

Hoofdstuk 2

Methodologie in vogelvlucht

2.1 Inleiding

In dit hoofdstuk behandelen we algemene begrippen die van belang zijn voor de multivariate technieken die we verderop gaan bespreken. De behandeling is telkens kort en op hoofdlijnen. De lezer zij verwezen naar Bijleveld & Van der Geest (2021) voor een nadere inleiding in de methodologie van criminologisch onderzoek.

2.1.1 Meetniveaus

Meten is het toekennen van getallen aan waarnemingen. Dat toekennen van getallen kan op verschillende manieren gebeuren. Zo kunnen we bijvoorbeeld delicten classificeren als ‘vermogen’, ‘geweld’, ‘openbare orde’, enzovoort. We kunnen de verschillende categorieën ook getallen geven: ‘1’, ‘2’, ‘3’, enzovoort. We kunnen ook het aantal eerdere veroordelingen tellen: personen hebben 0, 1, 2, enzovoort eerdere veroordelingen. Dit zijn voorbeelden van verschillende variabelen met verschillende *meetniveaus*. Het zij duidelijk dat de getallen 1, 2 en 3 van de soorten delicten geen getallen zijn waarmee te rekenen valt. De getallen 0, 1, 2, enzovoort voor de aantallen eerdere delicten zijn daarentegen wel getallen waar bijvoorbeeld een gemiddelde voor uit te rekenen valt. De vraag in hoeverre en op welke manier te rekenen valt met de getallen die wij aan waarnemingen hebben toegekend, hangt af van het niveau waarop gemeten is. Wat een meetniveau is, zullen wij zo dadelijk uit de doeken doen. Daarvóór vermelden wij echter al dat het van belang is goed vast te stellen wat zo’n meetniveau is waarop gemeten is, aangezien dat meetniveau in veel gevallen bepalend is voor het soort berekeningen en analyses dat op de gegevens uitgevoerd mag worden.

(1) *Nominaal meetniveau*

Op het nominale meetniveau kunnen getallen worden toegekend aan waarnemingen, maar die hebben geen andere functie dan de waarnemingen een label te geven, ze te classificeren. De soorten delicten worden met de labels die we ze geven geassocieerd in ‘geweld’, ‘vermogen’, enzovoort, maar de getallen representeren niet dat het ene delict ‘meer’ zou zijn dan het andere. De getallen dienen slechts om de verschillende soorten delicten te onderscheiden. Met andere woorden, in plaats van 1, 2 en 3 hadden we de delicten ook de codes 27, –83 en 1366.88899 mogen meegeven. De informatie besloten in die laatste getallen is even groot als in de eerste getallen: we weten of delicten gelijk zijn of verschillend zijn. Een nominale variabele is daarmee een variabele die de respondenten indeelt in ongeordende klassen. Voorbeelden van andere nominale variabelen dan type delict zijn godsdienst en haarkleur.

(2) *Ordinaal meetniveau*

Bij het ordinale meetniveau zit er een ordening, een rangorde in de getallen die aan de metingen zijn toegekend. Dat wil zeggen dat als de ene waarneming een grotere waarde heeft gekregen dan de andere, die eerste waarde meer van de gemeten eigenschap heeft dan de tweede. Een voorbeeld van dat soort metingen zijn zogenoemde *preferentieoordelen*. We kunnen respondenten bijvoorbeeld vragen welk soort delict zij erger vinden: een vermogensdelict of een openbare-ordedelict, of fraude of winkeldiefstal, en dan krijgen we preferentieoordelen. Bij ordinale variabelen mogen we de getallen ook veranderen, of ‘transformeren’, zolang we maar zorgen dat de oorspronkelijke ordening gehandhaafd blijft: als A meer had van de eigenschap dan B, dan moet dat na transformatie zo blijven. Als ‘fraude’ dus de ergheidsscore 8 heeft, ‘vandalisme’ de ergheidsscore 7 en ‘winkeldiefstal’ ook de ergheidsscore 7, dan mogen we deze getallen veranderen, bijvoorbeeld in 16 voor ‘fraude’, en 10.3 voor ‘vandalisme’ en 10.3 voor ‘winkeldiefstal’ – als we maar zorgen dat de oorspronkelijke ordening gelijk blijft, en dat scores die gelijk waren gelijk blijven. We hadden dus ook kunnen kiezen voor 283.1114 voor ‘fraude’, en 2.899 voor ‘vandalisme’ en ‘winkeldiefstal’. Een dergelijke transformatie van de getallen noemen we een *monotone transformatie*.

In sommige gevallen wordt een iets ruimere vorm van transformatie ook toegestaan. Bijvoorbeeld dat opeenvolgende getallen (waartussen een verschil was) na transformatie gelijk worden: de rij ordinale waarden 12, 10, 8, 7, 6, 4, 3, 2 wordt veranderd in 10, 9, 8, 7, 7, 5, 4, 3. De oorspronkelijke getal-

len 7 en 6 zijn nu allebei 7 geworden. Dit noemen we een *zwak monotone transformatie*.

(3) *Interval meetniveau*

Bij het interval meetniveau is nog meer informatie besloten in de metingen: niet alleen is het zo dat de metingen laten zien hoe de waarnemingen geordend zijn, nu is het ook zo dat intervallen tussen de getallen even groot zijn. Een klassiek voorbeeld van een meting die op interval schaalniveau is verricht, is een zogeheten *rating scale*, een schaal met categorieën die lopen van 1 tot 5 waarop de respondenten kunnen aangeven of ze het 'volstrekt oneens (1)', 'oneens (2)' ... 'volstrekt eens (5)' zijn met een bepaalde stelling. Veel schalen die wij gebruiken in onderzoek (bijvoorbeeld attitudemetingen) zijn op interval schaalniveau gemeten. Bij interval variabelen mogen we de labels ook anders kiezen, bijvoorbeeld als lopend van -2 naar $+2$. De aanname is dat de afstand tussen 1 en 3 twee keer zo groot is als die tussen 3 en 4 (en in het tweede voorbeeld die tussen -2 en 0 twee keer zo groot als die tussen 0 en 1).

Een ander voorbeeld is de meting van temperatuur. Of ik nu in graden Fahrenheit of in graden Celsius gemeten heb, als het temperatuurverschil tussen Cambridge en Oegstgeest twee keer zo klein is als het temperatuurverschil tussen Nashville en Denver, dan geldt die conclusie ongeacht in welke eenheid ik de temperatuur gemeten heb. Interval metingen hebben *geen vast nulpunt*.

De transformatie die wordt toegestaan bij interval variabelen is de lineaire transformatie: $f(x) = a + bx$.

(4) *Ratio meetniveau*

Bij metingen op ratio schaalniveau ligt er nog meer informatie in de metingen besloten. Nu is het niet alleen zo dat de verhouding van de verschillen tussen metingen vastligt, maar ook dat de verhouding zelf van de metingen vastligt. Een klassiek voorbeeld is de prijs van een product. Als een bepaald boek twee keer zo duur is als een tijdschrift, dan is dat zo ongeacht of ik die prijs uitdruk in euro's of in ponden of in Amerikaanse dollars. Een ratio meting heeft dan ook een *vast nulpunt*: gratis is gratis, ongeacht welke munteenheid men hanteert.

De transformatie die wordt toegestaan bij metingen op ratio niveau is van de vorm: $f(x) = ax$.

(5) *Absoluut meetniveau*

Bij het absolute meetniveau aangekomen ligt eigenlijk alle informatie vast (vandaar de naam). Een typisch voorbeeld van dit soort metingen zijn frequentietellingen: als een respondent 7 delicten heeft gepleegd, is het niet zinnig om dit in een andere eenheid uit te drukken: er staan eenvoudigweg 7 delicten, en dat is het gemeten getal.

Enigszins nuffig wordt wel gezegd dat de enige toegestane transformatie hier de identieke transformatie is: $f(x) = x$.

Samenvattend is het dus zo dat hoe lager het meetniveau (bijvoorbeeld nominaal), des te meer vrijheid de onderzoeker heeft om de metingen te transformeren. Maar hoe lager het meetniveau, des te minder informatie ligt er ook in de metingen besloten (vandaar natuurlijk ook de vrijheid om te mogen transformeren). Hoe hoger het meetniveau (met als hoogste absoluut), des te geringer de mogelijkheden om de scores in andere getallen om te zetten.

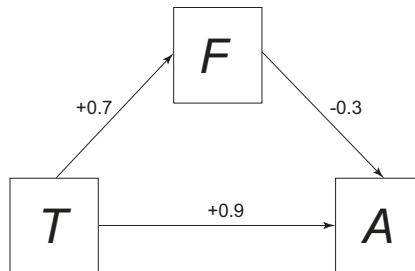
Een speciaal geval vormen de *dichotome variabelen*, dat wil zeggen variabelen met slechts twee categorieën, zoals slaag/faal voor een tentamen, of variabelen met alleen ‘ja’ of ‘nee’ als antwoordcategorie. Omdat er door twee categorieën altijd een rechte lijn kan worden getrokken (waarna de twee categorieën ook perfect op die rechte lijn liggen), kunnen deze speciale nominale variabelen toch vaak als interval variabelen in analyses worden meegenomen. Er zijn ook speciale technieken, erop toegesneden om dit soort variabelen mee te kunnen analyseren.

Het meetniveau is van belang voor de keuze van de analysetechniek. In de praktijk worden voor interval, ratio en absolute metingen dezelfde standaard (kwantitatieve) technieken gebruikt; voor ordinale en nominale metingen worden andere statistische maten en technieken ingezet. Zie bijvoorbeeld Tabel 3.1 in hoofdstuk 3.

2.2 Afhankelijke en onafhankelijke variabelen

Als we wetenschappelijk onderzoek doen, dan proberen we doorgaans een fenomeen of een kenmerk, een verschijnsel, te verklaren. We proberen te beschrijven met welke andere fenomenen het verschijnsel verband heeft, en ook waarmee het niet samenhangt, en daarbij proberen we uit te sluiten dat we onze waarneming door spurieuze verbanden laten vertroebelen. Idealiter proberen we onze vinger achter het verschijnsel te krijgen, dat wil zeggen dat we proberen uit te vinden wat het verschijnsel veroorzaakt. Hoe komt het dat ex-gedetineerden recidiveren? Welke factoren maken dat sommige slachtoffers van een zedendelict zich na aan-

Figuur 2.1: Model voor het effect van kijken naar gewelddadige films op agressief gedrag



gifte nog beroerder voelen? Hoe kunnen we de celcapaciteit van gevangissen voorspellen?

In die laatste voorbeelden is er duidelijk telkens één factor of variabele die de centrale maat is waarin we geïnteresseerd zijn: recidive, welbevinden, celcapaciteit. Dit noemen we de *afhankelijke variabele*. Een afhankelijke variabele is een variabele die we proberen te voorspellen uit andere variabelen. Anders gezegd: de afhankelijke variabele is het gevolg waarvan we de oorzaken willen vinden. De afhankelijke variabele wordt meestal geschreven als Y .

Die onafhankelijke variabelen zijn dan, enigszins kort door de bocht geformuleerd, de oorzaken. De onafhankelijke variabelen kunnen zelf ook weer uit andere onafhankelijke variabelen voorspeld worden, of ingrijpen op het effect van een andere onafhankelijke variabele, in welk geval ingewikkelder terminologie om de hoek kan komen kijken, zoals intermediërende en interveniërende variabele, of moderator variabele. Waar het in dit boek om gaat, is dat de lezer goed onderscheid kan maken tussen aan de ene kant die variabelen die in het model als sturend of veroorzakend of genererend worden beschouwd, en aan de andere kant de variabelen die daarop reageren. Onafhankelijke variabelen worden ook *predictoren* en in sommige analyses zelfs *covariaten* genoemd. Zij worden meestal aangeduid als X_1, X_2, X_3 , enzovoort.

Een voorbeeld wordt gegeven in Figuur 2.1, waar een fictief model is geschetst van het effect op agressief gedrag van blootstelling aan gewelddadige films. Agressief gedrag (A) is hier de afhankelijke variabele, temperament (T) is een onafhankelijke variabele, en het kijken naar gewelddadige films (F) in zekere zin zowel een onafhankelijke als een afhankelijke variabele. In de figuur is met getallen – die voorlopig even het beste als correlatiecoëfficiënten gelezen kunnen worden – weergegeven hoe de variabelen elkaar beïnvloeden. Zoals de figuur laat zien, is er een sterk verband tussen het hebben van een agressief temperament (T) en

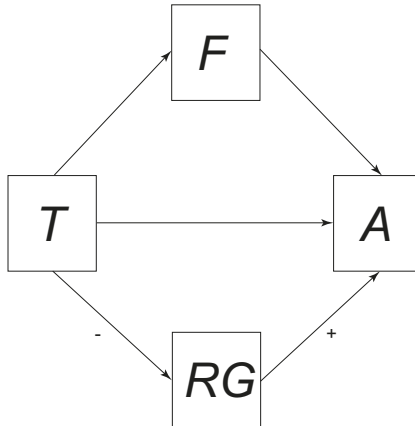
agressief gedrag (A): de coëfficiënt is 0.9. Er is tevens een sterk verband tussen het hebben van een agressief temperament en het kijken naar gewelddadige films: de coëfficiënt is 0.7. Tussen het kijken naar de films en het daadwerkelijk agressieve gedrag is echter zelfs een (zwak) negatief verband: de coëfficiënt is -0.3 . Het lijkt er dus op alsof, als er al een verband is tussen het kijken naar gewelddadige films en agressief gedrag, het kijken naar dit soort films eerder agressieregulerend en -dempend werkt dan andersom. Het verband tussen het temperament en het gedrag is verreweg het sterkst. We zien dat het temperament in zekere zin ook bepaalt of iemand naar dit soort films gaat kijken. Mogelijk hadden we als we alleen gekeken hadden naar het verband tussen film en agressiviteit, een redelijk sterk verband gevonden; in deze opzet, waar de achterliggende variabele temperament in het verhaal betrokken wordt, is de conclusie veel genuanceerder.

2.3 Modellen

Modellen gebruiken wij in het dagelijks leven ook: schaalmodellen of maquettes, modellen van interactie in bedrijven, computersimulatiemodellen, enzovoort. Dit soort modellen heeft de functie om de belangrijkste structuren van het fenomeen in kwestie weer te geven, de kern, of het wezen van datgene wat het model beoogt voor te stellen. Zo is een maquette een verkleinde, maar ook geabstraheerde voorstelling van bijvoorbeeld een bouwproject. Modellen hebben dus structuur-overeenkomst met de werkelijkheid die zij beschrijven. Modellen staan echter los van die werkelijkheid, in de zin dat zij ons in staat stellen om onafhankelijk daarvan zelf te experimenteren: wat gebeurt er in de computersimulatie als ik de waarde van een bepaalde variabele verander? Het veranderen van die waarde dient geen invloed te hebben op de werkelijkheid die gemodelleerd wordt, anders is het niet meer mogelijk om te onderzoeken wat er in die werkelijkheid zou gebeuren als ...

Als laatste is het zo dat – en dat is precies de reden dat wij regelmatig met modellen werken – wij het model makkelijker kunnen waarnemen, beter kennen dan de werkelijkheid die het model representeert. Dat is de crux, en de bruikbaarheid van het model. Vaak is de werkelijkheid complex, en onoverzichtelijk. Het kleinere, in structuur overeenkomende en onderzoekbare en manipuleerbare model geeft ons de gelegenheid om, zonder te hoeven letten op allerlei verstorende en complicerende factoren, te onderzoeken hoe die werkelijkheid zich gedraagt en verandert, wat de eigenschappen van die werkelijkheid zijn. Samengevat is een constellatie een *model* van de werkelijkheid als het model ten eerste *onafhankelijk* is van de werkelijkheid, ten tweede *bekender* is dan de werkelijkheid, en ten derde in structuur *overeenkomt* met die werkelijkheid.

Figuur 2.2: Uitgebreider model voor het effect van kijken naar gewelddadige films op agressief gedrag



Er zijn diverse soorten modellen. Wij noemden al de maquette, die men kan zien als een schaalmodel, die zowel fysiek kan zijn als op de computer getekend. Een dergelijk model wordt ook wel een empirisch model genoemd. Voorbeelden van andere empirische modellen zijn robots, of het concept van het ‘toverleitje’ van Freud, als empirisch model voor de werking van het geheugen. Als op het leitje geschreven wordt, staat het geschrevene in de waslaag die op het leitje zit. Door het papier weg te halen, waarop geschreven is (of het wissertje op en neer te halen in de modernere versies), verdwijnt het geschrevene, zoals wij dingen vergeten. Tegelijkertijd blijft er echter een afdruk zichtbaar van de tekst, er blijven dus sporen.

Het schema dat in Figuur 2.1 werd gegeven is ook een model. Het is een schematische, simpele weergave van een gecompliceerdere werkelijkheid. De werkelijkheid is gecompliceerder omdat in de werkelijkheid waarschijnlijk de context waarin de films worden gekeken een rol speelt, het commentaar dat ouders of leerkrachten of partners geven op de films van belang kan zijn, er mogelijk verschillende dimensies van agressiviteit (reactieve of proactieve agressiviteit) een rol spelen, en mogelijk allerlei andere persoonlijkheidstrekken interacteren (bijvoorbeeld neuroticisme, spanningsbehoefte, enzovoort).

Het simpelere model vertelt ons hoe de belangrijkste factoren op elkaar inwerken om een bepaald resultaat te genereren. Het model maakt het ook mogelijk om te toetsen of de relaties die geschetst worden daadwerkelijk zo bestaan.

Stel dat we nog niet de coëfficiënten hadden die we hiervoor bespraken, maar dat we zouden willen uitzoeken hoe de diverse verbanden liggen. Als we dat zou-

den doen en zouden vinden dat het verband tussen T en F gelijk is aan 0, dan zouden we concluderen dat temperament geen rol speelt bij de keuze om naar gewelddadige films te gaan kijken. De pijl tussen T en F kan dan dus verwijderd worden. Daarmee wordt het model, en daarmee onze theorie over hoe de werkelijkheid in elkaar zit, simpeler en dus weer anders.

We zouden ook andersom te werk kunnen gaan. We zouden een variabele kunnen toevoegen die aangeeft dat de mate van agressief gedrag mede beïnvloed wordt door de ruimte die de gezinsomgeving (RG) geeft aan het uiten van agressief gedrag. Als we zouden vinden dat die RG zelfstandig meebepaalt hoeveel agressief gedrag iemand kan vertonen, zouden we een pijl aan ons model moeten toevoegen, zie Figuur 2.2. Het model wordt dan gecompliceerder, we hebben immers een extra verklarende factor toegevoegd. En stel dat we nog verder gaan, en dat we vinden dat het temperament meebepaalt hoeveel ruimte het gezin geeft (bijvoorbeeld ouders die al veel te verduren hebben gehad van kinderen pikken minder agressieve expressie), dan komt er een pijl tussen T en RG (met een negatieve coëfficiënt). Het model wordt dan weer gecompliceerder, we hebben nu immers een extra relatie toegevoegd. Men zegt dan ook wel: er is een extra *parameter* aan het model toegevoegd – de coëfficiënten in dit soort modellen worden parameters genoemd.

In het algemeen geldt dat simpele modellen makkelijker werken, eenvoudiger te toetsen zijn, maar minder verklaringskracht hebben. Zij kunnen, gemiddeld gesproken, eenvoudigweg minder goed de doorgaans gecompliceerde werkelijkheid beschrijven. Hoe goed een model de werkelijkheid beschrijft, wordt veelal uitgedrukt in een getal dat weergeeft hoe dicht de voorspelde afhankelijke variabele in de buurt van de echte, de waargenomen afhankelijke variabele zit. Als de afhankelijke variabele volgens de voorspelling sterk lijkt op de waargenomen afhankelijke variabele, dan zeggen we dat de *model fit*, ook wel de *goodness of fit* genoemd, hoog is. Het model past dan goed op de gegevens uit de werkelijkheid. Dat is een wenselijke situatie. We hebben dan een model dat het ‘goed doet’.

Complexe modellen beschrijven de werkelijkheid zoals gezegd gemiddeld beter. Als we meer variabelen meenemen, als we alle mogelijke wijzen waarop variabelen elkaar onderling kunnen beïnvloeden in het model specificeren, dan hebben we allicht de mogelijkheid om de causale mechanismen beter te beschrijven en dus een betere fit te krijgen – en krijgen we ook vaak een betere fit. Complexere modellen hebben echter als nadeel dat ze ook te gecompliceerd kunnen worden: te ingewikkeld om te bevatten, of te gecompliceerd om gegeven het aantal respondenten op de data te toetsen. Dat soort heel complexe modellen kunnen ook triviaal worden, in de zin dat zij misschien de scores op de afhankelijke variabele goed beschrijven, maar statistisch kapitaliseren op allerlei toevalligheden en toevallige combinaties van scores die het voor deze dataset toevallig goed doen in de voorspelling van de afhankelijke variabele. Zeker als er, bij heel gecompliceerde

modellen, bijna evenveel of meer variabelen dan respondenten zijn, kunnen we eigenlijk nauwelijks meer van een model spreken. Een model dient immers een vereenvoudigde schematische representatie van de werkelijkheid te zijn.

Voor modellen geldt dus in zijn algemeenheid dat simpelheid ('parsimonie' of zuinigheid) de voorkeur geniet. Andersom geformuleerd: het is niet zo moeilijk om een perfecte fit te krijgen, als je maar ontelbaar veel voorspellers meeneemt. De kunst is om dat met een zo eenvoudig mogelijk model voor elkaar te krijgen. Daarom gebruiken veel technieken tegenwoordig een maat voor de fit die de fit niet in absolute zin berekent, maar relatief op het aantal parameters, en dus relatief op de complexiteit van het model. Dit soort fitmaten 'straf' in zekere zin voor complexiteit. Een goede fit met behulp van veel parameters is dan minder aantrekkelijk dan een goede fit met weinig parameters. Een voorbeeld is het *AIC*, het Akaike Informatie Criterium, dat in paragraaf 7.1.4 wordt besproken.

Dat een model goed past, wil overigens niet zeggen dat we het causale mechanisme 'te pakken hebben', dat we ontrafeld hebben wat de motor is achter onze afhankelijke variabele. Mogelijk bekend is dat de oude Grieken hadden opgemerkt dat als men in de bergen verbleef, men geen malaria opliep. Dit was veelvuldig beschreven, en de redenatie die dit in de antieke tijd had opgeleverd, was dat men malaria opliep van het laagland. Als we een analyse zouden doen op het aantal gevallen van malaria in een tropische streek en we zouden de hoogte van de locatie waar de onderzochte personen verbleven als onafhankelijke variabele opnemen, dan zouden we waarschijnlijk inderdaad een vrij goede fit vinden. Toch is het laagland niet de oorzaak van malaria. Er is een tussenliggende variabele, namelijk de aanwezigheid van malariamuggen, die de malaria overbrengen. In zo'n situatie spreken we van *modelmisspecificatie*. Zie verder Bijleveld & Van der Geest (2021, hoofdstuk 1).

2.4 Operationaliseren en meten

Als wij uitspraken willen doen over de geldigheid van theorieën, als wij beschrijvende vragen willen beantwoorden, dan zullen wij moeten meten. Wij zijn alleen geïnteresseerd in empirisch onderzoek, en dus moeten wij waarnemingen aan de werkelijkheid verrichten, en zo aan gegevens komen die ons kunnen vertellen in hoeverre een bepaalde theorie houdbaar is, of wat het antwoord is op een vraag, bijvoorbeeld naar recidive. Wij zullen daartoe nu behandelen welke stappen er meestal gezet moeten worden om te meten wat wij willen meten.

2.4.1 Operationalisatie

Theorieën, in de psychologie, criminologie, sociologie, bevatten allerhande begrippen. Zo beschrijven Gottfredson & Hirschi (1990) hoe een enkele factor, die zij *self-control* noemen, verklaart of mensen delinquentie gaan plegen. Durkheim (1895,1965) en later Merton (1938) gebruikten in hun theorieën het concept ‘anomie’. Psychologen hebben het over ‘intelligentie’ of over ‘extraversie’, pedagogen over ‘hechting’. Ook relevant kunnen zijn zogenoemde ‘constructen’ als ‘vertrouwen in de politie’, ‘sociale cohesie’ en ‘armoede’. Los van theoretische noties zijn begrippen zoals ‘zinloos geweld’ en ‘veelpleger’. Al deze begrippen zijn abstracte noties, waarvan we wel ongeveer aanvoelen wat ermee wordt bedoeld.

Op het moment echter dat we in de praktijk onderzoek willen gaan doen, moeten we veel preciezer aangeven wat we dan bedoelen met die intelligentie of die anomie. Deze begrippen zijn namelijk niet direct afleesbaar, zoals lengte, schoenmaat of bloeddruk wel zouden zijn (alhoewel overigens ook daar meetprotocollen zijn). Waar lezen we intelligentie aan af? Of armoede? Of gewetensontwikkeling? Dit soort niet direct afleesbare begrippen wordt daarom wel *constructen* genoemd. Ook komt men wel tegen de aanduiding *hypothetisch construct*, waarmee bedoeld wordt een abstracte term waarmee naar een niet-uitwendig waarneembaar kenmerk van personen wordt verwezen, dat verondersteld wordt aanwezig te zijn. Aangezien die trekken niet uiterlijk waarneembaar zijn, betekent dat dat zij afgeleid zullen moeten worden uit andere wel waarneembare kenmerken.

Voordat wij dat kunnen doen, en aan het meten kunnen slaan, zullen wij eerst precies moeten definiëren wat wij onder die constructen verstaan. Wij zullen zoals dat heet een *conceptuele definitie* moeten geven van de constructen. Zo definiëren Gottfredson & Hirschi (1990) lage self-control als de neiging van een individu om delicten te plegen, door een gebrek aan remmen, net als hoge self-control de neiging van een individu verklaart om zich te conformeren aan sociale normen en de wet. Een ander voorbeeld van een construct dat definitie behoeft, is anomie. In de sociologie wordt anomie onder andere gedefinieerd als een verzamelnaam voor een toestand van apathie, vervreemding en onbehagen vanwege het verlies van doelen. Martindale (1960) echter definieerde anomie weer als het omgekeerde van solidariteit, oftewel ‘(...) anomie is a state of confusion, insecurity, “normlessness” (...)’. Deze voorbeelden maken duidelijk dat verschillende auteurs aan eenzelfde label andere betekenissen kunnen verlenen. Het is dus van belang niet alleen op het label van een construct af te gaan, maar precies te onderzoeken wat in een theorie onder een construct verstaan wordt.

Na de eerste stap waarin de conceptuele definitie van een construct gegeven wordt, komt de tweede stap waar het in deze paragraaf om gaat: hoe gaan wij vervolgens, als wij weten wat onder dat construct verstaan wordt, dat construct

meten? Wij willen immers alleen onderzoek doen naar empirisch meetbare fenomenen; zie hierover in het kader van criminologisch onderzoek Bijleveld & Van der Geest (2021). Wij zullen dus moeten vaststellen met behulp van welke operaties wij gaan vaststellen óf, en zo ja hoeveel van, dat construct in de werkelijkheid voorkomt. We kunnen immers niet aan personen vragen: ‘Bent u anomisch?’, ‘Is er hier een subcultuur van geweld?’, of ‘Bent u intelligent?’ Zelfs als we redelijk uitgewerkte conceptuele definities hebben, bijvoorbeeld in onderdelen of dimensies, zullen we moeten verzinnen hoe we dan kunnen vaststellen in hoeverre de onderdelen of dimensies aanwezig zijn.

Stel dat wij gedefinieerd hebben dat gewetensontwikkeling voor ons vier componenten of dimensies omvat: schuld, schaamte, spijt en empathie. Dan nog moeten wij vaststellen hoe wij dat gaan meten. Wij kunnen niet volstaan met vier vragen ‘Voelt u zich schuldig over het delict?’, ‘Schaamt u zich dat u het delict gepleegd hebt?’, enzovoort. En dat nog eens los van de vraag hoe we gewetensontwikkeling denken te gaan meten als iemand het delict ontkent, of hoe we gewetensontwikkeling gaan meten in een algemene populatie die niet noodzakelijkerwijs delicten heeft gepleegd.

Dat betekent dat wij een zogeheten *operationele definitie* van het construct nodig hebben. Wij zullen moeten vastleggen welke handelingen wij moeten verrichten, welke vragen wij moeten stellen, welk gedrag wij moeten meten, welke indicatoren wij moeten gebruiken, om een meting van het construct te krijgen. Als al die handelingen, al die operaties, zijn vastgelegd, en als ook is vastgelegd hoe uit de combinatie van de metingen de vaststelling van de aanwezigheid van het construct in de empirie volgt, dan hebben wij de operationele definitie gegeven. Dan is duidelijk wat we moeten *doen* om het construct vast te stellen. We zeggen wel eens: als de operationele definitie gegeven is, dan kan de onderzoeker onder de tram komen – een opvolger weet dan precies hoe hij of zij moet gaan meten.

In het voorbeeld hiervoor over gewetensontwikkeling kan dat op allerlei manieren. We zouden de respondenten open vragen kunnen voorleggen: ‘*Ik denk dat mijn slachtoffer <naam> zich voelt*’, of een aantal vragen zoals: ‘*Wat er is gebeurd, is de eigen schuld van <naam slachtoffer>’ ‘mee eens/niet mee eens*’.

Voor belangrijke in veel theorieën gebruikte constructen zijn gevalideerde methoden ontwikkeld om de constructen te meten. Intelligentie wordt veel gemeten met de WISC, of de Raven, morele ontwikkeling met de Morele Ontwikkelings-Lijst, zie Stams et al. (2008), terwijl impulsiviteit met items van gestandaardiseerde persoonlijkheidsvragenlijsten wordt bepaald. Zie hierover verder de teksten over validiteit in Bijleveld & Van der Geest (2021, hoofdstuk 2).