeten zijn. Dergelijke tweelingbuurten bestaan in de praktijk niet.

In de vierde plaats is het niet goed mogelijk om de resultaten van quasi-

experimenteel effectonderzoek te interpreteren. Een van de belangrijkste drijfveren achter summatief effectonderzoek is het idee dat van evaluaties lering getrokken kan worden voor de toekomst. Binnen het quasi-experimentele frame betekent dat vooral dat interventies waarvan achteraf een effect is aangetoond, worden voortgezet of elders worden toegepast. Bij wat complexere interventies en projecten is dat niet goed mogelijk. Het effect kan dan namelijk afhangen van een te groot aantal factoren. Welke factoren essentieel zijn, valt niet te achterhalen. Dat geldt des te meer voor community-projecten, die bijzonder complex zijn en qua vorm en inhoud sterk uiteenlopen. Daarom is niet duidelijk wat uit effectieve projecten 'gekopieerd' moet worden naar mogelijke nieuwe projecten.

Mogelijkheden voor effectonderzoek: meer maatwerk en meer kans

Omdat de praktijk en de behoeften van community-projecten zoals ALOP en VW niet lijken te sporen met de quasi-experimentele traditie, heeft het evaluatieonderzoek een sterk formatief karakter gekregen: het ondersteunt verbetering in de richting die het project al gekozen heeft. Het summatieve effectonderzoek lijkt daarentegen te verdwijnen. Dat zou betekenen dat evaluatieonderzoek geen rol meer zou spelen bij de keuze voor een bepaald type project in de toekomst: alle projecten lijken dan evenveel waard. Meer wetenschappelijk maatwerk en potentieonderzoek zijn dan alternatieven die sturing op basis van evaluatie meer mogelijk maken.

Wetenschappelijk maatwerk houdt in dat het evaluatieonderzoek alleen rekening houdt met validiteitsproblemen die gezien de aard van de interventie waarschijnlijk zijn. Bij onderzoek naar korte termijn-effecten worden jongeren nauwelijks ouder, hoeft eigenlijk geen rekening gehouden te worden met 'maturation' en is er wellicht minder reden om een controlegroep in het onderzoek op te nemen. Als het gaat om effecten van training van heel herkenbare vaardigheden, zoals de conflicthanteeringsmethodiek bij VW, geldt hetzelfde. Het ligt meer voor de hand om in dit geval meer aandacht te besteden aan andere validiteitsproblemen zoals regressie naar het midden. Gek genoeg staat dit soort maatwerk nog in een kwade reuk. In de quasi-experimentele traditie ligt de nadruk namelijk op het streven naar het perfecte onderzoeksdesign, bij voorkeur een 'gouden standaard' die o.a. uitgaat van voor- en nametingen en een controlegroep waarbij proefpersonen willekeurig toegewezen worden aan de experimentele groep (deelnemers aan het experiment) dan wel de controlegroep.

Praktisch gezien is deze standaard overkill en bovendien contraproductief: omdat hij meestal niet haalbaar is, krijgen de effectonderzoeken zelf vaker de onvoldoendes dan de onderzochte interventies. Er is overigens geen wetenschappelijke of wetenschapstheoretische reden voor het streven naar een gouden standaard en hij komt ook niet voor in het standaardwerk van Cook en Campbell (1979).³

Potentieonderzoek is een tweede alternatief voor effectonderzoek. Projecten worden dan beoordeeld op hun potentiële effect in plaats van op hun (onmeetbare) werkelijke effect. Zo veronderstelt het Nederlands Jeugdinstituut dat de kans op effectiviteit toeneemt als er een goede omschrijving is van de interventie (niveau 1) of een goede onderbouwing (niveau 2).⁴ De veronderstelling dat wie het beter zegt het ook beter doet, lijkt nogal zwak. Organisatorische voorwaarden (niveau 3) zoals het berekenen van de doelgroep lijken meer plausibel. Het argument wordt sterker als metterdaad kan worden aangetoond dat organisatie en inhoud van de interventies voldoen aan op theorie en eerder empirisch onderzoek gebaseerde voorwaarden voor effectiviteit. Rossi, Lipsey en Freeman (2004) stellen voor om de algemene theoretische veronderstellingen achter (interventies binnen) projecten expliciet uit te schrijven en deze veronderstellingen vervolgens te toetsen aan wetenschappelijke literatuur. Baerveldt, Horjus en De Winter (2008) toetsen projecten aan empirisch ondersteunde 'grote' modellen zoals planned behavior, rationele keuze, de participatieladder en criteria voor empowerment. De voordelen van deze vormen van potentieonderzoek is dat ze niet alleen laten zien waar lopende projecten 'gerepareerd' kunnen worden, maar ook welke criteria vooraf gesteld kunnen worden aan nieuwe projecten.

Van evaluatie van community-projecten naar evaluatie van jeugdnetwerken.

Hoewel summatief onderzoek meer mogelijk is dan de huidige praktijk suggereert, blijven er problemen die volgens mij duidelijk maken dat het evaluatieonderzoek naar communities over een geheel andere boeg zal moeten. Het zijn namelijk niet de projecten (organisatie en interventies) die moeten worden geëvalueerd, maar de relatie tussen kinderen en het netwerk van alle mogelijke relevante instellingen en personen (jeugdnetwerken) om hen heen.

Ik heb daarvoor drie argumenten. Ten eerste lijkt de inhoud van de interventies binnen community-projecten minder belangrijk dan de orga-

nisatie; met name de omvattendheid van het project is belangrijk. VW is gestart met een inhoudelijke interventie, namelijk een methodiek voor conflictbeheersing door kinderen. Gaandeweg is het accent echter verschoven naar het organiseren van een breed draagvlak voor de VW. Bij ALOP gaat het eigenlijk nauwelijks nog om de inhoud van de interventie: het programma is primair gericht op de organisatie van alle mogelijke partijen. De doelstelling van community-projecten is minder inhoudelijk en veel meer procesgericht geworden. Het gaat er uiteindelijk niet om of de interventies in het project effectief zijn, maar of de community effectief is. Dat is geen toeval; feitelijk was dat ook altijd al de bedoeling van dit type projecten. De consequentie is dat evaluatievormen niet de interventies maar de rol van de community voor het kind centraal moeten stellen. En dat gebeurt nu nog niet.

Ten tweede kan de omvang (het domein) van de X niet goed gedefinieerd worden binnen een project. Projecten denken vanuit bestaande organisaties en vanuit de problemen die organisaties hebben, bijvoorbeeld qua expertise, continuïteit en onderlinge afstemming. Kinderen leven in een groot netwerk van personen en instellingen die zich met hen bemoeien. Het bestaat uit een groot aantal uiteenlopende zelfstandige organisaties die een rol spelen voor kinderen en jongeren, zoals scholen en kinderdagverblijven, jeugd- en buurtcentra, politie, artsen, orthopedagogen en andere jeugdspecialisten, sportverenigingen, lokale jeugdwelzijnsinstellingen en culturele initiatieven. Een deel daarvan blijft onzichtbaar zolang evaluaties gaan over projecten. Projectevaluaties laten niet zien wat er voor het kind ontbreekt, omdat dat domein niet zichtbaar is voor de deelnemende organisaties in het project. Kinderen weten bijvoorbeeld of het nuttig voor hen is om een speciale plek te hebben voor huiswerkbegeleiding; maar dat geldt niet automatisch voor de organisaties die het project vormen. Ze laten ook niet zien welke afstemmingsproblemen tussen organisaties direct voelbaar en belangrijk zijn voor kinderen.

Ten derde moet de evaluatiemethode in lijn komen met het principe van empowerment in community projecten. Empowerment betekent dat mensen zelf bepalen of, hoe en waarvoor ze ondersteund worden en vereist dat de communities zelf gaan bepalen of en welke interventies nodig zijn. Omdat evaluaties een onderdeel vormen van een democratisch verbeterproces, zal ook de focus van die evaluaties gedemocratiseerd moeten worden. Dat betekent dat de evaluatie moet gaan over de ervaringen van kinderen (en andere burgers) met het netwerk om hen heen. Dit past ook goed in de lijn die Micha de Winter (1995) heeft uitgezet in zijn werk over de rol van kinderen als medeburgers. Ik stel daarom voor dat het summa-

tieve onderdeel van de evaluatiemethode van community projecten gaat over de kwaliteit van het netwerk zoals kinderen die ervaren. Het identificeren van belangrijke verbeterpunten, een belangrijk doel van formatief onderzoek, moet daar ook starten.

Mijn voorstel is dus om de primaire focus van de inhoud van community-evaluaties van het project te verleggen naar de ervaringen van kinderen met de jeugdnetwerken: het gehele netwerk van personen en instanties om hen heen. Het gaat dan niet meer primair om de inhoud en organisatie van het project, maar om het signaleren van kwaliteit (summatief) en het wegnemen van bestaande knelpunten voor kinderen in het hele netwerk van instellingen en personen waarmee ze te maken hebben (formatief).

De kwaliteit van de community wordt getoetst door criteria die voor kinderen belangrijk zijn en het hele netwerk als een geheel betreffen (zie bijvoorbeeld Van Beek & Rutjes, 2009). Dat zijn vaak simpele criteria zoals zichtbaarheid van wensen, gedrag of problemen van kinderen bij minstens een van de betrokken instellingen of personen; de snelheid waarmee door minstens een van hen daarop gereageerd wordt; consistentie tussen de reacties tussen verschillende betrokken personen of organisaties; begrijpelijk en doeltreffend doorlopen van de totale weg naar een oplossing. Community-verbetering zou er niet op gericht moeten zijn om een bepaalde aanpak 'uit te rollen' over een wijk, maar om ervoor te zorgen dat aan deze criteria voor kinderen wordt voldaan. Summatief onderzoek is daarbij nodig om aan te geven wat de actuele kwaliteit is van het netwerk voor kinderen. Uitkomsten kunnen zichtbaarheidspercentages zijn, reactiekanalen en -tijden, waarderingen van kinderen voor consistentie, communicatie en doeltreffendheid. Het formatieve onderzoek geeft aan waar de knelpunten liggen en genereert suggesties vanuit de betrokkenen.

Conclusie

Er zijn weinig redenen voor zorg om het feit dat de evaluatiemethoden van community-projecten VW en ALOP op veel punten afwijken van de gebruikelijke standaards bij quasi-experimenteel onderzoek. Quasi-experimentele onderzoekdesigns blijken om diverse onderzoekstechnische redenen namelijk ongeschikt voor (summatieve) evaluaties van community-projecten en zijn ook minder geschikt (formatief) ter ondersteuning van verbetering van die projecten. Er zijn bovendien nog wat mogelijkheden, door meer methodologie op maat en door potentieonderzoek, om het summatieve onderzoek van community-projecten met een andere metho-

dologie te versterken. Voor zover er op dit moment sprake zou zijn (wat ik overigens betwijfel) van een methodenstrijd in evaluatieland, dan betreft dat vanuit een community-perspectief slechts een achterhoedegevecht. Er zijn nieuwe uitdagingen die belangrijker zijn.

Voor communities zijn de grote projecten met de grote gedachten minder belangrijk dan de 'kleine' verbeteringen die (dat is democratie) worden aangegeven door de directbetrokkenen. Evaluatiemethoden moeten daar beter op aansluiten. De nieuwe uitdaging is om bij evaluaties van communities niet een project en bijbehorende interventies centraal te stellen, maar het kind en het volledige netwerk van alle mogelijke relevante instanties en personen om haar/hem heen. Ik stel daarom voor dat de focus van evaluaties verschuift naar de ervaringen van kinderen met dat netwerk. Onderzoek zal dan in de eerste plaats gericht zijn op kwaliteit voor en volgens kinderen en vervolgens op de verbeteringen die zij nodig achten. De eerste ervaringen met dit type onderzoek laten zien dat het mogelijk is om dergelijk onderzoek uit te voeren en betaalbaar te houden, maar ook dat er nieuwe methodologische en inhoudelijke uitdagingen ontstaan. Laten we ze aannemen.

Noten

1. Meer informatie: <http://www.dekoloniehof.nl/wp>
2. Het verschil tussen de beide typen evaluaties is overigens iets minder groot dan gesuggereerd. Zo gebruiken de ALOP- en VW-onderzoekers ook kwantitatieve methoden (een enquête) in hun onderzoek. De rol van dit materiaal is echter beperkt: het is meestal bedoeld om veronderstellingen van het project of generaliseerbaarheid van bepaalde bevindingen te testen. Net zo gebruiken onderzoekers in quasi-experimenteel onderzoek, meestal zonder ze zo te noemen, kwalitatieve methoden om te bepalen in hoeverre projecten op elkaar lijken (X-analyse).
3. Het is wel overigens wel mogelijk om de procedures te standaardiseren en meer te objectiveren. Dat kan bijvoorbeeld inhouden dat onderzoekers of een specialistenforum per project de meest voor de hand liggende en bedreigende validiteitsproblemen identificeren en concretiseren, en alleen daar verder rekening mee houden. De aanpassingen van het design, stringente validiteitscontroles bij dataverzameling en extra monitoring kunnen dan beperkt worden tot wat echt nodig is. Wie dit graag wil kwantificeren, zou zelfs een Bayesiaans aandoende waarschijnlijkheidsscore kunnen berekenen voor het effect, die qua interpretatie enigszins vergelijkbaar is met die van een statistische toets.
4. De effectladder van het Nederlands Jeugdinstuut, <http://www.nji.nl/De-effectladder>.

Referenties

- Baerveldt, C. (1993). Het gebruik van rationele keuzereconstructies bij de evaluatie van interventies. *Tijdschrift voor criminologie*, 35(2): 158–76.
- Baerveldt, C., Horjus, B., & De Winter, M. (2008). From evaluation to interpretation: Assessing theoretical elements in residential programs. *Evaluation*, 14(3), 345-362.
- Cook, T.D., & Campbell, D.T. (1979). *Quasi experimentation: Design and analysis issues for field settings*. Boston, MA: Houghton Mifflin Co.
- De Winter, M. (1995). *Kinderen als medeburgers: Kinder- en jeugdparticipatie als maatschappelijk opvoedingsperspectief*. Utrecht: De Tijdstroom.
- De Winter, M. (2011). *Verbeter de wereld, begin bij de opvoeding: Vanachter de voordeur naar democratie en verbinding*. Amsterdam: SWP.
- Horjus, B., & Van Dijken, M. (2014). De ecologie van maatschappelijke betrokkenheid. In K. Jurrius & H. Toppinga (Eds.), *Opvoeden in betrokkenheid. Lessen uit de wetenschap*. (pp. 69-88). Amsterdam: SWP.
- Kesselring, M.C. (2016). *Partners in Parenting: A study on shared childrearing responsibilities between parents and nonparental adults*. Enschede: Ipskamp.
- Pauw, L. (2013). *De Vreedzame School: Onderwijs en de maatschappelijke vorming van de leerling*. Amsterdam: SWP.
- Polder, W., & Van Vlaardingen, F.C. (1992). *Preventiestrategieën in de praktijk: Een meta- evaluatie van criminaliteitspreventieprojecten*. Arnhem: Gouda Quint.
- Robson, C. (1993). *Real world research*. Oxford, UK: Blackwell.
- Rossi, P.H., Lipsey, M.W., & Freeman, H.E. (2004). *Evaluation: A Systematical Approach. Seventh Edition*. Thousand Oaks, CA: Sage.
- Van Beek, F., & Rutjes, L. (2009). *Kwaliteitstandaarden Jeugdzorg: Wat kinderen en jongeren belangrijk vinden als ze niet thuis wonen*. Houten: Bohn Stafleu van Loghum.

Over de auteur

Chris Baerveldt is socioloog en werkt momenteel als wethouder sociaal domein in IJsselstein.

E-mail: chrisbaerveldt@mac.com