

Het ontwerpen van case-studies: enkele keuzen

Peter Swanborn*

Summary

Designing case-studies: Some dilemmas

Two problems in designing case-studies are tackled: (a) selecting one versus selecting several cases; (b) how to select cases. Existing strategies such as selecting cases on the basis of maximizing or minimizing variance between the cases are discussed. Introduction of the often implicit difference between selecting on the independent or the dependent variables leads to a rejection of selection on the dependent variables.

1. Inleiding

Het methodologisch erf wordt tegenwoordig gedomineerd door multivariate analyse, of, iets algemener, het kwantificerende onderzoek. Een niet onbelangrijke nevenstroom, vooral in het moderne opdrachtonderzoek, wordt echter gevormd door 'case-studies'. Op verschillende terreinen – bijvoorbeeld het bestuurs-, organisatie- en onderwijskundig onderzoek – is deze vorm van onderzoek sterk vertegenwoordigd. Vooral de combinatie van case-studies met grootschalig survey-onderzoek is, zeker in Nederlands opdrachtonderzoek, bijna standaard geworden.

De methodologische literatuur terzake heeft echter een uiterst bescheiden omvang. Bekend zijn Yin (1989, 1984) en Miles en Huberman (1984). Voor wat betreft Nederlandstalige teksten beschikken we over van Hoesel (1985) en

* Hoogleraar methoden en technieken aan de Faculteit der Pedagogische en Onderwijskundige Wetenschappen van de Universiteit van Amsterdam.

recentelijk Hutjes en Van Buuren (1992). De doorsnee onderzoeker heeft het al een heel eind geschopt als hij deze titels bestudeerd heeft.¹ Slechts enkelen zijn gedoken in klassieke bijdragen op dit terrein, zoals die van Campbell (1975), Eckstein (1975), George (1979), Kennedy (1979), Lijphart (1971) en McClintock (1979).

Het karakter van case-studies is de laatste decennia sterk veranderd. Terwijl vroeger het fundamentele onderzoek via de idiografische, kwalitatieve, explorerende 'single-case'-aanpak de toon aangaf, is de laatste jaren de aandacht verschoven naar praktijkgericht toepassingsonderzoek, dat kwalitatieve en kwantificerende elementen combineert. De huidige populariteit van termen zoals 'beleidsexperiment' en 'evaluatieonderzoek' (die immers tegelijkertijd vaak als 'case-studies' getypeerd kunnen worden) illustreert dit.

Mede door de betrekkelijk geringe aandacht voor methoden van de case-study is een aanzienlijke spraakverwarring hiervan het gevolg. Terwijl sommigen het begrip case-study inweven in een kwalitatief paradigma, stellen anderen het voor als zeer wel passend in de empirisch-analytische traditie. Hutjes en Van Buuren (1992) worstelen met de problematiek van een inductieve benadering doordat zij hun startpunt kiezen in het kwalitatieve onderzoek; bij Merriam (1988) en Patton (1990) constateren wij tot mislukking gedoemde pogingen om aanbevelingen te formuleren die uit contrasterende paradigma's worden afgeleid. Het begrip case-study is sowieso moeilijk te omlijnen als onderzoeksstrategie, omdat typologieën van strategieën immers gebaseerd zijn op uiteenlopende dataverzamelmethodeën. Bij een case-study ligt echter, zoals het woord zegt, het accent op het bestuderen van één eenheid; een case-study is in principe combineerbaar met bijna alle vormen van dataverzameling, en daardoor moeilijk te plaatsen in een opsomming van strategieën.

Een motief achter deze bijdrage is een zekere ergernis over de vanzelfsprekendheid waarmee juist de combinatie van survey en case-studies in praktijkgericht onderzoek tot onderzoeksvoorstellen en rapporten leidt waaruit onvoldoende duidelijk wordt wat de case-studies toevoegen aan de kennis die een enquête oplevert, of omgekeerd. Veel onderzoekers hebben de slechte gewoonte om de resultaten van hun case-studies niet samen te vatten of te systematiseren, wat zij wel met survey-resultaten doen (vgl. ook Yin, 1993, p. 70). Men schijnt ervan uit te gaan dat beschrijvingen en interviewverslagen voor zich zelf spreken (dat lijkt zeker de veronderstelling te zijn als interviewcitaten e.d. gebruikt worden om de rapportage van een enquête te verlevendigen), en dat hangt samen met het impliciet laten van een specifieke probleemstelling, conform de situatie bij één van de historische bronnen van de case-study: het etnografische kwalitatieve onderzoek. Richtinggevend voor deze bijdrage is daarom de gedachte dat de keuzen die een onderzoeker in het ontwerp maakt in

de eerste plaats bepaald worden door de probleemstelling, en dat in het kader van een 'multi-strategy design' voor elk onderdeel, dus ook voor eventuele case-studies, (a) zorgvuldig vooraf verantwoord moet worden welke problemen men ermee wil oplossen, en (b) achteraf verantwoord moet worden in hoeverre dat gelukt is. Onduidelijkheid over de probleemstelling manifesteert zich in het design van een case-study, en wel primair in de keuze van de cases. In dit artikel inventariseren we daarom de oplossingen die onderzoekers kiezen op twee centrale keuzemomenten, en we onderwerpen deze aan een kritische analyse:

- a. het kiezen van één resp. méér cases;
- b. indien meer cases, hoe deze te kiezen.

2. Overwegingen rondom de probleemstelling

Bij het ontwerpen van onderzoek in het algemeen, en dus ook bij het stellen van een probleem voor een case-study, spelen enkele basisoverwegingen een rol:

- met welke type vragen hebben we te maken: beschrijvings-, verklarings- of ontwerp- (veranderings-)problemen?
- hoe is de stand van onze kennis over het object van studie, en hoe kunnen we met het oog daarop efficiënt onze probleemstellingen formuleren?

Wat de eerste overweging betreft: het traditionele onderscheid tussen beschrijvings- en verklaringsproblemen is voor het ontwerpen van case-studies weinig relevant. Vrijwel altijd komen beide aan de orde, met dien verstande dat één type van beschrijvingsvragen gewoonlijk niet via case-studies beantwoord (kan) worden: die betreffende frequentieverdelingen: 'hoeveel?-vragen'. Maar: de vraag of veranderingsproblemen aan de orde komen leidt ons tot het belangrijke onderscheid tussen het fundamenteel- versus praktijkgericht zijn van de onderzoeksdoelstelling, dat wil zeggen: willen we kennis produceren die voor het onderzoekersforum bedoeld is, dan wel kennis die in ieder geval bruikbare resultaten voor (veranderings)beleid moet opleveren. En als dit niet uit de probleemstelling duidelijk wordt, moet het als achtergrondkennis van de onderzoeker toch meespelen. Praktijkgericht (of: beleids-)onderzoek kent in ieder geval als extra regulerend idee dat het als oplossing voor veranderingsproblemen bruikbare resultaten moet opleveren.

Wat de tweede overweging betreft, de efficiency van een onderzoek neemt toe naarmate de probleemstelling qua uitwerking beter afgestemd is op de stand van onze kennis. Naar de mate van uitgewerktheid kunnen we probleem-

stellingen onderverdelen in 'onderwerpaanduidingen'; 'globale vragen'; 'concrete vragen' en 'hypothesen' (Swanborn, 1991, p. 61-69). In de communis opinio heeft een case-study nog vooral een explorerende functie: het opdoen van ideeën over welke variabelen in het model de relevante zouden kunnen zijn (en welke waarden deze moeten hebben), kortom: globale vragen. Echter, talloos zijn de puur beschrijvende case-studies waarin concrete vragen gesteld worden; daarnaast zijn vele voorbeelden van hypothese-toetsende case-studies bekend. De mate van uitgewerktheid van de probleemstelling dient met andere woorden de stand van onze a priori kennis te weerspiegelen, en die kan bij case-studies lopen van 'nauwelijks enig inzicht' tot 'zeer uitgebreide kennis'.

3. Cases 'op zich', of cases als 'pars pro toto'

In de probleemstelling moet bovendien het domein duidelijk worden gemaakt: gaat het om 'cases op zich' of cases als 'pars pro toto'? Willen we kennis verzamelen over deze bepaalde case hier en nu, of selecteren we één of meer cases als steekproef, als pars pro toto uit een (veel) breder domein waarover we uitspraken willen doen?²

Bij case-studies uit vroeger tijden ging het meestal om deze case hier en nu; ook in tegenwoordig opdrachtonderzoek is dat nogal eens het geval (bijv. de opdracht betreft dit ziekenhuis, of dit politiebureau). Maar meestal toch staan de onderzochte cases voor een veel grotere verzameling. Yin (1989) vat case-studies zelfs uitsluitend als pars pro toto studies op.

Dat dit zeer relevante onderscheid zo weinig aandacht krijgt, is eveneens oorzaak van spraakverwarring. Vaak wordt zonder meer gesproken over de (tekortschietende) externe validiteit of generaliseerbaarheid van 'de' case-study. Men realiseert zich dan blijkbaar niet dat in het geval van 'cases op zich' het domein van onderzoek zich niet verder uitstrekt dan de bestudeerde cases, en dat de betreffende discussie dus irrelevant is.

4. Eén case of meer cases?³

Bij 'case(s) op zich' is er geen keuzeprobleem voorzover het opdrachtonderzoek betreft: de te bestuderen sociale systemen zijn eenvoudig gegeven (tenzij er een groot aantal cases, elk voor zich, bestudeerd moet worden, zodat een steekproef nodig is; maar zodra iets dergelijks het geval is gaat het toch meestal om een 'pars pro toto'-studie). En als het fundamenteel onderzoek betreft,

wordt, als er in principe een keuze mogelijk is, gewoonlijk op triviale gronden (nabijheid tot de universiteit!) gekozen.

Als er daarentegen sprake is van een echte 'pars pro toto'-studie, kiezen we dan één case, of meerdere cases? Laten we eerst enkele situaties bekijken waarin volgens Yin met één case kan worden volstaan.

Het is niet verwonderlijk dat hier in eerste instantie naar analogie van het 'cruciale experiment' de 'cruciale case-study' wordt genoemd. Wanneer één of meer theorieën zodanig gepreciseerd zijn dat concurrerende voorspellingen kunnen worden afgeleid onder gespecificeerde condities, en die condities zijn in een bepaald geval gerealiseerd, dan zou het onderzoek van één case kunnen volstaan om een theorie te steunen of te weerleggen (althans tot modificatienopen). Allison's *Cuban missile crisis*, Lipset's *Union democracy* en Festinger's *When prophecy fails* zijn standaardvoorbeelden. Verderop zal duidelijk worden (een case-study is geen experiment!) dat wij hier vraagtekens bij plaatsen.

Een tweede situatie ontstaat wanneer cases van het gezochte type tamelijk uniek zijn, zoals patiënten met een zeldzaam syndroom. Het is dan echter eerder praktische noodzaak dan methodologisch wijs om zich tot één case te beperken. Een derde situatie doet zich voor bij een 'revelatory case', dat wil zeggen wanneer de onderzoeker bij unieke uitzondering toegang heeft tot een bepaalde groep, situatie of persoon. Soms ontstaan dergelijke situaties bij toeval, zoals wanneer een onderzoeker aangewezen wordt als lid van een recht-bankjury of betrokken raakt bij een gijzelingsactie. Linschoten (1964, p. 87) verhaalt in zijn prachtige boek over Ernst Mach, die tijdens een treinreis een beroerte kreeg, en van de situatie gebruik maakte door systematisch na te gaan welke bewegingen hij nog wel, en welke niet kon maken, waarbij hij zichzelf als een black box tussen wil en beweging beschouwde.

Over een fraai voorbeeld dat overigens onder alledrie de bovengenoemde condities tegelijk past werd gerapporteerd door Koenen (1994) in *NRC-Handelsblad* van 3 februari 1994: een meisje dat vanaf haar geboorte tot op dertienjarige leeftijd in vrijwel volstrekte afzondering had doorgebracht vormde een testcase bij uitstek voor de concurrerende taalontwikkelingstheorieën van Chomsky en Lennenberg.

Wanneer we verder echter afzien van dit type unieke (maar overigens buitengewoon belangrijke) situaties doet zich de vraag voor of we één, dan wel meerdere cases zullen onderzoeken. Yin stelt – en dat is een principieel punt – dat niet de simpele 'sample logic' moet worden gevolgd van 'hoe meer cases, hoe beter'. Er zijn nu eenmaal vrijwel nimmer voldoende cases te onderzoeken om betrouwbare conclusies over het bedoelde domein van cases te formuleren. Volgens Yin moeten we cases niet als onderzoekseenheden beschouwen, maar, in plaats daarvan, een case-study zien als een experiment in de traditie van de

laboratoriumpsychologie: het onderzoek dient om een theorie te toetsen. Er is niet sprake van een 'sample-to-population generalization', maar van 'analytic generalization' van een case-study naar een theorie (of model).

De term 'analytic generalization' heeft inmiddels, als contrasterend met statistische generalisatie, in de literatuur op dit terrein vaste voet gekregen: vergelijk bijvoorbeeld Firestone (1993). Firestone noemt in navolging van Kennedy (1979) nog een derde vorm van generalisatie, die specifiek zou gelden voor case-studies: de 'case-to-case generalization'. Het betreft hier echter de inschatting van de bruikbaarheid van case-study-resultaten vanuit het perspectief van de gebruiker, dus met het oog op de 'eigen case'. Men kan hier dan ook moeilijk van een wetenschappelijke strategie of procedure voor de onderzoeker spreken.

Yin stelt naar onze mening de vergelijkbaarheid van case-study en experiment in zijn redenering wel erg rooskleurig voor. Kenmerkend voor een experiment is de zorgvuldigheid waarmee getracht wordt de onderzochte situatie te modelleren zodat zoveel mogelijk alleen modelvariabelen een rol spelen en geïsoleerd worden van foutenbronnen; kenmerkend voor een case-study is juist de natuurlijke situatie. Het interpreteren van resultaten van één case in termen van een theorie, is, afhankelijk van de stand van onze kennis, een grote dan wel een zeer grote gok; de kans dat niet in het model opgenomen variabelen een rol spelen, of dat we simpel met een toevalstreffer te maken hebben, is aanzienlijk.

Zonder dat Yin dit dilemma expliciet erkent of ontkent noemt hij toch enkele manieren om hier boven uit te stijgen, zoals het werken met concurrerende theorieën binnen één case. Deze remedie is naar onze mening buiten de context van het formuleren van een nulhypothese niet van veel praktisch belang (immers: hoe vaak blijkt niet uit grootschalig onderzoek dat zowel de ene 'cruciale variabele' als de andere een rol speelt?). Een andere remedie is het zo ver mogelijk expliciteren van het nomologisch netwerk, met andere woorden het formuleren van zo veel mogelijk toetsbare voorspellingen. Een derde remedie, ten slotte, is dat niet volstaan wordt met één case, maar dat gezocht wordt naar replicatie van resultaten via meerdere cases. Met deze suggesties kunnen we alleen maar instemmen.

Een voorlopig antwoord op de vraag 'hoeveel cases moeten we kiezen' hangt primair af van de rol die we aan het onderzoek van diverse cases toekennen; gaat het om 'wat zekerder weten', of gaat het er juist om verschillen te ontdekken. In de volgende paragraaf wordt hierop ingegaan. Naar onze mening zal echter meestal een triviaal eerste antwoord gegeven kunnen worden: hoe meer cases, des te betere controlemogelijkheid op de betrouwbaarheid van het resultaat en de geldigheid van model of theorie. Binnen het raamwerk van de financiële ruimte in opdrachtonderzoek gesteld trachten we daarom zo veel

mogelijk cases te onderzoeken. Een grens wordt echter gesteld door de kosten/baten ratio: zo heeft bij voorbeeld de zoveelste replicatie onder exact gelijke condities een zeer kleine kans om weerlegd te worden als alle vorige niet weerlegd zijn, en is dus nauwelijks meer interessant (vgl. Hofstee, 1980). Bij de keuze van cases zal vooral de efficiëncy in het oog moeten worden gehouden.

5. Hoe cases te kiezen?

Wanneer het de onderzoeker vrij staat om uit een groter universum één of enkele cases te kiezen, resteert de vraag op welke criteria cases geselecteerd worden. De importantie van deze keuze wordt onderstreept door Yin (Yin 1993, p. 12) die voor een bepaald geval signaleert dat 20 procent van het onderzoeksbudget ging zitten in een zorgvuldige voorbereiding van de selectie van cases. We gaan hieronder in op verschillende selectieprocedures.

Een traditioneel onderscheid, voor zover mij bekend geïntroduceerd door Glaser en Strauss (1976, hfst. 3) is dat tussen minimalisering en maximalisering van variantie. Met andere woorden: zoeken we cases die zo veel mogelijk op elkaar lijken, of juist cases die zo veel mogelijk van elkaar verschillen? De auteurs bevelen aan om eerst te minimaliseren teneinde een stevige basis voor 'de theorie' te leggen, daarna te maximaliseren om de reikwijdte te beproeven. Het is een voor de hand liggende werkwijze die we ook in modern onderzoek veel zien. Glaser en Strauss laten in het midden om welke variabelen het gaat; het probleem is natuurlijk dat we over een model of theorie moeten beschikken om te weten op welke variabelen we juist niet of juist wel variantie moeten nastreven. De oplossing van dit typisch methodologische probleem ligt in een 'bootstrap'-benadering waarbij uitgaande van zwakke subjectieve kennis een eerste poging wordt gedaan; op basis daarvan een volgende stap wordt gezet (of een stap terug wordt gedaan), enzovoort. Kortom: exploratie, ook in de keuze van de volgende eenheden.

Glaser en Strauss spreken niet over afhankelijke en onafhankelijke variabelen (ze spreken überhaupt niet over variabelen en eenheden, maar vervallen in een jammerlijke terminologische vervaging door over categorieën en hun kenmerken te spreken; het heeft aan kwalitatieve onderzoekers veel tijd gekost om hier vanaf te komen, maar dat terzijde!). Wanneer we echter over selectieprincipes voor case-studies spreken is dit onderscheid van groot praktisch belang. Tot nu toe is in de literatuur, voorzover ons bekend, in dit verband het onderscheid tussen afhankelijke en onafhankelijke variabelen niet gemaakt. Vanuit de lange 'kwalitatieve' voorgeschiedenis van de case-study, waarbin-

nen onderzoekers zich niet van dit onderscheid bedienen, is dat ook wel begrijpelijk. Maar zeker in het huidige beleidsonderzoek, waarin het vrijwel altijd om de bepaling van effecten gaat, is het cruciaal.

In eerste instantie passen we de principes minimalisering en maximalisering toe op relevant geachte onafhankelijke variabelen of condities. Minimalisering van variantie op onafhankelijke variabelen kan men zien als: replicatie 'to be sure': Yin spreekt hier van literal replication. Leveren cases die op de door ons relevant geachte onafhankelijke variabelen sterk op elkaar lijken hetzelfde resultaat op? Het is een verstandige eerste optie als we nog weinig op het terrein thuis zijn. We voegen hieraan toe: bij de selectie van cases zullen we, ongeacht verdere principes, cases kiezen van een type dat frequent voorkomt. Het is niet verstandig om, als de financiële mogelijkheden beperkt zijn, alle categorieën van een onafhankelijke variabele te vertegenwoordigen als daar enkele zeer weinig vóórkomende bij zijn.

Het uitsluitend afgaan op frequentie van voorkomen van bepaalde 'conditie-combinaties' kan soms het doorslaggevende criterium zijn bij de selectie van cases, vooral als we hoegenaamd niets over de effecten onder die verschillende condities weten. We spreken dan van selectie van enkele 'doornee-' of 'typische' gevallen. Maximalisering van variantie, qua fasering volgend op de voorafgaande procedure, komt bij exploratie neer op onderzoek naar de reikwijdte van het model of de gevoeligheid voor het variëren van parameters, en bij een toetsende aanpak op het kiezen van cases die volgens de theorie tegenstelde, of in ieder geval uiteenlopende, resultaten zouden moeten opleveren. Heel grof gebeurt dit in de praktijk wanneer we de variantie maximaliseren op enkele relevant geachte onafhankelijke variabelen (bijv. door uit elke cel van het cartesisch produkt der variabelen één case te kiezen). Ook nu beperken we ons tot die gecombineerde condities waarvan bekend is dat ze redelijk veel voorkomen. Yin spreekt, bij een toetsende aanpak, van 'theoretical replication'. Methodologisch komt dit neer op de bekende raadgeving om zoveel mogelijk voorspellingen af te leiden en te toetsen. Wederom geldt echter het dilemma: een verstandige keuze is alleen mogelijk op basis van kennis die we meestal nog niet hebben. Toch kan aan de praktijk heel wat verbeterd worden. Zo worden in de Nederlandse onderwijsresearch scholen voor case-studies vrijwel strijk en zet geselecteerd op landstreek en zuil, zonder dat er duidelijke redenen zijn om te veronderstellen dat dit voor het te onderzoeken probleem relevante variabelen zijn. Wanneer het automatisme om te selecteren op bepaalde standaardvariabelen wat meer plaats maakt voor reflectie, en raadpleging van het wetenschappelijk forum en velddeskundigen vooraf, kan dit leiden tot een meer efficiënte variabelenkeuze. In eerste instantie kan men zich daarbij richten op die variabelen waarover een zekere communis opinio be-

staat; in tweede instantie op die variabelen waarover deskundigen tegengestelde meningen hebben. Consultatie van deskundigen is de aangewezen manier om niet alleen contrasten-opleverende en daarom interessante variabelen op het spoor te komen, maar ook de redenen daarvoor: de veronderstelde werking. Naïeve exploratie van materiaal kan daarmee een eindweegs in de richting van toetsing van veronderstellingen verschoven worden.

De situatie ligt anders als we selecteren op de afhankelijke variabele, en dat komt in de huidige onderzoekspraktijk, bij voorbeeld in het raam van evaluatieonderzoek, het meest frequent voor.

Kijken we eerst naar selectie op basis van minimalisering van variantie in de afhankelijke variabele. Selectie van enkele 'success-stories' waarbij een analyse van de veroorzakende condities en hun werking het doel is, komen we frequent tegen; beleidsinstanties vragen nogal eens om 'best practice'-s case-studies! Een dergelijk design is zwak, omdat niet bekend is of wellicht dezelfde condities ook bij de minder succesvolle cases optreden: een controle- of contrastgroep ontbreekt. Tot veel meer dan enkele tentatieve vooronderstellingen kunnen we niet komen.

Meestal komt selectie op de afhankelijke variabele(n) dan ook neer op het selecteren van enkele succes-cases en enkele mislukkings-cases: met andere woorden: maximalisering van variantie. Soms wordt zelfs in aanbestedingen voor opdrachtonderzoek met zoveel woorden één van deze designs, waarbij 'gezocht moet worden naar kenmerken van enkele goed, resp. slecht lopende, cases' al gesuggereerd. Toch zijn er genoeg redenen om uiterst terughoudend te zijn met een dergelijk ontwerp. Bij selectie van cases op de afhankelijke variabele (of: variabelen, hetgeen het er nog lastiger op maakt) schuilen er immers enkele kolossale adders onder het gras. De eerste daarvan is het regressie-effect: we selecteren meestal volgens de meetresultaten van één moment, en we kunnen er geenszins zeker van zijn dat deze zo stabiel zijn dat de selectie op een ander tijdstip niet anders uitgevallen zou zijn. Bij meetfouten en instabiliteit is te verwachten dat de geselecteerde extreme groepen in feite minder extreem zijn dan we dachten. Wanneer we er in slagen om onze selectie te baseren op de extremen van scores die gemiddeld zijn over enkele tijdstippen kunnen we het regressie-effect enigszins ontlopen. Maar bij de evaluatie van vaak nog maar kortlopende experimenten zijn we vaak al blij als er ten minste één effectmeting kon plaatsvinden... De tweede is de 'confounding' van onafhankelijke variabelen. Blalock (1964) heeft al gewezen op het grote nadeel van selectie op de afhankelijke variabele: door succesvolle van mislukte cases te onderscheiden, construeert men in feite twee groepen die contrasteren op een aantal oorzaken tegelijk; het resultaat is dat die oorzaken nauwelijks uit elkaar geplozen kunnen worden; de onafhankelijke variabelen zijn hopeloos

'confounded'. We beschrijven dan hoogstens enerzijds situaties waarin een aantal factoren tegelijk als 'positief' gelden, en enkele andere situaties waarvoor het omgekeerde geldt. Maar inzicht in de relatieve sterkte van deze oorzaken, in hun onderlinge verwevenheid en mogelijke interacties, kunnen we hoe dan ook niet verwerven.

De bezwaren tegen deze wijze van selectie zijn zo ernstig, dat een algemene ontrading passend is. Hoogstens zou er aan te denken zijn als een eerste check op de groslijst van variabelen waar we hierboven over spraken. Maar zelfs dan zou het efficiënter zijn om, als we bij voorbeeld vier cases kunnen onderzoeken, meteen enige contrasten aan te brengen op de belangrijkste onafhankelijke variabelen.

Een verbijzondering van deze selectiemethode, die meer aan het eind van een programmacyclus past, treffen we nog in de praktijk aan in de vorm van onderzoek van 'afwijkende cases': hoe komt het dat op een groot aantal gevallen een theorie past, maar op enkele niet? Zorgvuldige conditionering van een theorie zou in principe van een dergelijke analyse het resultaat kunnen zijn, alsmede het formuleren van een diepere (meeromvattende) theorie waarin ook de afwijkende cases kunnen worden ondergebracht. Maar in de praktijk leidt dit type van onderzoek meestal tot niet veel meer dan een beschrijving van cases in de vorm van een complexe structuur van variabelen waarmee dan inzichtelijk gemaakt zou worden waarom deze case niet aan het algemene model voldoet; aan de lezer moet worden overgelaten 'waarin het 'm nu eigenlijk zit'.

Ten slotte, als laatste selectieprincipe, kan vermeld worden de selectie van cases die in een verschillende fase van ontwikkeling verkeren (voor zover mij bekend voor het eerst genoemd door Van Hoesel (1985)). Deze procedure zien we nogal eens in het raam van het meest populaire toepassingsgebied van case-studies: dat naar de factoren die innovatie bevorderen resp. remmen. Waar een compleet innovatietraject soms vele jaren beslaat terwijl voor het onderzoek een beperkte periode beschikbaar is, ligt het voor de hand om op deze wijze cases te selecteren. Nadat de keuze van subpopulaties op basis van tijdstippen of cohorten gemaakt is, kan men een random verdere selectie van cases volgen, of één van de bovengenoemde principes van selectie op onafhankelijke variabelen toepassen (selectie op de afhankelijken is niet van toepassing omdat zeker de cases in de eerste 'fasen' nog niet zover zijn). Selectie op ontwikkelingsfase kan zeer efficiënt zijn, maar berust op de meestal niet toetsbare veronderstelling dat de verschillende generaties van cases vergelijkbaar zijn. Daardoor ontstaat enige interpretatie-ambigüiteit waarmee onderzoekers in de praktijk nogal eens worstelen. De problematiek is een zwakke afspiegeling van

de niet-identificeerbaarheid van effecten in longitudinaal onderzoek van het age/period/cohort (APC)-type.

Samenvattend komen we tot de volgende lijst van selectieprincipes:

1. minimalisering van variantie op veronderstelde oorzaken, waaronder 1a: selectie van 'doorsnee'-gevallen.
2. maximalisering van variantie op veronderstelde oorzaken;
3. minimalisering van variantie op gevolg(en);
4. maximalisering van variantie op gevolg(en), waaronder 4a: selectie van 'deviant cases';
5. selectie van cases in verschillende fasen van ontwikkeling.

Werkwijze 1 staat in dienst van onderzoek naar de betrouwbaarheid van onze conclusies, de 'stevigheid'; werkwijze 2 staat in dienst van onderzoek naar de reikwijdte; werkwijze 1 gaat logisch vooraf aan werkwijze 2. De werkwijzen ad 3 en 4 dienen in het algemeen afgeraden te worden. Zelfs indien we het regressie-effect kunnen vermijden zijn deze procedures niet efficiënt omdat we geen vat kunnen krijgen op het onderlinge gewicht van de diverse oorzaken. Werkwijze 5 ten slotte lijkt, al dan niet in combinatie met 1 of 2, een goede aanpak.

Dat 'de methodologie van de case-study' nog steeds in het licht staat van een explorerend en/of beschrijvend wetenschapsdoel wordt geïllustreerd door de literatuur over de analyse in het geval men met verschillende cases te maken heeft. We komen daarbij de volgende alternatieven tegen.

Eén van de genoemde mogelijkheden is dat men de cases los van elkaar beschrijft, analyseert en rapporteert; in zekere zin de minst veeleisende en goed in een idiografische benadering passende werkwijze. Of het onderzoek van de diverse cases gelijktijdig dan wel achtereenvolgens plaatsvindt is dan van weinig belang. In de literatuur staat deze procedure bekend als de case-comparison approach (Yin & Gwaltney, 1981).

Het andere uiterste is dat men case A analyseert; een voorlopig model bouwt; dit model toetst aan case B; tot bijstelling komt zodat A en B gedekt worden door het model; vervolgens daarmee naar case C gaat, enzovoort, tot alle cases binnen een gemeenschappelijk model passen. Deze benadering staat bekend als de 'controlled-comparison approach' (Diesing, 1971; George, 1979).

Een derde alternatief, dat veelal wordt gekozen als men het tweede probeert maar daarmee geen succes heeft, is dat men typen ontwerpt: voor sommige cases kan men een model maken, voor andere cases een ander model. Niet zelden een verlegenheidsoplossing: een typologie is nu eenmaal van een lager

niveau dan een theorie. Voor verklaringsproblemen is uiteraard het meest informatief om in eerste instantie te proberen één theorie te formuleren die voor alle onderzochte cases geldt.

Het onderscheid zoals hier beschreven is vooral in een explorerende aanpak relevant, en de betekenis ervan is daarom beperkt. Werkt men immers vanuit een theorie dan is tevoren duidelijk of 'een volgende case' een zelfde, dan wel juist een contrasterend resultaat in vergelijking met een eerdere case moet opleveren. Bij een op toetsing gerichte probleemstelling volgt zowel de keuze van de cases als de wijze van analyse uit de theorie, hetgeen natuurlijk niet uitsluit dat men bij voorbeeld tot de ontdekking kan komen dat verschillende sets van condities, verschillende modellen, tot eenzelfde resultaat kunnen leiden.

Ook als we echter de hypothesenvormende functie van case-studies, zeker in de gebruikelijke combinatie met een survey, centraal plaatsen, is, mede vanwege de hoge kosten, het hanteren van een efficiënt design en daarmee de keuze van de case(s) van groot belang.

Noten

1. Yin (1989, 1984) is kort, helder en duidelijk geënt op praktijkgericht (organisatie)onderzoek. De 'herziening' van 1989 is een verkooptruc, die slechts gebaseerd is op toevoeging van enkele pagina's over het gebruik van theorie bij case-studies. Het is overigens de beste korte tekst die we ter oriëntatie kunnen aanraden. Recentelijk verscheen Yin (1993) waarin tal van uitgewerkte toepassingen te vinden zijn. Het vormt een nuttige aanvulling op zijn eerdere werk. Yin gaat hier ook in op 'ethnographic research' en 'the grounded theory approach', waartegen hij zijn eigen werkwijze duidelijk afzet (Yin 1993, pag. 46; 57 e.v.). Hutjes en Van Buuren (1992) behandelen in de eerste honderd pagina's eigenlijk 'kwalitatief onderzoek' in meer algemene zin, en beschrijven in de laatste honderd pagina's nogal onkritisch de analyseprocedures van Miles en Huberman. Het boek van Miles en Huberman zelf (1984) heeft als belangrijkste bezwaar dat het veel te dik (en te kostbaar!) is, en dat de parade van analysetechnieken (die grotendeels triviaal zijn, in die zin dat een verstandige onderzoeker ze meestal zelf kan bedenken) buitengewoon onoverzichtelijk blijft. Een verwijzing naar 'Miles en Huberman' wordt naar mijn ervaring door velen gezien als een 'must' in subsidieaanvragen voor onderzoek waarin iets kwalitatiefs wordt gedaan. Helaas wordt dan vrijwel nooit erbij gezegd welke techniek men precies gaat gebruiken; alsof in kwantitatief onderzoek een verwijzing naar 'Hays'; 'Kerlinger', 'Pedhazur' voldoende zou zijn om te verduidelijken wat men gaat doen!

Zeer recentelijk verscheen een geheel vernieuwde (paperback) editie van Miles en Huberman, met 'more than hundred new techniques', hetgeen na de reeds vele tientallen matrix- en tabeltechnieken van de eerste editie niet vrolijk stemt.

Van Hoesels werk (1985) over case-studies is veel minder bekend, waarschijnlijk omdat het betreffende gedeelte wat verborgen is in een dissertatie onder een geheel andere titel. Van Hoesels eigen case-study, en zijn analyse van een aantal andere, is echter verdienstelijk.

2. De criteria fundamenteel/praktijkgericht, en case(s) op zich versus cases als pars pro toto hebben we in onderstaand schemaatje gecombineerd, en elke cel van een voorbeeld voorzien om te laten zien dat elk van de vier typen voor de praktijk van belang is.

	fundamenteel		praktijkgericht	
case(s) op zich	Jansens 'De straat'	A	B	Utrechtse politie
cases als pars pro toto	Boys in White	C	D	Amersfoorts politiebureau

Toelichting bij A. De Utrechtse socioloog G.H. Jansen schrijft een boekje over een typisch volksbuurtstraatje waarin iedereen de ander kent, ouderwetse feesten worden gehouden, enzovoort. De auteur wil niet generaliseren naar andere straten, laat staan naar volksbuurten in het algemeen; er is geen beleidsbelang aan verbonden: typisch fundamenteel wetenschappelijke interesse (de vraag is wel: wat doet de wetenschap ermee? en: is het wel wetenschap?).

Toelichting bij B. In verband met de toekomstige huisvesting, en de wens om afdelingen die veel contact met elkaar hebben in elkaars nabijheid te huisvesten, wordt onderzocht hoe intensief de interactie is tussen verschillende afdelingen van de Utrechtse politie. Er is geen enkele generalisatiepretentie.

Toelichting bij C. Een groep Amerikaanse sociologen doet onderzoek op een medische faculteit naar de wijze waarop aankomende studenten zich identificeren met hun nieuwe vakgebied en met hun docenten. Het is een 'case', maar met generalisatiepretenties naar vele andere socialisatiecontexten.

Toelichting bij D. Op een Amersfoorts politiebureau waar naast mannelijke sinds kort ook een aantal vrouwelijke agenten werkt wordt door een stagiaire onderzoek gedaan in het kader van de anti-discriminatiewetgeving naar mogelijk discriminerend gedrag bij mannelijke agenten. De onderzoeker werkt aan en binnen een 'case', maar deze is voor haar slechts een voorbeeld dat staat voor vele andere politiebureaus die het beleidsdomein vormen. Ook zogeheten beleidsexperimenten passen in deze categorie.

Deze 2x2 typologie kan desgewenst een derde dimensie krijgen door de indeling van probleemstellingen naar uitgewerktheid toe te voegen. Dat is alleen van academisch belang; ons doel is natuurlijk niet om een sluitende typologie te ontwerpen alswel om de relevante dimensies waarlangs men case-studies kan onderscheiden te signaleren.

3. In diverse typologieën van case-studies treffen we dit onderscheid aan. Yin (1989, 1984) werkt met twee criteria:

- single case designs versus multiple case designs;
- holistic (single unit of analysis) versus embedded (multiple units of analysis).

Het 'single case/multiple case'-onderscheid is simpel, en leidt – zoals Yin benadrukt – niet tot uiteenlopende methoden. Het holistic/embedded-onderscheid leidt wel tot uiteenlopende werkwijzen. Het erbij halen van een subniveau – soms zelfs meerdere subniveaus – is een mogelijkheid om, zoals men wel zegt, 'de N te vergroten', of: 'kwantitatieve analyse mogelijk te maken'. Dat laatste is echter middel, geen doel.

McClintock (1979) heeft zich in een vroeg stadium verdiept in het werken met subeenheden

binnen een case-study. Hij gebruikt hiervoor de term 'case-cluster'-methode; Miles en Huberman zijn hiermee veel verder gegaan. In het algemeen heeft de multi-niveau-analyse sindsdien mogelijkheden verschaft om dergelijke data adequaat te analyseren (hetgeen nog niet inhoudt het adequaat conceptualiseren van een multi-niveau-model).

Literatuur

- Blalock, H.M. (1964). *Causal inferences in non-experimental research*. Chapel Hill: University of North Carolina Press.
- Campbell, D.T. (1975). Degrees of freedom and the case-study. *Comparative Policy Studies*, 8, 178-193.
- Diesing, P. (1971). *Patterns of discovery in the social sciences*. Chicago: Aldine Atherton.
- Eckstein, H. (1975). Case study and theory in political science. In: F.I.Greenstein & N.W. Polsby (red.), *Strategies of inquiry* (pp. 79-137). Reading, MA: Addison-Wesley.
- Firestone, W.A. (1993). Alternative arguments for generalizing from data as applied to qualitative research. *Educational Researcher*, 22, 4, 16-23.
- George, A.L. (1979). Case studies and theory development: The method of structured, focused comparison. In: P.G. Lauren (red.), *Diplomacy: New approaches in history, theory and policy* (pp. 43-68). New York: Free Press.
- Glaser, B.G. & Strauss, A.L. (1967). *The discovery of grounded theory*. Chicago: Aldine.
- Hoesel, P.H.M. van (1985). *Het programmeren van sociaal beidsonderzoek: analyse en receptuur* (dissertatie). Maastricht: Rijksuniversiteit Limburg.
- Hofstee, W.K.B. (1980). *De empirische discussie*. Amsterdam: Boom.
- Hutjes, J.M. & Buuren J.A. van, (1992). *De gevalstudie; strategie van kwalitatief onderzoek*. Amsterdam/Heerlen: Boom/Open Universiteit.
- Kennedy, M.M. (1979). Generalizing from single case studies. *Evaluation Quarterly*, 3, 4, 661-678.
- Koenen, L. (1994). Een modern wolvenkind. *NRC Handelsblad*, 3 februari.
- Lijphart, A. (1971). Comparative politics and the comparative method. *American Political Sciences Review*, 65, 682-693.
- Linschoten, J. (1964). *Idolen van de psycholoog*. Utrecht: Bijleveld.
- McClintock, C.C. e.a. (1979). Applying the logic of sample surveys to qualitative case studies: The case cluster method. *Administrative Science Quarterly*, 24, 612-629.
- Merriam, S.B. (1988). *Case study research in education: A qualitative approach*. San Francisco: Jossey-Bass.
- Miles, M.B. & Huberman, A.M. (1984). *Qualitative data analysis*. Beverly Hills: Sage.
- Patton, M.Q. (1990). *Qualitative evaluation and research methods* (2e dr.) Newbury Park: Sage.
- Ragin, C.C. (1987). *The comparative method*. Berkeley: University of California Press.
- Swanborn, P.G. (1991). *Basisboek sociaal onderzoek*. Amsterdam: Boom.
- Yin, R.K. (1989/1984). *Case study research; Design and methods* (herziene dr.). Newbury Park: Sage.
- Yin, R.K. (1993). *Applications of case study research*. Newbury Park: Sage.
- Yin, R.K. & Gwaltney, M.K. (1981). Knowledge utilization as a networking process. *Knowledge, Creation, Diffusion, Utilization*, 2, 555-580.