

Basisboek sociaal onderzoek

Basisboek sociaal onderzoek

Peter G. Swanborn

Zesde druk

Boom Lemma uitgevers
Amsterdam
2015

Voorwoord

Dit boek is vooral gericht op het begrijpen van wat onderzoeken in de maatschappij- en gedragswetenschappen nu eigenlijk is. De nadruk ligt op basisprincipes van wetenschappelijk onderzoek, of dat nu kwalitatief of kwantitatief is. Het verwoorden van een probleemstelling, het kiezen van een onderzoeksstrategie en van de benodigde soort data en het vermijden van fouten en valkuilen krijgen alle aandacht. Het is een ‘verbindende tekst’ waarin de lezer de gemeenschappelijke basis kan vinden voor allerlei specifieke vormen van onderzoek, gebruikte termen, procedures en technieken.

Tegenwoordig is het immers heel goed mogelijk om via internet actuele informatie te krijgen over allerlei specifieke methoden en technieken van onderzoek. In een ‘basisboek’ komt het daarom aan op het begrijpen *waarom*.

De hoofdstukken over beschrijvende statistiek die in eerdere edities waren opgenomen, zijn dan ook verhuisd naar de bijbehorende website www.basisboeksocialonderzoek.nl. Op deze website is ook informatie over methoden te vinden die eerder in bijlagen was opgenomen. De omvang van het boek is daarmee aanzienlijk afgenomen. Op deze manier is het met recht ‘een basisboek’ geworden dat de kern bevat van waar het in sociaalwetenschappelijk onderzoek om gaat.



Het boek is – samen met de website – goed te gebruiken als algemene inleiding in methoden van onderzoek. In vergelijking tot andere inleidingen in de methoden en technieken geeft dit boek een wat dieper gaande benadering voor degenen die zich vragen stellen over het ‘waarom’ van bepaalde methodische of technische procedures.

Het tegenwoordige sociaalwetenschappelijke onderzoek is in belangrijke mate praktijkgericht. Dat is dit boek eveneens, maar er wordt ook ruimschoots aandacht besteed aan het werken met theorieën en de rol van modellen in onderzoek. Bovendien zijn bij elk hoofdstuk nogal wat literatuurtips en vermeldingen van websites toegevoegd, die aanknopingspunten bieden voor lezers die meer willen weten over bepaalde onderwerpen. Ook is elk hoofdstuk ruimschoots voorzien van oefeningen.

De tekst is gebaseerd op de Engelse versie (2009) van het *Basisboek*, dat na de eerste Nederlandse editie van 1991 verschillende keren werd aangevuld en gewijzigd. Het is bedoeld voor studenten in het eerste of tweede jaar van hun opleiding, maar ook voor belangstellenden uit alle hoeken van de samenleving, die meer willen weten over hoe modern sociaalwetenschappelijk onderzoek nu eigenlijk werkt en waarom bepaalde keuzen worden gemaakt.

Inleidingen op dit terrein verouderen snel, en dat hoort ook zo bij een vakgebied dat zich snel ontwikkelt. Het is steeds weer een uitdaging om te beslissen wat verouderd is en weggelaten kan worden, en welke nieuwe ideeën en werkwijzen de moeite waard zijn om vermeld te worden. Ik hoop dat veel lezers er het een en ander van opsteken, maar vooral ook dat zij plezier krijgen in het boeiende vakgebied van sociaalwetenschappelijk onderzoek.

Peter G. Swanborn
Oosterbeek, april 2015

Inhoud

Voorwoord	5
1 Sociaalwetenschappelijk onderzoek in vogelvlucht	11
1.1 De fasen van een onderzoeksproject	11
1.2 De gereedschapskist van onderzoekers	17
1.3 Praktijkgericht onderzoek	28
Oefeningen	35
Literatuurtips	36
Kernbegrippen	37
2 Probleemstellingen	39
2.1 Waarom beginnen we met een probleemstelling?	39
2.2 Een voorbeeld: onderzoek naar schoolverzuim	40
2.3 Beschrijvingsproblemen	45
2.4 Verklaringsproblemen	47
2.5 Ontwerpproblemen	52
2.6 Het begin van een onderzoeksproject: oriëntatie	54
Oefeningen	61
Literatuurtips	63
Kernbegrippen	63
3 Data en onderzoeksstrategieën	65
3.1 Inleiding	65
3.2 Een panorama van databronnen	67
3.3 Bestaande data	70
3.4 Nieuwe data	79
3.5 De relatie tussen databronnen en onderzoeksstrategieën	94
Oefeningen	96
Literatuurtips	98
Kernbegrippen	100
4 Case studies	101
4.1 Wat is een case study?	101
4.2 Probleemstellingen bij case studies	105
4.3 Het kiezen van de case(s)	110
4.4 Rijkdom aan data	112
4.5 De participerende onderzoeker	113
4.6 Soorten observaties	115
4.7 Analyse van de data	116
4.8 Samenvatting	120

	Oefeningen	123
	Literatuurtips	124
	Kernbegrippen	124
5	Surveys: soorten, steekproeven en non-respons	125
	5.1 Wat is een survey?	125
	5.2 Een vergelijking van surveytypen	129
	5.3 Herhaalde surveys	135
	5.4 Respondenten als informanten	137
	5.5 Aselecte steekproeven	138
	5.6 Selecte steekproeven	147
	5.7 Non-respons	150
	Oefeningen	155
	Literatuurtips	157
	Kernbegrippen	157
6	Surveys: de vragenlijst	159
	6.1 Inleiding	159
	6.2 Vragen	161
	6.3 Antwoordvormen	168
	6.4 'Niet van toepassing' en 'geen mening'	171
	6.5 Samengestelde meetinstrumenten	172
	6.6 De lay-out van een vragenlijst	181
	Oefeningen	184
	Literatuurtips	184
	Kernbegrippen	185
7	Experimenten	187
	7.1 Wat is een experiment?	187
	7.2 Het uitschakelen van de effecten van andere variabelen	192
	7.3 Uitbreidingen van het basisexperiment	192
	7.4 Tijdreeksen	193
	7.5 Efficiency	194
	7.6 Alternatieven voor een experiment	196
	7.7 De bestudering van langetermijneffecten	202
	7.8 Evaluatieonderzoek	205
	Oefeningen	208
	Literatuurtips	209
	Kernbegrippen	209
8	Kwaliteitscriteria	211
	8.1 Het wetenschappelijk forum	211
	8.2 Weerlegbare kennis	213
	8.3 Kwaliteitscriteria voor onderzoek	216
	8.4 Betrouwbaarheid	216

8.5	Validiteit van instrumenten	218
8.6	Validiteit van onderzoeksconclusies over causaliteit	220
8.7	Intermezzo	221
8.8	Generaliseerbaarheid	222
8.9	Bruikbaarheid (voor praktijkgericht onderzoek)	224
	Oefeningen	225
	Literatuurtips	227
	Kernbegrippen	228
9	De logica van wetenschappelijk onderzoek	229
9.1	Modellen	229
9.2	De empirische cyclus	233
9.3	Theorieën	237
9.4	Stijlen van onderzoek	241
	Oefeningen	252
	Literatuurtips	253
	Kernbegrippen	253
10	Causale conclusies bij het gebruik van surveydata	255
10.1	Inleiding	255
10.2	De tijdsvolgorde	256
10.3	De uitschakeling van derde variabelen	257
10.4	Andere modellen	265
	Oefeningen	268
	Literatuurtips	270
	Kernbegrippen	270
11	Ethiek van sociaalwetenschappelijk onderzoek	271
11.1	Inleiding	271
11.2	De onderzoeker en zijn proefpersonen	272
11.3	De onderzoeker en zijn 'missie' tegenover de wetenschap	276
11.4	De onderzoeker en de setting waarin hij werkt	277
11.5	De onderzoeker en het publiek	281
	Literatuurtips	282
	Websites over ethische codes	282
	Bijlage A: Cohen's kapp	283
	Bijlage B: Variantie en het ontbinden van variantie	287
	Register	293
	Over de auteur	297

Sociaalwetenschappelijk onderzoek in vogelvlucht

1

Sociaalwetenschappelijke onderzoeksprojecten zijn er in allerlei vormen. In dit hoofdstuk bespreken we de vijf achtereenvolgende fasen die we bij vrijwel elk onderzoek tegenkomen. Daarna kijken we naar 'de gereedchapskist' van een onderzoeker. We bestuderen 'wetenschappelijke uitspraken', waarin eenheden, variabelen en scores aan de orde komen. In bijna alle onderzoeken zijn dit belangrijke begrippen. Ze helpen ons ook om de gemeenschappelijke structuur te zien in uiteenlopende manieren om onderzoek aan te pakken.

Aan het eind van dit hoofdstuk ligt het accent op praktijkgericht onderzoek. Praktijkgericht onderzoek mikt op de beantwoording van vragen die betrekking hebben op sociale problemen. Meestal worden dergelijke problemen op het bureau van de onderzoeker neergelegd door klanten, zoals beleidsinstanties of bedrijven.

1.1 De fasen van een onderzoeksproject

In het algemeen kunnen we vijf fasen onderscheiden:

- het formuleren van de probleemstelling;
 - het ontwerpen van het onderzoek;
 - het verzamelen van de benodigde gegevens;
 - de analyse van die gegevens;
 - het schrijven van het onderzoeksverslag en het overdragen van de resultaten aan anderen.
-

Hierna beschrijven we elk van deze vijf fasen in het kort. Maar let op: niet altijd komen deze vijf zo keurig na elkaar in een onderzoek voor. Zo kan het zijn dat tijdens het verzamelen van de gegevens de probleemstelling opnieuw geschreven moet worden, omdat de bedoelde gegevens niet allemaal verkregen kunnen worden. Of het verzamelen en het analyseren van de gegevens lopen door elkaar, doordat na een voorlopige eerste analyse blijkt dat je deels andere, of nog aanvullende, gegevens nodig hebt. En ten slotte: je doet er meestal verstandig aan

om al meteen bij de start van het onderzoek voor ogen te hebben hoe je rapport eruit moet zien, en vanaf het begin stukjes tekst te schrijven.

Probleemstellingen

Omdat elk onderzoek bedoeld is om een probleem op te lossen, starten we elk onderzoeksproject met één of meer *probleemstellingen*, ook wel *onderzoeksvragen* genoemd. Een probleemstelling is niets anders dan de vraag die de onderzoeker wil beantwoorden (in hoofdstuk 2 staan hier veel meer details over). Voorbeelden zijn:

- Hoe is de groep werkzoekenden in deze stad samengesteld naar leeftijd, geslacht en opleiding?
- Is de kans dat je in de openbare ruimte met agressieve scheldpartijen wordt geconfronteerd, afhankelijk van het aantal aanwezige personen?
- Komen psychosomatische aandoeningen meer voor bij lager geschoolden en in bepaalde leeftijds- en beroepsgroepen, en hoe is dit te verklaren?
- Welke beleidsmaatregelen kun je goed gebruiken om winkeldiefstal in grote steden tegen te gaan?
- Wat voor type kinderen loopt kans in de schoolklas het slachtoffer te worden van pesten?
- Welke knelpunten doen zich voor bij de inkoop van informatietechnologie door ondernemingen in de culturele sector?
- Maakt de aanwezigheid en toegankelijkheid van expliciete seks op tv de rol van de ouders bij de seksuele opvoeding gemakkelijker of juist moeilijker?
- Gegeven drie bètablokkers: welke heeft het sterkste effect op het rijgedrag?

Vaak is het vaststellen van de precieze probleemstelling een taak voor de onderzoeker, maar vooral in praktijkgericht onderzoek zijn soms al probleemstellingen geformuleerd door de opdrachtgever. Toch moet de onderzoeker ook dan de gestelde vragen meestal opnieuw formuleren, zodat ze werkelijk onderzocht kunnen worden. Dit komt vaak neer op het inperken van de oorspronkelijke vragen.

Dikwijls is er sprake van meer dan één onderzoeksvraag. Soms is er één centrale, overkoepelende vraag, waarbij het aan de onderzoeker wordt overgelaten om deze te herformuleren tot een reeks van onderzoekbare, preciezer genoteerde probleemstellingen (zie hoofdstuk 2 voor een voorbeeld).

Probleemstellingen of onderzoeksvragen moeten niet worden verward met de *doeleinden van het onderzoek*. Het doel van een onderzoek is altijd het vinden van een antwoord op de onderzoeksvraag, het oplossen van het probleem. Of, wat indrukwekkender gezegd: het verkrijgen van kennis, het verdiepen van het inzicht. Maar die verkregen inzichten kunnen heel verschillende doeleinden

dienen. We maken een onderscheid tussen onderzoek dat vooral gericht is op het maken en uitbreiden van wetenschappelijke theorieën, het zogeheten theoriegerichte of fundamentele onderzoek, en onderzoek dat vooral dient om te helpen bij het oplossen van praktische problemen: het praktijkgerichte onderzoek. De resultaten van praktijkgericht onderzoek worden gebruikt om bestaande situaties waarmee mensen niet tevreden zijn, te veranderen in meer gewenste situaties. Praktijkgericht onderzoek mikt op het ontwerpen van beleidsingrepen (daarom wordt het ook vaak beleidsonderzoek genoemd, al is dit niet precies hetzelfde), en het wordt dikwijls uitgevoerd op basis van een contract met een cliënt, de opdrachtgever en/of een subsidiërende instantie (daarom wordt het ook wel eens contractonderzoek genoemd).

Probleemstellingen komen niet uit de lucht vallen. Vooral bij praktijkgericht onderzoek is het nuttig en zelfs noodzakelijk voor de onderzoeker om het een en ander te weten over waar die vraag om onderzoek te doen eigenlijk vandaan komt. De posities en rollen van diverse belanghebbenden, de verwachtingen die men heeft van de onderzoeksresultaten en de vraag wat men eigenlijk met die resultaten wil doen, kortom, de gehele ‘setting’ waar de opdracht vandaan komt, is van belang. Al deze zaken vormen de *achtergrond van de probleemstelling*, niet te verwarren met de probleemstelling zelf. Een korte bespreking van die achtergrond stellen we uit tot aan het eind van dit hoofdstuk. Op dit moment is het van meer belang iets te leren over de elementaire fasen van elk onderzoek dan om je te verdiepen in de ingewikkelde rollen en onderlinge relaties van allerlei belanghebbenden bij het onderzoek.

Ten slotte moet je een probleemstelling niet verwarren met een *onderzoeksplan of onderzoeksvoorstel*. Deze termen verwijzen naar een systematisch en gedetailleerd overzicht van de activiteiten die de onderzoeker zich voorneemt te gaan doen: de tijd die elke fase in beslag gaat nemen, de kostenbegroting, de strategie die gevolgd gaat worden, waar men de gegevens wil halen, enzovoort. Al deze zaken zijn belangrijk, maar zij maken geen deel uit van de probleemstelling.

Het ontwerpen van het onderzoek

Na het formuleren van de probleemstelling(en) vragen we ons af hoe het onderzoek eruit moet zien om tot een oplossing van die probleemstelling te komen. Een van de eerste zaken waar het hier om gaat, is de vraag welke soort gegevens je nodig hebt. Historische gegevens? Gegevens over de situatie nu? Gegevens over het gedrag van mensen, over hun houdingen en meningen? Gegevens van een heel grote verzameling van mensen of gegevens over een bepaalde lokale groep? De aard van de gewenste gegevens bepaalt in belangrijke mate de algemene onderzoeksstrategie die gevolgd moet worden.

Er zijn vijf hoofdgroepen van strategieën:

- *bureauonderzoek* (met gebruikmaking van internet, rapporten van anderen, archieven, persoonsgegevens, beschikbare documenten en statistisch materiaal (zie hoofdstuk 3));
- *groepsbenaderingen*, zoals focusgroepen en Delphi-procedures (zie hoofdstuk 3);
- *kleinschalig onderzoek*, zoals het interviewen van een klein aantal personen, gedragsobservatie ‘in het veld’, case studies, historisch onderzoek (zie hoofdstuk 4);
- *grootschalig onderzoek*, zoals een enquête (zie de hoofdstukken 5 en 6);
- *experimenten* (zie hoofdstuk 7).

Binnen een gekozen strategie moeten verdere keuzen worden gemaakt. Bij een enquête kun je bijvoorbeeld kiezen voor een survey¹ via internet, via de telefoon, via de post of via een interviewcampagne. Soms worden verschillende benaderingen, waarvan elke is afgestemd op een bepaalde deelgroep, naast elkaar gebruikt in hetzelfde onderzoeksproject.

Bij sommige probleemstellingen ligt de te kiezen strategie erg voor de hand. Bij andere is het niet zo gemakkelijk uit te maken aan welke strategie je de voorkeur moet geven. Het maken van een keuze veronderstelt dat je precies weet welke verzameling mensen, groepen of organisaties je wilt bestuderen en wat voor steekproeven je kunt trekken. We moeten ook weten welke eigenschappen heel gemakkelijk zijn te meten, en voor welke eigenschappen je nog meetinstrumenten moet maken. Het vastleggen van de hoofdlijnen van de analyse van de gegevens valt ook onder het ontwerpen van een onderzoek. Het document waarin alle beslissingen en keuzen ten slotte worden vastgelegd, wordt het *onderzoeksontwerp* genoemd.

Het verzamelen van de benodigde gegevens

Daarna volgt het verzamelen van de gegevens die je nodig hebt. ‘Gegevens’ zijn antwoorden op vragen, waarnemingen van het gedrag van mensen en aflezingen van instrumenten. Voor ‘gegevens’ gebruiken onderzoekers tegenwoordig altijd het woord ‘data’. Ook als een onderzoeksontwerp er volkomen plausibel uitziet, gezien de probleemstelling, kunnen er praktische hindernissen zijn bij het verzamelen van data. Rapporten van andere onderzoekers zijn soms niet op tijd beschikbaar, statistische gegevens kunnen verouderd zijn, een survey kan te kampen krijgen met veel non-respons, enzovoort. Daarom is het soms een beetje improviseren tijdens de dataverzameling, en moet je het vervolg van

1 We gebruiken de woorden *survey* en *enquête* door elkaar; de betekenis is hetzelfde.

een onderzoeksproject wel eens wat anders opzetten dan in het oorspronkelijke ontwerp voorzien was.

Het analyseren van de gegevens

Als je de gegevens hebt verzameld, ben je er nog lang niet. Het resultaat moet eerst worden geanalyseerd, wat betekent dat je de resultaten gaat interpreteren met de probleemstelling in het achterhoofd. Als de gegevens zijn verzameld en in een datamatrix zijn weergegeven (zie voor een voorbeeld kader 1.4), ligt een oplossing van de probleemstelling nog niet meteen voor de hand. We geven een voorbeeld:

Voorbeeld

De erfelijkheid van politieke voorkeur



Om een vraag over de 'erfelijkheid' van politieke voorkeur te beantwoorden hebben we bijvoorbeeld een grote groep van jonge mensen (laten we zeggen vierhonderd personen) gevraagd naar hun persoonlijke politieke voorkeur, en ook naar die van hun ouders. De resultaten worden neergelegd in een datamatrix. In dit heel eenvoudige voorbeeld is dit een tabel met drie kolommen (de eigen voorkeur, die van de vader en die van de moeder) en vierhonderd rijen, voor elke respondent één. Het bekijken van zo'n datamatrix geeft ons niet onmiddellijk een antwoord op de gestelde vraag. Er moeten een paar berekeningen worden uitgevoerd voordat de vraag kan worden beantwoord.

Kader 1.1

Het onderzoeksrapport

De onderzoeksresultaten worden neergelegd in precieze uitspraken over de empirische werkelijkheid.² Maar geschreven rapporten zijn niet de enige manier om de resultaten van wetenschappelijk onderzoek te verspreiden. Andere manieren om kennis over te dragen zijn formules, formele en niet-formele modellen, instrumenten, essays, audio- en video-opnamen, zelfs films en tv-documentaires. Voorlopig houden we ons alleen bezig met op schrift gestelde wetenschappelijke beweringen. Tenslotte kun je de meeste andere vormen van rapportage uiteindelijk ook in normale taal weergeven, al zou je veel tekst nodig hebben om een film in woorden uit te drukken. Aan de andere kant,

² De empirie is 'de wereld om ons heen'. De empirische wetenschappen zijn wetenschappen waarvan de inhoud kan worden geverifieerd of weerlegd door observatie of experiment. Regelmatigheden die we vinden in de waarneming van de wereld om ons heen, worden 'empirische wetten' genoemd (denk aan Newtons wetten). De logica, de filosofie, de theologie en de wiskunde zijn niet-empirische wetenschappen. De meeste andere wetenschappen zijn (in elk geval voor een groot deel) empirisch.

een formule is veel preciezer dan het gebruik van woorden, maar een formule is zeker niet overal op toe te passen, en is bovendien voor veel mensen moeilijk te begrijpen.

Vanzelfsprekend dient een onderzoeksrapport om de resultaten van het onderzoek (de 'producten') weer te geven. En als de probleemstelling niet, of alleen maar gedeeltelijk, kan worden opgelost, moeten we uitleggen waarom dit zo is. In het rapport geven we ook een gedetailleerd overzicht van alle gevolgde stappen en gemaakte keuzen in het onderzoek (de 'procedures' van het onderzoek). Om te zorgen dat onderzoeksresultaten werkelijk gelezen worden, kan het verstandig zijn om tabellen en dergelijke in een bijlage op te nemen. Of je kunt een afzonderlijk, samenvattend rapport(je) schrijven, en het complete verhaal in een apart document opnemen.

Een rapport van praktijkgericht onderzoek eindigt meestal met aanbevelingen voor het beleid. Bovendien is het bij praktijkgericht onderzoek langzamerhand de gewoonte geworden om met brochures, handleidingen en beschrijvingen van *good practices* te werken. Ook worden websites, conferenties, speciale bijeenkomsten en workshops gebruikt om de verkregen inzichten te delen met mogelijke gebruikers van de onderzoeksresultaten.

Het schrijven van een rapport moet niet worden uitgesteld tot aan het einde van het hele onderzoeksproces. We bevelen zelfs aan om er vanaf het begin aan te werken. Ervaren onderzoekers starten hun project meestal met het schrijven van een *dummy*, dat wil zeggen een inhoudsopgave van de hoofdstukken die later nog moeten volgen. Zo snel mogelijk wordt elk hoofdstuk gevuld met voorlopige stukjes tekst. Op die manier kun je voorkomen dat na dataverzameling en -analyse nog een uitgebreide langdurige schrijffase nodig is. Dit betekent natuurlijk niet dat alles wat geschreven is daarmee ook definitief vaststaat. Voortdurende aanpassingen zijn eerder regel dan uitzondering. Maar het vroegtijdig schrijven brengt met zich mee dat je je voortdurend bewust bent van je eigen beslissingen en de veranderingen in je gedachten over het onderzoeksobject. Dan is het ook veel gemakkelijker om je te houden aan het principe van het 'verantwoording afleggen'³ dan wanneer je het schrijfwerk uitstelt tot wanneer alle andere fasen achter de rug zijn – en je veel vergeten bent van wat er gebeurd is en hoe je eerder over een en ander gedacht hebt.

3 Dit betekent dat een wetenschappelijk onderzoeker de gebruikte procedures en de gemaakte keuzen nauwgezet moet rapporteren, zodat andere onderzoekers op dezelfde wijze onderzoek kunnen doen om te controleren of zij tot dezelfde of mogelijk andere resultaten komen (zie hoofdstuk 8).

1.2 De gereedschapskist van onderzoekers

Voordat we in de volgende hoofdstukken uitvoerig ingaan op onderzoeksontwerpen en op vormen van dataverzameling hebben we een simpel gereedschapskistje nodig om de taal van wetenschappelijke onderzoekers te begrijpen. Die taal verschilt van de taal van bijvoorbeeld zorgverleners, schrijvers of journalisten. Eigenlijk is de taal van de wetenschap nogal arm. Terwijl het in een literair essay mooi wordt gevonden om niet in één zin dezelfde term twee- of driemaal te gebruiken, maar om juist naar verschillende bewoordingen te zoeken, hechten onderzoekers eraan om het aantal te gebruiken begrippen zo beperkt mogelijk te houden. Dat heeft te maken met het streven naar precisie:⁴ het gebruik van zo weinig mogelijk verschillende begrippen komt de duidelijkheid ten goede. Termen moeten worden gedefinieerd, zodat de betekenis voor iedereen op dezelfde manier vastligt, en bij synoniemen rijst soms de vraag of nu hetzelfde is bedoeld of net iets anders.

In wetenschappelijke uitspraken staan drie aspecten centraal. Het eerste aspect is het *domein*: de verzameling van elementen of *eenheden* waarop de uitspraak betrekking heeft. Het tweede aspect betreft de kenmerken of eigenschappen van de eenheden waarover in de uitspraak iets gezegd wordt (zie kader 1.2). We noemen die eigenschappen gewoonlijk de *variabelen* (zie kader 1.3). Het derde aspect is de score of waarde die elke eenheid op elke variabele heeft: de *waarde*.

Het domein van de eenheden

Het domein is de verzameling van eenheden waarover de onderzoeker conclusies wil trekken. Statistici noemen het domein meestal de populatie of het universum; met al deze termen wordt hetzelfde bedoeld. In kader 1.2 zijn wat voorbeelden gegeven. Eigenlijk hoort er ook een tijdsbepaling bij, omdat de verzamelingen op het ene tijdstip geheel of gedeeltelijk anders kunnen zijn dan op een ander tijdstip. Korthedshalve zijn de meeste tijdsbepalingen hier weggelaten.

Een onderzoeker zal gewoonlijk niet *alle* elementen van het domein waarin hij geïnteresseerd is onderzoeken, omdat dit op veel praktische problemen stuit. Uit het domein wordt een *steekproef* getrokken. Een steekproef is niets anders dan een deelverzameling. Onder het begrip steekproef vallen zowel hele specifieke, geselecteerde deelverzamelingen (bijvoorbeeld van alle bedrijven alleen die uit de metaalsector) als zogenoemde *aselecte steekproeven* (waarbij een blinde kansprocedure wordt gevolgd om een steekproef uit de populatie te trekken).

4 Zo precies mogelijk zijn is een ander criterium voor de kwaliteit van wetenschappelijk onderzoek (zie hoofdstuk 8).


Voorbeeld

Enkele voorbeelden van een domein

- Alle inwoners van Ierland van achttien jaar en ouder op 1 oktober 2009
- Alle Nederlanders die een 'bijna-doodervaring' gehad hebben
- Alle huishoudens binnen de gemeente Arnhem
- Alle alleenwonende bejaarden in de gemeente Gent
- De eerstejaarsstudenten van de Hogeschool van Amsterdam
- De productiebedrijven met meer dan honderd werknemers in de provincie Antwerpen
- De landen die bij de Verenigde Naties zijn aangesloten
- Conversaties van studenten tijdens koffiepauzes
- Vijf Vlaamse dagbladen
- De meldingen gedurende één week in het logboek van een politiebureau

Kader 1.2

Variabelen en scores

Een eigenschap van eenheden wordt een variabele genoemd wanneer de eenheden binnen het domein van elkaar verschillen in waarde. Een variabele heeft ten minste twee waarden. In plaats van het woord 'waarde' worden ook wel andere aanduidingen gebruikt, zoals categorieën of condities.

Een groep mensen bestaat gewoonlijk uit mannen en vrouwen. Daarom noemen we in zo'n gemengde groep de eigenschap sekse een variabele. De waarden van de variabele zijn 'man' en 'vrouw'. Binnen een groep van alleen vrouwen is sekse geen variabele, omdat het dan geen eigenschap is op grond waarvan je eenheden van elkaar kunt onderscheiden.

In het onderzoek moet de onderzoeker ervoor zorgen dat een waarde van een variabele wordt toegekend aan *elke* eenheid (meten moet 'uitputtend' zijn), en dat niet meer dan *één* waarde wordt toegekend aan elke eenheid (meten moet 'exclusief' zijn).⁵

Je kunt het idee van 'variabele' ook toepassen op een eigenschap van één en dezelfde eenheid als die *varieert in de tijd*. Als je gedurende een langere periode steeds opnieuw van iemand de gezondheid vaststelt, zal die gezondheid wel een variabele zijn. Eigenschappen zoals geboortjaar, sekse en dergelijke variëren niet bij één persoon, maar meningen en houdingen wél.

5 De uitdrukking 'meten' betekent in het algemeen het toekennen van getallen aan eenheden volgens bepaalde regels. Het is hier gebruikt in een heel algemene zin, of het er nu om gaat getallen aan mensen toe te kennen voor hun haarkleur of geboortestad, of voor kwantificeerbare eigenschappen, zoals lengte of gewicht.

Je kunt variabelen onderscheiden naar het type eenheden waar ze bij horen, zoals mensen, organisaties, landen, stakingen, auto's, conversaties, arrestaties. In kader 1.3 staat een aantal voorbeelden (de eenheden zijn schuin gedrukt).

Voorbeeld

Variabelen als eigenschappen van ...



- *mensen*: lengte, gewicht, mening over een nieuwe pensioenregeling, opleiding, rookgedrag, vrijetijdsgedrag
- *bedrijven*: aantal werknemers, de bedrijfstak, de productiviteit, gemiddeld aantal ziektedagen
- *landen*: aantal inwoners, percentage van het BNP dat aan onderwijs wordt besteed, natuurlijke hulpbronnen
- *scholen*: aantal leraren, grootte van de gemeente waarin de school ligt, percentage allochtone leerlingen, percentage oud-leerlingen dat een universitaire opleiding volgt
- *ziekenhuizen*: aantal bedden, aantal specialismen, lokale of regionale functie
- *leerlingen*: sekse, schoolresultaat, sociaaleconomische achtergrond, etnische herkomst
- *'beurten' in de schoolklas*: of de leraar vragen stelt aan iedereen of aan bepaalde leerlingen; of de leerling(en) het goede antwoord we(e)ten; of de leraar goedkeuring of afkeuring laat blijken
- *kranten*: oplage, politieke kleur, percentage sportpagina's
- *auto's*: aantal deuren, maximumsnelheid, motorvermogen, type remmen
- *conversaties*: aantal deelnemers, onderwerp(en), lengte, evenwicht van de bijdragen van de verschillende deelnemers
- *aantekeningen in het logboek van een politiebureau*: tijd, type, wat ermee gedaan is

Kader 1.3

De datamatrix

Eenheden, variabelen en waarden worden in het onderzoek bij elkaar gebracht in een 'datamatrix' (zie kader 1.4 en 1.5). De datamatrix dient om netjes en systematisch de verzamelde gegevens vast te leggen. De datamatrix vormt daarmee het eindpunt van de dataverzameling, en tegelijk het beginpunt voor de data-analyse (in principe; niet altijd zijn dataverzameling en data-analyse zo precies gescheiden).

Stel dat ons onderzoek gaat over de vraag hoe jonge stellen die geen kinderen kunnen krijgen, maar dat wel graag willen, hiermee omgaan; welke stappen zij ondernemen, en de volgorde daarvan, om toch een kind te kunnen krijgen. Meer in het bijzonder gaat het om de vraag of zij wel eens over adoptie hebben gedacht. We verzamelen data via uitgebreide interviews met beide partners. Bij het begin van het interview vragen we man en vrouw spontaan te vertellen over hun persoonlijke gedachten en ervaringen. We zouden dan een paar variabelen kunnen maken waarop we de mensen die geïnterviewd worden van elkaar

onderscheiden, en bij elk jong stel (de rijen in de datamatrix) en voor elke variabele (de kolommen in de datamatrix) een opmerking kunnen schrijven:

- Wie treedt bij het begin van het interview als gesprekspartner op, man of vrouw?
- Is er overeenstemming tussen de partners?
- Probeerde men een kind te krijgen anders dan via de normale weg (bij deze variabele kun je aan een simpel ja of nee denken)?
- Welke methoden worden spontaan genoemd (dus zonder dat de interviewer er een paar opsomt)?
- Als adoptie wordt genoemd, werden daar spontaan negatieve dan wel positieve kwalificaties aan toegevoegd? Welke woorden werden gebruikt?

We komen dan misschien tot het volgende ‘stukje van’ een datamatrix.


Voorbeeld

Een datamatrix in woorden

Eerste gesprekspartner	Mate van overeenstemming	Pogingen	Methoden	Kwalificatie
paar 01 vrouw	we hebben er erg veel over gepraat, maar na een jaar of twee zijn we daarmee maar opgehouden. Mijn man wi nog steeds kinderen	ja	medisch onderzoek liet geen afwijkingen zien	het was niet gemakkelijk voor me om naar een specialist te gaan zien
paar 02 vrouw	ik ben wanhopig. Mijn man maakt er zich niet meer druk om. We hadden afgesproken om ons uiterste best te doen, maar nu zijn we het niet meer eens	ja	we hebben ons beiden medisch laten keuren; 3x ivf	ik kreeg de pest aan al die ‘behandelingen’. Op een gegeven moment heb je zoiets van ‘het doel heiligt niet meer de middelen’
paar 03 man	vier jaar lang hebben we ons best gedaan. Eindeloos hebben we erover gepraat, ten slotte besloten we een kind te adopteren	ja	medisch onderzoek; 5x ivf; adoptie	we wilden gewoon heel graag een kind. Al die moeite daarvoor was niet erg. We zijn nu bezig een Vietnamees kind te adopteren, hoewel er risico’s (gezondheid) aan verbonden zijn