

Uitwerkingen opdrachten *Basisboek sociaal onderzoek* (6e druk)

Hoofdstuk 1

- 1.1** Het domein: collegestudenten; hier niet gespecificeerd naar tijd en plaats, wat eigenlijk wel moet in een rapport. De variabelen: moedertaal (categorieën: moedertaal Engels/moedertaal niet-Engels); geslacht; eerste- of oudere-jaars; aantal verbale bijdragen tijdens de les.
- 1.2** Mannelijke college studenten. Geslacht is hier geen variabele. Je kunt er natuurlijk wel een variabele van maken door vrouwelijke studenten dezelfde taak te laten uitvoeren. De variabelen zijn: (a) kwaliteit van het essay (goed of niet goed geschreven); (b) de waardering voor het uiterlijk van de schrijver (drie categorieën, bijvoorbeeld 'onaantrekkelijk'; 'houdt het midden' en 'aantrekkelijk'); (c) de waardering van de capaciteiten van de schrijver (verschillende schalen, hier niet nader gespecificeerd); (d) de waardering door de mannelijke studenten van de kwaliteit van het essay (bijvoorbeeld op een tienpuntsschaal, zoals schoolcijfers).
Om de beoordeling van het uiterlijk van de schrijfster in verband te brengen met de kwaliteit van het essay kun je een tabel maken met variabele a in de kolommen, en variabele b in de rijen. Voor de goed geschreven essays zouden de berekende gemiddelden voor elke cel, als ze ingevuld worden in de rijen, voor de goedgeschreven essays van boven naar beneden kunnen oplopen van bijvoorbeeld 8, via 8.5 naar 9, terwijl voor de slecht geschreven essays de gemiddelden 4, 5 en 6 zouden kunnen zijn. Ga na dat deze getallen in overeenstemming zijn met de gegevens. Je kunt ook een ingewikkelder tabel maken, waarin alle vier variabelen betrokken zijn. Probeer dat te doen, waarbij je aan alle beweringen in de vraag recht doet. Als je dat kunt, ben je al een heel eind gevorderd in het 'denken in variabelen'.
- 1.3**
- Diverse variabelen in de kolommen, en diverse eenheden in de rijen;
 - één variabele als kolom, en diverse eenheden in de rijen;
 - diverse variabelen in de kolommen, en één rij;
 - verschillende eenheden in de kolommen, en dezelfde verzameling eenheden in de rijen;
 - één variabele, verschillende tijdstippen van meting in de kolommen; eenheden in de rijen;
 - verschillende variabelen, elk op drie tijdstippen gemeten, in de kolommen, en diverse eenheden in de rijen.
- 1.4** Lengte en gewicht houden verband met elkaar; in het algemeen geldt: hoe langer, des te zwaarder. Waarschijnlijk laat de leeftijd weinig variatie zien, en zal deze ook nauwelijks verband houden met de andere variabelen.

Hoofdstuk 2

2.1 Het hangt er natuurlijk van af welk soort verschil wordt gevonden. Als in stedelijke gebieden het absentisme groter is dan in niet-stedelijke gebieden, denk je misschien aan de grotere gelegenheid in steden voor allerlei bijbaantjes; aan meer winkels en meer uitgaansgelegenheden; coffeeshops in de buurt van de scholen, enzovoort. Dit is een meer voor de hand liggende verklaring dan een verklaring die uitgaat van hogere echtscheidingscijfers, minder sociale controle, afwezigheid van normen en waarden in stedelijke gebieden. Maar als het absentisme groter is in de niet-stedelijke gebieden, kun je denken aan kinderen die helpen op de boerderij; aan grotere afstanden, aan minder gemotiveerde kinderen, enzovoort. Achteraf een verklaring bedenken is nooit moeilijk. Dit maakt duidelijk waarom we een toetsende benadering nodig hebben, waarbij hypothesen voorafgaande aan het onderzoek worden geformuleerd, en waarbij een streng gereguleerde onderzoeksprocedure wordt doorlopen. De resultaten zijn dan gemakkelijker als geldig aan te merken dan wanneer je blijft exploreren.

2.2

- a. Na de waarneming beginnen we met te zoeken naar de aard en de oorzaak van de vlek; je krijgt dan vermoedens. Ieder vermoeden leidt weer tot een specifieke behandeling (die afgeleid kan worden uit een chemische of een mechanische theorie). Uiteindelijk toetsen we het idee dat het meest waarschijnlijk lijkt. Als we geen succes hebben, is het volgende idee aan de beurt om getoetst te worden.
- b. Eerst moet een beschrijvingsprobleem worden opgelost: is het waar dat de prestaties van het team een stuk slechter zijn dan die van andere teams? Is dat al lange tijd zo, of kwamen die slechte prestaties na een bepaalde verandering (trainer, tijdstip, deelnemers)? Het antwoord dat we in eerste instantie op deze vragen geven, brengt ons op het spoor van één of meer oorzaken, bijvoorbeeld interne conflicten, tekort-schietende begeleiding. Ook nu weer, na het kiezen van de meest waarschijnlijke oorzaak maken we 'een model' van de factoren die op hun beurt die oorzaak beïnvloeden. Dat kan een heel praktische oplossing zijn, zoals het kiezen van een ander trainingsschema. Het kan ook een meer 'groepsdynamische aanpak' zijn om conflicten op te lossen. De eindfase bestaat, zoals altijd, uit het nemen van de proef op de som.
- c. Ook nu weer kom je pas na een precieze beschrijving van de stand van zaken op het spoor van de identiteit van niet bereikte doelgroepen, en wellicht van mogelijke verklaringen, bijvoorbeeld gebrekkige nazorg. In de eindfase kan de veronderstelling dat een betere nazorg tot een groter succes leidt, worden getoetst.
- d. Om te beginnen moet je inzien dat er gaten zijn in je kennis, en dat je je moet inspannen om deze op te vullen. Om een strategie te ontwerpen- zoek je naar de oorzaken van deze lacunes, zoals bijvoorbeeld onvoldoende kennis verworven op de middelbare school of ontbrekende motivatie tot dusver om eraan te werken. Daarna stel je jezelf de vraag: wat is de beste manier om de situatie te verbeteren?

2.3 Beginfase, beschrijvingsvragen: hoeveel vuilnis wordt er per straat en per dag van de week op straat gegooid? Welke beleidsmaatregelen zijn er tot nu toe gebruikt, en hoeveel succes hebben die?
Beginfase, verklaringsvraag: hoe komt het dat er verschillen tussen straten zijn?

Welke factoren beïnvloeden de hoeveelheid rommel die op straat gegooid wordt, en hoe verhouden die zich tot de genomen maatregelen?

Beginfase, ontwerpvraag: wat is de beste mix van beleidsmaatregelen om het probleem op te lossen?

Als we nu eens aannemen dat die 'beste mix' wordt uitgevoerd (waaronder bijvoorbeeld het op strategische plaatsen zetten van vuilcontainers), komen we op een paar vragen voor de volgende onderzoeksfase.

Tweede fase, beschrijvingsvraag:

Hoeveel vuilnis wordt er op straat gegooid een week na de invoering van de nieuwe maatregel; een maand later, en drie maanden later?

Tweede fase, verklaringsvraag: waarom heeft het nieuwe beleid aanvankelijk wel succes, maar is het effect na enige tijd verdwenen?

Hoe komt het dat de aanvankelijke doeleinden in buurt X bereikt worden, maar niet in buurt Y?

Tweede fase, ontwerpvraag:

Hoe passen we onze beleidsingrepen zodanig aan dat deze wel blijvend succes hebben?

Hoe moeten we ons beleid aanpassen zodanig dat het ook in buurt Y succes heeft?

2.4 Beschrijvingsvragen: Wat is de wetgeving met betrekking tot dit onderwerp in deze landen?

Zijn er beleidscampagnes geweest op dit terrein; bijvoorbeeld informatiecampagnes gericht op de leerlingen van middelbare scholen? Wat is de inhoud, en wat de reikwijdte van deze campagnes geweest? Zijn er kwantitatieve aanwijzingen voor raciaal geweld; wat blijkt daaruit?

Verklaringsvragen: Waarom heeft land X een veel strengere wetgeving op het gebied van raciaal geweld dan land Y, terwijl de omvang van het geweld nagenoeg hetzelfde is? Waarom werkt een bepaalde beleidsmaatregel in het ene land wel, in het andere land niet?

Ontwerp vragen: Wat kan er gedaan worden door de nationale overheid, wat door de privésector in de verschillende landen? Wat kan verwacht worden van een strengere wetgeving? Deze laatste vraag kan ook worden gezien als een verklaringsvraag: waarom zijn strengere wetten effectiever; wat zijn hiervan de oorzaken?

2.5 De eerste twee vragen zijn beschrijvend. Het kan zijn dat de derde vraag een ontwerp probleem betreft, maar de woordkeuze is niet goed. Het is beter om te vragen: 'Welk protocol komt tegemoet aan de eisen van ...?' Vaak worden vragen als volgt geformuleerd: 'Hoe ziet een ... maatregel eruit?' Het valt meteen op dat er wordt verondersteld dat het antwoord op de beschrijvingsvragen niet afdoende zal zijn, en dat er veel mis is met de bestaande situatie. Het onderzoeksontwerp moet rekening houden met zowel beschrijvings- als verklaringsvragen. Een verklaringsvraag ontbreekt hier, en het idee dat de vragen logisch uit elkaar moeten volgen, is afwezig. Een subsidiërende instantie zou hier een vraagteken zetten: als je al weet dat het allemaal anders moet, waarom moet je dan nog een beschrijvend onderzoek doen? De vierde vraag, die met betrekking tot de 'werkbaarheid', is een onderdeel van de derde. De werkbaarheid is een van de eisen die aan een protocol moeten worden gesteld. De laatste vraag refereert aan de evaluatie van een nieuw protocol. Met andere woorden, het voorlopige eind van het ontwerpproces.

2.6

1. *
2. *
3. Beschrijvingsvragen: a, c, d, e, h, i, k, m, o, p.
Verklaringsvragen: g, l, n.
Ontwerp vragen: b, f, j, q.
4. *

Hoofdstuk 3

- 3.1** a. overheidsstatistieken en historische bronnen;
b. overheidsstatistieken en publicaties in de pers;
c. survey (waarom zullen de officiële verkiezingsuitslagen niet voldoende zijn?);
d. en e.: laboratoriumexperimenten;
f. survey.
- 3.2** Selecteer eerst een tijdsperiode, bijvoorbeeld vijf jaar. Kies enkele dagbladen. Beslis of je alleen naar de voorpagina gaat kijken, of naar alle pagina's. Beslis over hoe je zelfdoding definieert. Neem je bijvoorbeeld ook de combinatie van moord en zelfmoord van de moordenaar mee? Wat doe je met politieke zelfmoorden? Hongerstakingen waar de dood op volgt? Neem een beslissing over rouwadvertenties: moet je die meenemen wanneer zelfmoord (min of meer) duidelijk als doodsoorzaak is aangegeven? Hoe moet je, in de praktijk, al die oude kranten vinden? Ten slotte: hoe veel tijd en geld heb je om codeurs te trainen?
Ten tweede heb je betrouwbare statistieken over de gekozen periode nodig. Waar vind je die? Dit moet een van de eerste onderwerpen van onderzoek zijn, want als je geen gegevens over zelfdodingen in die periode (liefst per dag, maar per week zou eventueel genoeg gedetailleerd kunnen zijn) kunt vinden, is dit type onderzoek niet mogelijk. Er zijn voorts allerlei moeilijkheden te verwachten rondom de betrouwbaarheid en vergelijkbaarheid van politiearchieven, waar meestal de statistische gegevens op berusten.
- 3.3** Hoeveel tijd besteden mijn collega's en ik aan dit onderwerp? Hoe is die tijd over verschillende categorieën verdeeld, zoals: a. bijwonen van colleges; b. lezen van boeken enz.; c. oefeningen maken; d. discussie met collega's; e. uiteindelijke voorbereiding op het tentamen. Hoe is deze totale tijds-investering verdeeld over de hele periode van studie?
De totale werktijd is misschien overwegend besteed aan individuele studie, of juist in het samenwerken met anderen, of weegt dit ongeveer tegen elkaar op? Wat is de frequentieverdeling over deze categorieën?
Het ontwerp van het onderzoeksproject: begin zo mogelijk bij het begin van de periode die bestudeerd wordt. Een dagboek is essentieel. Laat iedereen bij wijze van spreken een wekker zetten, zodat ze niet vergeten een dagboek bij te werken. Anderzijds kan dit misschien de studiegewoonten beïnvloeden. Als een dagboek niet mogelijk is, ga dan herhaaldelijk je collega's interviewen, vanaf het begin. In het algemeen is achteraf interviewen nogal onbetrouwbaar, tenzij je heel nauwkeurig bent en heel

gedetailleerde vragen stelt over de studieactiviteiten.

3.4 Het trekken van een toevalssteekproef uit het bevolkingsregister van enkele gemeenten (waarbij je gemakkelijk op leeftijd kunt selecteren) zou tot een goede steekproef kunnen leiden. Maar de privacywetgeving en de kosten zouden dit waarschijnlijk verhinderen. Meestal kan een min of meer random sample worden verkregen via de telefoon met 'random digit dialing' waarbij personen van 55 jaar en ouder kunnen worden geselecteerd. Een survey is de enig mogelijke strategie om voorkeuren vast te stellen. Of het een telefonische survey moet zijn, een interviewcampagne, een postenquête of een mailenquête is een punt voor verder overleg. Omdat het onderwerp niet zo bijzonder moeilijk is en de respondenten waarschijnlijk wel geïnteresseerd zullen zijn, hoef je hier niet zo bang te zijn voor een hoge non-respons. Als het om een plaatselijk onderzoek gaat, bijvoorbeeld wanneer de gemeente graag gegevens hierover wil hebben van de eigen bevolking in verband met de planning van woonwijken en verzorgingstehuizen, moet je eigenlijk de leeftijdsgrens wat lager leggen, immers, 60-jarigen van nu zullen al wel 70 zijn als de eerste gebouwen klaar zijn.

3.5

- a. Er is blijkbaar een survey gedaan. Een gevonden verband tussen twee variabelen wordt veel te snel als causaal geïnterpreteerd. De leeftijd zou een gemeenschappelijke oorzaak kunnen zijn.
- b. Er wordt niets beweerd over de mate van geluk van echtparen met kinderen; de tevredenheid heeft misschien niets met het al dan niet hebben van kinderen te maken.
- c. Er wordt een conclusie getrokken die gebaseerd zou moeten zijn op longitudinaal onderzoek (dezelfde personen worden daarbij over een aantal jaren gevolgd), terwijl de data verwijzen naar een transversale vergelijking (verschillende groepen op hetzelfde moment gemeten). Het gaat hier waarschijnlijk om verschillende *cohorten*: jongere cohorten gaan minder naar de kerk dan oudere.
- d. Gesuggereerd wordt dat vroeger het geboortecijfer hoger was, maar er zijn geen data om dat te bevestigen of te ontkennen.
- e. Op zijn best komen de data uit longitudinaal onderzoek, maar er worden alleen mensen met kinderen ondervraagd. Het is ook mogelijk dat op één moment mensen zijn gevraagd naar hun geluk nu en toen; dat leidt tot zwaar vertekende antwoorden. Wil de conclusie overigens gefundeerd zijn, dan kan dat alleen met een vergelijkingsgroep van mensen die nooit kinderen hebben gehad.

3.6 Het is een veldexperiment, data worden uitgelokt. Maar de situatie is erg natuurlijk, en we kunnen veronderstellen dat het gedrag van de verkoper en ander personeel niet beïnvloed wordt door de identiteit van de koper omdat deze persoon in niets afwijkt van andere consumenten. Dus moet de 'mystery shopper' oppassen dat hij of zij zich bij het invullen van een vragenformulier niet te sterk door de eerste indruk laat leiden: er treedt dan 'contaminatie' op.

3.7

- a. Zou zowel betrouwbaar als valide kunnen zijn; overtuig je ervan dat de respondenten het maken van de quiz als een serieuze bezigheid opvatten.

- b. Het kan veel tijd kosten voordat genoeg data zijn verzameld.
- c. Mensen kunnen sterk in de politiek geïnteresseerd zijn en er veel van af weten, maar discussiëren hier niet over met medestudenten, of ze kunnen er ook weinig van af weten, maar er toch veel over praten; daarom is deze benadering misschien niet valide.
- d. Op zichzelf is het deelnemen aan politieke bijeenkomsten waarschijnlijk wel een goede indicator van kennis, maar het komt zo weinig voor dat je met de antwoorden op de vraag nauwelijks mensen van elkaar kunt onderscheiden.

3.8

- a. Aardig gevonden worden is niet hetzelfde als populariteit.
- b. Deze benadering kan in principe gebruikt worden. De individuele rangordes kunnen samengevat worden in een groepsrangorde; dit heeft echter alleen zin als de individuele rangordes niet te veel van elkaar verschillen. Een praktische moeilijkheid zou kunnen zijn dat er te veel 'weet niet'-antwoorden zijn, of te veel mensen die eenzelfde rangpositie krijgen.
- c. Deze benadering kan ook worden gebruikt; de score die elke persoon krijgt is de som van alle keren dat hij of zij als vriend wordt genoemd. Het is misschien zelfs wel een betere maat dan b, omdat deze meer persoonlijk, meer direct is, en daarmee het 'algemene opinie'-achtige karakter van b omzeilt.

3.9 Hoewel het woord 'analyse' in de naam wordt genoemd, omvat deze benadering zowel ontwerp, dataverzameling, steekproeftrekking, kwaliteitscontrole e.d. als alle andere strategieën, en daarom wordt inhoudsanalyse beschouwd als een zelfstandige strategie.

Hoofdstuk 4

4.1 De staking: er wordt slechts één databron gebruikt, hoewel er wel van allerlei soorten documenten sprake kan zijn. Voor het overige is het een goed voorbeeld van een case study. Het 'follow-up'-onderzoek van een paar jongens: bijna alle kenmerken van een case study zijn van toepassing (behalve de bespreking achteraf van de resultaten). De opgenomen patiënten: de gestelde vraag heeft niet het open karakter met de nadruk op sociale relaties die gewoonlijk bij een case study aanwezig is, en 'opgenomen zijn' is niet een erg natuurlijke situatie. Het is een veldexperiment met een paar eigenschappen van een case study (slechts een paar mensen, longitudinaal).

4.2 *

4.3 Zoals altijd bij een veldexperiment wordt het gedrag van mensen uitgelokt door degene die het experiment uitvoert (de student die een vuilnisbak repareert enz.). In het algemeen zijn onderzoekers geneigd om de invloed van hun eigen gedrag op dat van anderen te onderschatten. Mogelijk komt een situatie zoals geschetst over als onnatuurlijk, hetgeen negatieve reacties kan uitlokken. Verondersteld wordt dat voorbijgangers de situatie niet herkennen als een onderzoekssituatie. Verder is het een relevante vraag of de gekozen situatie relevant is voor het concept 'verantwoordelijkheid dragen voor zaken van algemeen belang'. Maar het is moeilijk om een betere manier van onderzoek te bedenken om deze bereidheid te meten. Het

gebruik van vignetten (zie hoofdstuk 8) komt er het dichtstbij en kost minder tijd, maar is waarschijnlijk minder valide. Babbies onderzoek is een herhaald veld-experiment. Als het beschouwd wordt als een case study, kun je het het best typeren als een multi-pele case study, met elke situatie als een aparte case. Om je conclusies steviger te maken is het aan te raden om het gedrag van mensen in verschillende situaties te onderzoeken.

4.5 *

Hoofdstuk 5

- 5.1** Vergelijk de definitie van het survey. De gestelde vragen zullen zeker niet identiek zijn; in een diepte-interview worden ze aangepast aan de gesprekspartner. Maar de behandelde onderwerpen zullen grotendeels hetzelfde zijn, en je kunt de van persoon tot persoon uiteenlopende aanpassingen zien als een techniek om de validiteit van de antwoorden te verhogen. Slechts één databron wordt gebruikt, en iemand interviewen is iets anders dan iemand in een natuurlijke situatie te bestuderen. Of je nu dit onderzoek als een survey of als een serie case studies beschouwt, is op zich niet zo belangrijk; het gaat erom dat je de verschillende relevante aspecten kunt benoemen.
- 5.2** Ongeveer 2500. Gebruik de formule, of de tabel. NB: $(.02)^2 = .0004!$
- 5.3** De tijd van de dag is belangrijk. Van 9-12 uur en van 14-17 uur zijn werklozen en ouderen oververtegenwoordigd, en ook nog steeds: vrouwen. En hoe later op de avond, des te minder ouderen en gehandicapten. Quotasampling kan op alle tijden van de dag worden gedaan, met quota gebaseerd op geslacht, ruw geschatte leeftijds categorieën en werkend/werkloos (dat laatste kan gevraagd worden).
- 5.4** Actieve, gemotiveerde studenten (die regelmatig op de universiteit komen), maar ook enkele zeer ontevreden studenten, bijvoorbeeld degenen die recentelijk een in hun ogen onterecht gedrag van de faculteit jegens hen hebben ervaren.
- 5.5** Ja, er zal zeker een bepaalde selectie zijn. Mensen die het gevoel hebben dat zij in hun loopbaan gefaald hebben; werklozen; zij die geheel buiten het terrein van hun opleiding werkzaam zijn, zullen minder gauw geneigd zijn om aan zo'n onderzoek deel te nemen, en sommigen van hen zullen ook erg moeilijk bereikbaar zijn. Tegelijkertijd zullen er een paar zijn die een flitsende carrière in het vak gemaakt hebben, die hoge posities bereikt hebben, maar die dus ook een zeer drukbezet leven leiden. Zelfs als ze bereikt kunnen worden en ze even tijd hebben, liggen het studentenbestaan en de opleiding ver achter hen, en hebben ze de belangstelling daarvoor verloren.
- 5.6** Veronderstel dat de gekozen gemeente 932 telefoonnummers heeft; deze kunnen onderverdeeld worden in 40 intervallen, elk van ongeveer 23 nummers. Begin met het 'met de ogen dicht' kiezen van een nummer onder de 23. Selecteer daarna de andere nummers door steeds tot 23 te tellen tot je bij de volgende bent, of meet met een liniaal de gemiddelde afstand tussen 23 nummers, en gebruik de liniaal om de volgende nummers te vinden.

5.7 *

5.8 *

Hoofdstuk 6

6.1

- a. De eerste zin is overbodig en bovendien suggestief. 'Heb je die en die film gezien?' is genoeg.
- b. Suggestief. 'Wat vind je van de duur van de laatste kabinetsformatie' is genoeg.
- c. Een jaar is veel te lang om te kunnen overzien, en het is onduidelijk wat wel en wat niet onder boeken valt (studieboeken ook?).
- d. Wat moeten mensen zeggen die óf voor, óf tegen elke legalisering zijn? De vragen moeten worden gesplitst.
- e. Niet naar percentages, maar naar bedragen vragen (nog liever een lijst van bedrag-'klassen' voorleggen).
- f. Dubbele negatie, en de verwoording is veel te ingewikkeld.
- g. 'Niet van toepassing' toevoegen. Of de vraag alleen stellen na een voorafgaande filtervraag: ga je wel eens naar de schouwburg?
- h. uitsluitend = altijd!
- i. Zo'n vraagstelling altijd beperken tot vrouwen die in verwachting zijn of geweest zijn.
- j. Bus wel, trein niet?
- k. Wie is daartegen?
- l. Te lange periode om te kunnen overzien voor de meeste mensen.
- m. Het 'of juist niet' gezien de antwoordmogelijkheden weglaten.
- n. Wie niet? Het antwoord is veel te vrijblijvend; bij zoiets altijd de kosten of nadelen van het gewenste beleid in stelling brengen.
- o. 'Of minder' weglaten. Dan wel antwoordmogelijkheden: meer/minder.
- p. Wat moet iemand zeggen die nergens liever zou willen wonen?
- q. Grapje; dit is een dichotome vraag.
- r. Je moet ze dan eigenlijk allemaal noemen, omdat sommigen dan 'nee' zeggen omdat ze niet *de Volkskrant* of *het Parool* lezen, maar een andere krant.

Hoofdstuk 7

- 7.1 Toevalstoewijzing: laboratoriumexperiment: ja. Veldexperiment: ja (meestal). Natuurlijk experiment: nee, er wordt gewerkt met bestaande groepen.
Introductie en manipulatie van de experimentele variabele: laboratorium-experiment: ja. Veldexperiment: ja. Natuurlijk experiment: nee. Quasi--experiment: ja.
Conditie onder controle: laboratoriumexperiment: ja. Veldexperiment: gedeeltelijk.
Natuurlijk experiment: nee. Quasi-experiment: gedeeltelijk.
- 7.2 Het zal moeilijk zijn om vergelijkbare scholen te vinden (een toevalstoewijzing aan verschillende experimentele condities is misschien mogelijk, maar of de scholen vergelijkbaar zijn op relevante variabelen blijft vaak een punt van discussie). Verder kunnen er gedurende een zo lange periode veel factoren in de context veranderen, en,

ten slotte, scholen kunnen gedwongen worden hun beleid aan te passen gedurende de periode die bestudeerd wordt (door bijvoorbeeld een nieuwe wetgeving).

- 7.3** Bijvoorbeeld: in de kolommen een hoge, een middelmatige en een lage dosis van het medicijn. In de rijen: fysieke training zwaar, licht en helemaal niet.
- Cellen: eerste kolom 42.4, 36.7 en 31.5. Tweede kolom: 40.0, 34.7 en 29.2. Derde kolom: 35.7, 29.9 en 24.6. De rijvariabele heeft het sterkste effect.
 - De eerste kolom blijft hetzelfde, de tweede is (bijna) gelijk aan de eerste.
 - De eerste rij is gelijk aan de tweede rij, hetgeen betekent dat lichamelijke inspanning er niet toe doet.
 - Verwissel de cellen van de eerste en de derde kolom (of van de eerste en de derde rij).
 - In principe hoef je slechts één cel te veranderen, bijvoorbeeld de cel rechtsonder.
 - Een effect van de X-variabele wordt voorgesteld door een niet-horizontale lijn. Als de andere onafhankelijke variabele ook een effect heeft, zijn er drie niet-horizontale lijnen. Als deze lijnen parallel lopen, is er geen interactie. Niet-parallelle lijnen betekent de aanwezigheid van statistische interactie.

Hoofdstuk 8

- 8.1** Algemene aanwijzingen: geef de periode of de tijdstippen van meting precies aan; geef precies aan wat het domein is (wat betekent 'lagere sociaal-economische groepen', 'lagere betaalde baan' en 'exacte vakken')?

8.2

- 'Comfortabel' is niet precies genoeg, en het label hangt af van de persoonlijke 'bias' van degene die dit beweert.
- 'Van nature voorbestemd' kan niet weerlegd worden.
- Als niemand het spook kan vinden, kan men altijd volhouden dat het zich verdekt heeft opgesteld (of dat, nog beter, alleen de ware gelovigen het kunnen zien).
- Is gedeelte van een definitie.
- Wanneer? Zelfs als het wachten eindeloos lang duurt, kan de spreker, omdat hij nu eenmaal een optimistisch karakter heeft, volhouden dat de participatie zal toenemen; het is daarom niet weerlegbaar.
- Het woord 'betreuren' drukt een persoonlijk standpunt uit; zonder dat is de bewering weerlegbaar.
- Refereert aan een norm, die dus niet weerlegbaar is.
- Dit is een persoonlijk gezichtspunt waarmee niemand hoeft in te stemmen. Niet weerlegbaar.
- Niet precies genoeg.
- Dit is een weerlegbare bewering, althans voor een bepaalde verzameling landen of gemeenten waarvan gegevens over de politie en de criminaliteit bekend zijn gedurende een bepaalde periode. De manier van formuleren vertoont enig cynisme, omdat de achtergrond van de bewering waarschijnlijk is dat als de criminaliteit toeneemt, er wat meer politie op straat komt, maar dat dit laatste altijd achterloopt bij de groei van de criminaliteit. Op zich is de bewering niet onjuist als de statistiek een correlatie tussen die twee variabelen laat zien.

8.3 'Ben je tevreden met je studie?'

'Zou je anderen aanraden om deze studie te kiezen?'

'Voel je je in het algemeen plezierig als je met je studie bezig bent?'

'Als je met vrienden over je studie praat, ben je dan in het algemeen positief over je studie?'

'Als je 's morgens opstaat, wat zijn dan je eerste gedachten over je studie?'

'Vind je het leuk om over je studie te lezen?'

Al deze beweringen zijn positief. In een goede schaal worden positieve en negatieve items afgewisseld. Een heel andere benadering zou zijn om verschillende aspecten van de studie af te lopen (studielast boeken; colleges; practica; beroepsperspectieven) en bij elk daarvan enkele vragen te stellen.

8.4 Het eerste item lijkt precies gericht te zijn op wat de bedoeling is. Het item 'met de trein naar het werk gaan' is alleen bruikbaar voor mensen die tussen trein en andere middelen van vervoer kunnen kiezen (en die niet werken op loop- of fietsafstand van hun woonplaats).

'Praten over koetjes en kalfjes': dit is niet een vergelijkende bewering, maar beperkt tot de trein. Mensen die het met deze bewering eens zijn, zijn misschien ook positief over de auto. Voor 'rare mensen' geldt hetzelfde. Je kunt wel veel vragen bedenken over positieve en over negatieve motieven om met het ene, dan wel met het andere middel van vervoer te gaan, maar het probleem is om ze te combineren en om tot een uiteindelijke vergelijking te komen. Dit vraagt een erg verfijnde techniek van vragen formuleren. Daarom is het gebruik van vergelijkende items (zoals het eerste) te prefereren.

8.5

- a. Een praktisch probleem kan zijn dat niemand uit zichzelf geneigd is om afval te verzamelen, maar een hint kan genoeg zijn. Je moet zo'n experiment vaak herhalen om het effect van groepsdruk uit te schakelen: iemand die iedereen om hulp vraagt, iemand die anderen uitscheldt, enzovoort. Verder kunnen de algemene verwachtingen over de rol van mannen en vrouwen een rol spelen, in plaats van verschillen in milieubesef. Een laatste discussiepunt is: is 'opruimen' een kwestie die onder milieubesef valt?
- b. Onderzoek op dit terrein laat altijd grote verschillen zien tussen wat mensen zeggen, en wat ze doen. De validiteit van wat mensen zeggen kan flink vergroot worden door het gebruiken van 'scènes uit het dagelijks leven' (zie c).
- c. Deze benadering is efficiënt en werkt goed. Maar er is heel wat voorbereidend werk nodig om goede situatiebeschrijvingen op te stellen, die relevant zijn voor de onderzochte populatie. Het moeten bovendien situaties zijn waarin de gebruikelijke rolverschillen en rolverwachtingen geen rol spelen.
- d. Deze vragen slaan op werkelijk gedrag, en staan dus in principe dicht bij wat je wilt weten dan vragen over houdingen. Maar er zijn ook hier vertekeningen mogelijk: er zijn kosten (die niet voor iedereen hetzelfde wegen) aan verbonden, en voor sommige mensen is 'ergens lid van zijn' of 'contributie moeten betalen' belangrijker dan werkelijk milieuvriendelijk gedrag (dat meer kost in termen van persoonlijke offers).

- 8.6** Vergelijk het meten van de lengte van een voorwerp met een niet erg precies instrument. Je kunt dan een betere schatting van die lengte maken door het gemiddelde te nemen van een aantal afzonderlijke metingen. De gedachte daarachter is dat elke meting beïnvloed wordt door een aantal oorzaken (die ieder voor zich niet zo belangrijk zijn), en die op de lange duur elkaar neutraliseren. Daarom beginnen we met veel metingen, en eindigen met per variabele één waarde.

Hoofdstuk 9

9.1 *

- 9.3** Het aggregatieniveau is: industriële ondernemingen. De afhankelijke variabele is het aantal niet gewerkte dagen ten gevolge van ziekte, gedeeld door het totaal aantal dagen \times het totaal aantal werknemers, $\times 100$. Fysieke werkdruk, de gemiddelde leeftijd van de werknemers en de houding van de werknemers ten opzichte van het bedrijf zijn voorbeelden van mogelijke onafhankelijke variabelen die het ziekteverzuim bepalen. De werking wordt in een diagram voorgesteld door een pijl.
- 9.4** Het beroep van de grootvaders heeft een invloed op het beroep van de zoons, en het beroep van de zoons beïnvloedt de beroepskeuze van de kleinzoons. Er kan zelfs ook nog een directe invloed (grootvaders op kleinzoons) zijn (tegenwoordig zijn deze invloeden natuurlijk zwakker dan vroeger). Belangrijk is de toevoeging van andere oorzakelijke variabelen. Elk wordt voorgesteld met een punt en een pijl naar het beroep van de kleinzoons.
- 9.5** Het aggregatieniveau is: individuele leerlingen. Relevante variabelen kunnen zijn: het prestatieniveau op de verschillende schoolvakken; het beroep van vader/moeder; de houding ten opzichte van de te kiezen studie (deze laatste is een wat al te gemakkelijke variabele, die waarschijnlijk hoog correleert met de keuze, maar die inhoudelijk weinig interessant is). Op een hoger niveau van aggregatie kun je denken aan de voorkeuren van broers en zussen, aan die van de beste vrienden en aan het algemene klimaat wat betreft toekomstperspectieven in de schoolklas.

Hoofdstuk 10

- 10.1** Piet is vaker op tijd dan Jan, zowel overdag als 's nachts. Maar omdat Jan bijna alle nachtritten draait en Piet bijna alle ritten overdag (en overdag zijn de vertragingen veel groter dan 's nachts) is het eindresultaat dat Jan vaker op tijd is dan Piet. Het is een voorbeeld van het bizarre effect van statistische interactie: de variabele Jan/Piet interacteert met de variabele overdag/'s nachts in de werking op de afhankelijke variabele het al dan niet op tijd zijn.
- 10.2** In de subtabel voor HOGE opleiding is de odds ratio tussen de andere variabelen 1.07. Voor degenen met een LAGE opleiding is de odds ratio 1.06. Over de totale tabel is de odds ratio tussen motivatie en resultaten echter gelijk aan 1.65. Dit resultaat

suggereert dat het niveau van de voor-opleiding de oorzaak is dat er een correlatie is tussen motivatie en resultaten.

- 10.3** Geslacht is een oorzaak van het al dan niet lid worden van een sportclub, maar bovendien interacteren beide variabelen in hun invloed op het roken.
- 10.4** Een model van schijnbare causaliteit: het huwelijkscijfer kan als de gemeenschappelijke oorzaak aangemerkt worden.