

‘Vandaag wil ik u meenemen op een graafexpeditie, om u te laten zien wat er schuilgaat onder de oppervlakte van ons wetenschapsgebied, en wat daar voor bodemschatten te vinden zijn.’ Dat hield Denny Borsboom op 25 november jl. zijn gehoor voor in de aula van de Universiteit van Amsterdam bij de aanvaarding van zijn leerstoel ‘Grondslagen van de psychologie en de psychometrie’. Onderstaande tekst is een bewerking van zijn oratie. ‘Ik durf wel te stellen dat de psychologische test is voor de psychologie wat de atoombom is voor de natuurkunde.’

WAT GRONDSLAGENONDERZOEK

DE PSYCHOLOGIE TE BIEDEN HEEFT

BODEMSCHATTEN

De psychologie is overal. Zonder dat de meeste mensen er erg in hebben, is zij doorgedrongen in alle facetten van ons bestaan. Dat heeft onze wereld diepgaand veranderd. Wie stemmen hoort wordt vandaag de dag niet meer de schedel gelicht, maar naar een ziekenhuis gestuurd. Niet de afkomst van een kind bepaalt zijn of haar schoolopleiding, maar de score op een psychologische test. De meest succesvolle animatiefilm van 2015, *Inside Out*, visualiseert psychologische mechanismen en emoties waar twee eeuwen geleden niet eens een woord voor bestond. Onze maatschappij is doordrenkt van theorieën, ideeën, en termen die ontleend zijn aan de psychologie. Woorden als 'autistisch', 'iq', 'onbewust', 'cognitieve dissonantie', of 'aangeleerd hulpeloos', zouden iemand die in de 19^e eeuw leefde als koeterwaals zijn voorgekomen, maar zijn nu onderdeel van het dagelijks taalgebruik.

De wetenschappelijke psychologie, die ons dit alles opgeleverd heeft, bestudeert menselijk gedrag, de processen die aan dat gedrag ten grondslag liggen, en de factoren die erop van invloed zijn. Je kunt over de die wetenschap allerlei vragen hebben. Stel: u voelt zich ellendig, en u wordt door de huisarts doorverwezen naar een psycholoog. De psycholoog diagnosticeert een depressie en start een psychotherapeutische behandeling. U zou zich dan kunnen afvragen, waarop deze behandeling is gebaseerd. Voor het antwoord op die vraag zou u te rade moeten gaan bij de klinische psychologie – het vakgebied dat zich bezig houdt met psychische problematiek. Of misschien bent u geïnteresseerd in de hersenprocessen en genen die geassocieerd zijn met depressie. Dan moet u zich verdiepen in de biologische

psychologie – het vakgebied dat de biologische kant van psychologische processen onderzoekt. Mocht u daarentegen willen weten waarom mensen met een depressie zo gemakkelijk sociaal geïsoleerd raken, dan zou u zich kunnen wenden tot de sociale psychologie – het vakgebied dat zich bezig houdt met sociale processen en interacties tussen mensen. Zo is de psychologie opgedeeld in allerlei deelgebieden, die zich ieder met een specifiek onderdeel van menselijk gedrag bezig houden.

Maar je kunt je ook nog een heel ander soort vragen stellen. Een soort vragen dat misschien niet meteen betrekking heeft op één bepaald stukje van de psychologie, maar dat het algemene fundament van een begrip als depressie bevraagt. Wat is depressie eigenlijk voor iets? *Bestaat* depressie echt, zoals stoelen of magnetische velden of lichtbundels, of is het uiteindelijk maar een verzinsel, een etiket dat we gebruiken om bepaalde problemen te benoemen? Kunnen we depressie *meten*, zoals je gewicht of lengte meet, of is dat maar bij wijze van spreken? Is zoiets als depressie, dat een uiterst subjectief fenomeen lijkt te zijn, eigenlijk wel *toegankelijk* voor wetenschappelijk onderzoek, dat naar haar aard mikt op objectivering en kwantificering?

Deze vragen gaan over de grondslagen van de psychologie. Ze lijken op de vragen waarmee kinderen hun ouders tot waanzin brengen, als ze in de 'waarom'-fase zitten. Het zijn vragen die de diepte in willen, die niet zozeer de wetenschap maar de aannames *achter* de wetenschap willen verhelderen. Het zijn vragen die graven.

PSYCHOLOGISCHE TESTS

Wetenschappelijke ontwikkelingen beginnen vaak met een

kleine uitbreiding van onze waarnemingsmogelijkheden. De uitvinding van de telescoop bijvoorbeeld, waarmee Galileo net iets verder kon kijken zodat hij de manen van Jupiter kon zien. Als je het hebt over zulke uitbreidingen van ons waarnemingsvermogen, dan gaat het al snel over thermometers, microscopen, deeltjesversnellers, en fMRI-apparaten. Het komt niet vaak voor dat iemand in zo'n rijtje ook de psychologische test plaatst.

Dat komt omdat de meeste mensen psychologische tests banaal vinden. Ze zien er bijvoorbeeld nooit erg bijzonder uit. Meestal gaat het om een lijstje vragen die in doodgewoon inkt op doodgewoon papier zijn afgedrukt, of een paar puzzeltjes, zoals bij de IQ test. Heel wat minder imponerend dan een fMRI-scanner of de deeltjesversneller in Genève.

Toch is de psychologische test een van de belangrijkste wetenschappelijke uitvindingen die er bestaan. Met de ontwikkeling van de psychologische test richtte de mens voor het eerst de wetenschappelijke methode op zichzelf. Dat heeft enorme consequenties gehad. Het idee dat de mens een onderwerp kan zijn van wetenschappelijke studie, en dat je menselijke eigenschappen kunt meten heeft ons zelfbeeld en onze maatschappij diepgaand veranderd.

In mijn onderwijs gebruik ik om dit duidelijk te maken altijd het voorbeeld van de onderzoeker Gustave Le Bon, die in 1879 uitlegde waarom een intelligentie-onderzoeker geen vrouwen hoeft te bestuderen: 'Alle psychologen die de intelligentie van de vrouw hebben bestudeerd [...] erkennen vandaag de dag dat zij de meest inferieure vorm van menselijke evolutie representeert, en dat zij dichterbij kinderen en wilden staat dan bij de volwassen, beschaafde man. Ze blinkt uit in wispelturigheid, inconsistentie, gedachtenloosheid, afwezigheid van logica, en ze kan niet redeneren. Zonder twijfel bestaan er enkele uitzonderlijke vrouwen, die superieur zijn aan de gemiddelde man, maar zij zijn zo uitzonderlijk als een monstreuze misgeboorte, vergelijkbaar met een gorilla met twee hoofden, en derhalve kunnen wij ze volstrekt negeren' (Le Bon, 1879, geciteerd in Gould, 1980, p. 155, mijn vertaling).

Dat is waar wij vandaan komen, als het gaat om psychologie. Het is van belang even te noteren hoe ontzettend kort geleden het is dat Le Bon dit kon zeggen. Le Bon is nog geen honderd jaar dood. Uw grootmoeder had hem de hand kunnen schudden – of een knietje kunnen geven.

Als we Gustave Le Bon hierheen zouden kunnen halen met een teletijdmachine, en we konden hem even bijpraten over de stand van zaken in de wetenschappelijke psychologie, dan zou hij steil achterover slaan. Want vandaag de dag

blijken vrouwen helemaal niet dommer dan mannen. Het is nog geen honderdvijftig jaar geleden dat de eerste vrouw een universiteit betrad, maar er stromen inmiddels systematisch meer meisjes dan jongens in aan de universiteit.

De psychologische test zou in dat gesprek tussen ons en meneer Le Bon een centrale rol moeten spelen. Het is immers de intelligentietest die ons geleerd heeft dat de verschillen tussen mannen en vrouwen veel kleiner zijn dan vroeger aangenomen werd. Ik wil natuurlijk niet beweren dat de IQ-test in zijn eentje de seksediscriminatie verholpen heeft, maar de emanciperende werking van dit soort instrumenten moet ook zeker niet worden onderschat.

Ook in het doorbreken van klassenverschillen heeft de gestandaardiseerde test een belangrijke rol gespeeld. De allereerste intelligentietest, van de Fransman Alfred Binet, werd onder andere ontwikkeld omdat menselijke beoordelaars zich vooral lieten leiden door het voorkomen, de kleding, en de goede naam van de families van op intelligentie beoordeelde mensen. In ons land heeft de CITO-toets waarschijnlijk een belangrijke rol gespeeld in de emancipatie van kinderen uit achterstandsmilieus, die op basis van de subjectieve inschatting van de leraar nooit hadden kunnen doorleren, omdat een dubbeltje nu eenmaal nooit een kwartje kan worden.

Dat is actueel. In huidige discussies over testwoede en rendements-denken worden gestandaardiseerde metingen soms ideologisch verdacht gemaakt, en wordt gemakkelijk gesuggereerd dat de ongeleide subjectieve beoordeling van leerkrachten en visitatiecommissies eigenlijk meer geschikt zijn om beoordelingen op te baseren. De reden dat een test als de CITO-toets überhaupt bestaat is echter, dat mensen op dit punt niet altijd even betrouwbaar blijken, omdat zij hun medemens – gewild of ongewild – beoordelen naar zijn of haar lidmaatschap van allerlei irrelevante categorieën: man of vrouw, rijk of arm, Nederlands of Marokkaans. Gestandaardiseerde tests hebben daar minder last van: het maakt de CITO-toets niet uit of een kind 'wc' zegt of 'toilet', als het maar weet hoe je het spelt.

Ik durf wel te stellen dat de psychologische test is voor de psychologie wat de atoombom is voor de natuurkunde: een toepassing van het vakgebied die geen vezel in het weefsel van onze samenleving ongemoeid heeft gelaten.

PSYCHOMETRIE

Een uitbreiding van onze waarnemingsmogelijkheden, zoals de psychologische test in de psychologie realiseerde, gaat altijd samen met de vraag hoe de nieuwe instrumenten

werken. Dat is het domein van de psychometrie: de tak van wetenschap die zich bezighoudt met het meten van psychologische eigenschappen.

Als de psychologische test de atoombom van de psychologie is, dan is de psychometrie haar kernfysica. De psychometrie levert de statistische modellen waarmee psychologische tests gemaakt, geanalyseerd, en beoordeeld worden. De psychometrie is daarmee het brandpunt van de psychologie, want vrijwel al het psychologisch onderzoek maakt gebruik van testcores. Wie iets wil weten over mensen moet daarom ook iets willen weten over statistiek – je kunt zelfs volhouden dat wie helemaal niets wil weten van statistiek eigenlijk ook niet echt geïnteresseerd is in de psychologie.

De allereerste psychometricus was Francis Galton. Galton was geïnteresseerd in het meten van individuele verschillen in intelligentie. Hij dacht dat die verschillen verankerd

waren in onze biologie. Zo zouden mensen bij wie informatie sneller door de zenuwbanen reist intelligenter zijn. In zijn antropometrisch laboratorium probeerde Galton daarom intelligentie te meten door mensen allerlei elementaire, biologisch geïnspireerde taakjes te laten doen, vooral op het gebied van de sensorische discriminatie: hoe snel kan iemand bijvoorbeeld onderscheid maken tussen twee naast elkaar liggende kleuren? Galton dacht dat verschillen in intelligentie uiteindelijk terug te voeren waren op verschillen in zulke basale, biologisch geïnspireerde variabelen.

Het werd een ramp. Galton was de eerste die de harde les leerde die ook zoveel onderzoekers na hem moesten ervaren: als het om menselijk gedrag gaat is de biologie niet hard, maar boterzacht. De samenhang tussen biologisch gemeten variabelen en psychologisch gemeten variabelen is doorgaans zó zwak, dat er in de hele geschiedenis van de psychologie geen enkel voorbeeld bestaat van een psychologisch relevante variabele die betrouwbaarder gemeten kan worden met een biologische dan met een psychologische maat. Ook vandaag moeten we dat, ondanks de grote woorden die regelmatig vallen in de neurowetenschap, gewoon weer concluderen. Een simpele vraag als 'Bent U vaak moe?'

verklaart bijvoorbeeld al snel duizend keer zoveel variantie in relevante uitkomstmaten op het gebied van depressie, dan het nieuwe depressie-gen dat eerder in dit jaar in de literatuur opdook (Cai et al., 2015). Wanneer u iemand hoort beweren, dat hij op basis van biologische metingen kan achterhalen wat u kunt, denkt, voelt, of wilt, heeft u van doen met iemand die het verschil tussen wetenschap en *science fiction* niet helder voor de geest heeft.

Dat was honderd jaar geleden ook al zo en daarom zat Galton dus in de penarie. Het intelligentieproject werd gered door de eerder genoemde Binet, die een even praktische als briljante oplossing voor het meetprobleem introduceerde. In plaats van zich te richten op biologisch geïnspireerde maten, legde Binet mensen gewoon een aantal puzzeltjes van verschillende aard voor. Vervolgens telde hij hoeveel puzzeltjes ze konden oplossen: hoe meer puzzeltjes goed,

hoe slimmer (Brybaert & Rastle, 2013).

Het idee, dat je het aantal opgaven dat iemand goed maakt kunt tellen, en dat die simpele telling een maat is voor iemands vaardigheid, dat is echt een heel goed idee. Beter dan het onbewuste van Freud of de leertheorie van Skinner. Misschien wel het belangrijkste idee in de geschiedenis van de psychologie.

De telling in kwestie heet in de psychometrie de somscore. De somscore is berucht en gevreesd.

Eens in de zoveel tijd komt er

weer iemand die een manier gevonden heeft om tests op een andere manier te scoren. Dan worden er subscores vernieuwvuldig en gedeeld, er worden ingewikkelde schattingsalgoritmen gebruikt om allerlei wegingsfactoren te schatten. Dat werkt bijna nooit. De somscore wint vrijwel altijd, en die ene keer dat hij verliest is het verschil zó klein, dat het de moeite niet waard is.

Waarom werkt de somscore zo goed? Dat komt omdat de correlaties tussen testcores die met hetzelfde inhoudelijk domein te maken hebben een zogenaamde positieve topologische ruimte vormen. In de psychometrische volksmond wordt die ruimte aangeduid met de Engelse term *positive manifold*. In het geval van intelligentie houdt de *positive manifold* in dat verschillende soorten puzzeltjes altijd positief correleren. Dat wil zeggen: wie beter is dan

De psychologische test een van de belangrijkste wetenschappelijke uitvindingen die er bestaan

gemiddeld in ruimtelijke rotatie, is waarschijnlijk ook iets beter dan gemiddeld in spelling, en ook iets beter dan gemiddeld in rekenen. De positive manifold van intelligentie werd begin 20^e eeuw ontdekt door de Brit Charles Spearman (Spearman, 1904). Er zijn, naast de positive manifold van intelligentietests, in de psychologie nog andere voorbeelden van positive manifolds, zoals die van persoonlijkheid (alle 'gunstige' persoonlijkheidstrekken hangen positief samen) en psychopathologie (hoe meer problemen op het ene vlak, hoe meer problemen op het andere vlak).

Doordat in een positive manifold verschillende scores positief samenhangen, geeft de somscore al snel een goede samenvatting van de algemene toestand van een persoon op een domein, vooral als alle correlaties ongeveer in dezelfde orde van grootte zijn. Dit is een krachtig principe, dat de basis vormt voor eigenlijk alle metingen in de psychologie. Vrijwel alle theoretische constructen in de psychologie worden geoperationaliseerd met een somscore die gedefinieerd is op een positive manifold: intelligentie, persoonlijkheid, mentale stoornissen, attitudes, noem maar op. Van meegaandheid tot emotionele intelligentie en van depressie tot kwaliteit van leven: het zijn in de praktijk allemaal, zonder uitzondering, somscores.

De somscore is dus de Popeye van de psychometrie, en de positive manifold is zijn spinazie: Als Popeye maar voldoende spinazie te eten krijgt, dan is hij onverslaanbaar.

THEORETISCHE VERANKERING

De operationalisatie van psychologische begrippen met gestandaardiseerde tests, die gescoord worden met somscores, is in principe een heel goed idee. De psychometrie heeft de afgelopen eeuw, in allerlei schitterende wiskundige afleidingen, de basis onder deze praktijk uitgewerkt met meetmodellen. Met behulp van nieuwe technieken kan het psychometrisch gekwantificeerde mensbeeld nu ook gerelateerd worden aan hersenscans en genetica. Alle ingrediënten voor een succesvolle wetenschap zijn aanwezig. Er is maar één probleem. Het werkt niet.

Ondanks het ontegenzeggelijke belang van psychologische tests in de samenleving, en ondanks de prachtige psychometrische modellen die de afgelopen honderd jaar zijn opgesteld, moeten wij constateren dat de psychologie als wetenschap betrekkelijk weinig baat heeft gehad bij de standaardmethodes om psychologische constructen te meten. De precisie, voorspellende kracht, en controlemogelijkheden die worden uitgedrukt in de slogan 'meten is weten', die zijn in de psychologie zelden gerealiseerd.

Ik denk dat dat komt omdat wij er niet in geslaagd zijn een overtuigende theoretische verankering voor onze metingen te vinden. We weten eigenlijk niet goed wat onze tests meten, hoe ze werken, of welke psychologische processen ze oproepen, en we hebben daar ook nauwelijks theorieën over ontwikkeld. Dat is op zijn beurt een symptoom van een dieper probleem; namelijk, dat de psychologie er überhaupt niet in slaagt overtuigende wetenschappelijke theorieën op te stellen.

Het is niet mijn bedoeling hiermee mijn collega's te beledigen. Ik ken mijn eigen werk, dat voor een groot deel uit pogingen tot theorievorming bestaat, zonder meer dezelfde status toe. In mijn ogen zijn grote delen van de psychologie, waaronder ook mijn werk valt, nog steeds in een fase die de



wetenschapsfilosoof Kuhn als 'prewetenschappelijk' aanduidde, juist omdat de psychologie geen theoretische richting kent. De grote theoretische kaders in de psychologie – behaviorisme, psychoanalyse, biologisch reductionisme – hebben stuk voor stuk gefaald, en er is niets voor in de plaats gekomen. Daardoor ontbeert de psychologie een richtinggevend kader, zoals de natuurkunde dat vond in het werk van Newton, de biologie in de evolutietheorie van Darwin, en de computerwetenschap in het idee van de Turing-machine.

Als ik over dit soort zaken vertel, zakken sommige studenten en onderzoekers langzaam weg in een depressieve toestand. Is het allemaal wel ergens goed voor? Moet ik niet toch maar iets anders gaan doen met mijn leven?

Daar heb ik geen last van, en ik vind het ook onterecht.

Menselijk gedrag is, om met Skinner te spreken, 'het moeilijkste onderwerp dat ooit aan een wetenschappelijke analyse is onderworpen'

(Skinner, 1987, p. 780). Zelfs de meest eenvoudige vragen die je

over mensen kan stellen leiden tot moeilijkheden die onze collega's in de exacte wetenschappen zich nauwelijks voor kunnen stellen. Ik zou onze psychometrische problematiek bijvoorbeeld zonder enige aarzeling uitruilen tegen het beruchte meetprobleem in de kwantummechanica. Subatomaire deeltjes die uit hun superpositie vallen zijn ongetwijfeld ingewikkeld, maar ze vragen zich niet af wat de onderzoeker toch van ze wil met al die ingewikkelde proefopstellingen. Mensen doen dat wel.

Een tijdje geleden moest ik bij de avondapotheek een medicijn halen. Er stond een lange rij. Achteraan was een man stennis aan het schoppen. Hij had zijn medicijnen nú nodig, schreeuwde hij. De apotheker vertelde hem, dat hij net als de anderen op zijn beurt moest wachten. 'Maar dat kan ik nou juist niet' riep de man, 'want ik heb ADHD! Daarvoor moet ik juist die medicijnen hebben!' De man verklaarde zichzelf dus aan de hand van een stoornis die de wetenschappelijke psychologie hem heeft aangeleverd, en rechtvaardigde zijn gedrag daarmee. Zo hebben de categorieën, die de psychologie bedenkt om menselijk gedrag in te delen, weer invloed op datzelfde gedrag. Door het ongekende zelfreferentiële vermogen van de mens doet de psychologie

denken aan de mystieke slang Ouroboros, die zichzelf in de staart bijt. Niets in de ons bekende natuur bezit ook maar bij benadering zo'n diepgaande complexiteit. De mens is het moeilijkste onderzoeksobject dat er bestaat.

Het is dus misschien niet zo gek dat de psychologie, met haar krankzinnig ingewikkelde onderzoeksobject, na honderdvijftig jaar onderzoek nog weinig overtuigende theorie heeft opgeleverd. Toch maak ik me op dit punt zorgen. Het lijkt soms namelijk, alsof de psychologie het gevecht een beetje heeft opgegeven. Alsof men genoeg

neemt met de documentatie van basale effecten en relaties tussen variabelen. Daar ga ik niet mee akkoord. Ik denk, of weet eigenlijk wel zeker, dat er op het gebied van theorievorming in de psychologie veel meer mogelijk is dan men nu denkt. Ik zal nu een voorbeeld geven van hoe ik denk dat de psychologen theorievorming kunnen aanpakken.

Het maakt de CITO-toets niet uit of een kind 'wc' zegt of 'toilet', als het maar weet hoe je het spelt

LATENTE VARIABLEN EN ECOSYSTEMEN

Eerder heb ik het gehad over de positive manifold: het fenomeen dat gelijksoortige eigenschappen of vaardigheden in de psychologie doorgaans positief samenhangen. Zo zijn vrijwel alle psychische klachten waarmee een mens te maken kan hebben – van het horen van stemmen tot depressie en van verslaving tot angstigheid – positief gecorreleerd. Mensen met meer problemen van het ene soort, hebben ook meer problemen van het andere soort.

Maar hoe komt zo'n positive manifold nu tot stand?

Waarom hangen gelijkaardige testcores positief samen?

Daarop had de psychologie tot een paar jaar geleden eigenlijk maar één antwoord: omdat zij hetzelfde meten.

Dit verklaringsschema berust op een analogie. Geobserveerde variabelen worden namelijk opgevat als metingen van hetzelfde, net als bijvoorbeeld de uitslagen van verschillende thermometers allemaal afhangen van dezelfde grootte, namelijk de omgevingstemperatuur. Gaat die temperatuur omhoog, dan reageert iedere individuele thermometer daarop door te stijgen. Daardoor zullen de thermometerstanden een patroon van positieve correlaties vertonen: een positive manifold dus.

In de psychometrie worden correlaties op soortgelijke

wijze verklaard. Waarom zien wij een positieve samenhang tussen verschillende maten voor probleemoplossend vermogen? Omdat er een algemene intelligentiefactor is, die al die prestaties bepaalt. Waarom zien wij een positieve samenhang tussen symptomen van depressie? Omdat al deze symptomen worden veroorzaakt door dezelfde onderliggende stoornis. Waarom zien wij een positieve samenhang tussen verschillende soorten evaluaties van politieke partijen? Omdat al die evaluaties dezelfde attitude meten. Iedere keer wordt de samenhang tussen observeerbare variabelen verklaard met een zogenaamde latente variabele: een onderliggend construct dat als gemeenschappelijke oorzaak van die observeerbare variabelen fungeert.

In de afgelopen eeuw heeft psychologie ontzettend hard geprobeerd zulke latente variabelen ook empirisch te identificeren. Voor bijvoorbeeld depressie heeft men geprobeerd de centrale oorzaak van depressiesymptomen te identificeren met aangeleerde hulpeloosheid, verdrongen trauma's, hormonale disbalansen, chemische abnormaliteiten in het brein, en een stoet andere variabelen. Soortgelijke rijtjes kan men opstellen voor andere psychologische stoornissen, maar ook voor intelligentie en persoonlijkheid.

Deze benadering is mislukt. Er is, ook na een eeuw onderzoek, in vrijwel geen enkel geval een serieuze aanwijzing dat de in het psychometrisch model gepostuleerde gemeenschappelijke oorzaak ook echt bestaat. Maar waar komen al die positieve correlaties *dan* vandaan?

Daar heb ik jaren mijn hoofd over gebroken. Tot op een goede dag mijn collega Han van der Maas mijn kamer binnenkam, zwaaiend met een papiertje waarop hij allemaal rondjes met pijltjes ertussen had getekend. Hij lanceerde de hypothese dat de positive manifold ontstaat omdat verschillende IQ-subtests processen meten die elkaar direct beïnvloeden. Een beter werkgeheugen zorgt er bijvoorbeeld voor dat een kind meer strategieën kan gebruiken om rekensommen op te lossen, en toepassing van die strategieën verbetert weer het werkgeheugen. Dus waar hadden we die hele latente variabele dan eigenlijk voor nodig? Was het niet voldoende om aan te nemen dat er directe interacties bestonden tussen de betrokken processen, die ervoor zorgen dat zij als het ware verknoopt raken?

Dat zijn wat mij betreft de mooiste momenten in het wetenschapsbedrijf: dat iemand een figuurtje op papiertje tekent, waarna zich ineens een hele wereld van onderzoeksmogelijkheden opent.

Een psychologisch construct zou, als dit model klopte, immers helemaal niet als gemeten onderliggende variabele

Als het om menselijk gedrag gaat is de biologie niet hard, maar boterzacht

fungeren. Beter past een alternatieve analogie, die gebaseerd is op het idee van een ecosysteem. In een ecosysteem als een bos leven allerlei diersoorten, die van elkaar afhankelijk zijn. Als de konijnenstand stijgt, dan groeit het aantal vossen, want vossen eten konijnen. Deze directe causale interacties leiden ertoe dat diersoorten met elkaar verknoopt raken in een dynamisch systeem. De toestand van het gehele systeem – in ons geval zou dat bijvoorbeeld iemands algemene intelligentie kunnen zijn – *emergeert* uit die lokale interacties, maar is *niet* de oorzaak van de positive manifold. Dat is een heroverweging van het psychometrisch fundament waarop de psychologie rust, oftewel van de grondslagen van de psychologie.

NETWERKEN

In 2008 begon ik met een klein groepje onderzoekers, onder wie Lourens Waldorp, Angélique Cramer, Verena Schmittmann en Sacha Epskamp de mogelijkheden van complexe netwerkbenaderingen te verkennen. Al snel werd duidelijk dat het toepassingsgebied van netwerken veel breder is dan intelligentie alleen. Ook constructen in de psychopathologie, persoonlijkheid, en in het attitudeonderzoek leken zich eigenlijk beter te lenen voor netwerkbenaderingen dan voor het klassieke latente variabelen-model. Daarom vatten wij het plan op om een algemeen bruikbare methodologie te ontwikkelen, die door onderzoekers gebruikt kan worden om netwerkstructuren te analyseren en representeren.

Die methodologie bestaat inmiddels. We hebben data-analytische methoden ontwikkeld om op basis van allerlei soorten data netwerken te schatten, te simuleren, en te analyseren. Deze methodologie is geïmplementeerd in de programma's *qgraph* (Epskamp et al., 2012) en *IsingFit* (Van Borkulo et al., 2014), en wordt nu wereldwijd gebruikt door onderzoekers in allerlei onderzoeksgebieden. Netwerkbenaaderingen zijn succesvol toegepast op onderwerpen als posttraumatische stress (McNally et al., 2015), autisme (Ruzzano et al., 2014), attitudes (Dalege et al., 2015), depressie (Cramer et al., 2012), kwaliteit van leven (Kossakowski et al.,

2015), complexe rouw (Robinaugh et al., 2014), en persoonlijkheid (Cramer et al., 2012).

Het grootste succes van deze benadering tot zover ligt echter op het gebied van de psychopathologie. Daar bleek de koppeling tussen netwerken en mentale stoornissen zo sterk, dat zich – tot mijn verbazing, trouwens – als vanzelf een principiële theorie ontwikkelde over wat stoornissen eigenlijk zijn (Cramer et al., 2010). In deze opvatting ontstaan mentale stoornissen als depressie, gegeneraliseerde angststoornis, of posttraumatische stressstoornis uit interacties tussen de symptomen van die stoornissen. Dat zulke interacties bestaan is zeer plausibel en inmiddels ook

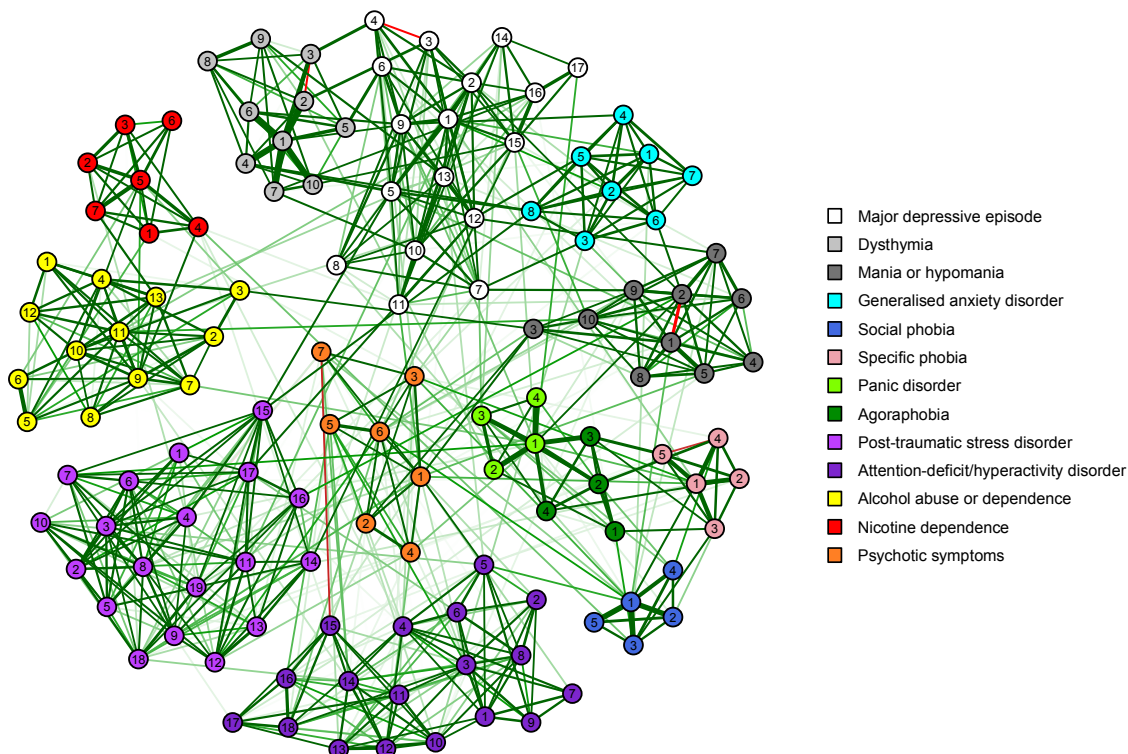
empirisch ondersteund. Bekijken wij de symptomen van depressie, dan vinden wij daar bijvoorbeeld onder andere: slaapproblemen, vermoeidheid, concentratieproblemen, en neerslachtigheid. Welnu: wie niet slaapt wordt moe; wie moe is, kan zich slecht concentreren; wie zich slecht kan concentreren komt nergens toe; wie nergens toe komt, voelt zich neerslachtig, en gaat misschien 's nachts piekeren, waardoor de slaapproblemen weer verergeren. Zo vormen symptomen van mentale stoornissen vicieuze cirkels van problemen, die kunnen culminereren in een neerwaartse spiraal waaruit ontsnappen moeilijk is, en in sommige gevallen zelfs onmogelijk.

FIGUUR 1

Figuur 1. De netwerkstructuur van de *Diagnostic and Statistical Manual of Mental Disorders-IV* zoals bepaald in een steekproef van ruim 34.653 inwoners van de vs. Ieder element in het netwerk representeert een symp-

toom. Positieve (negatieve) relaties tussen symptomen zijn weergegeven met groene (rode) connecties. Symptomen zijn gekleurd naar de dsm-iv-stoornis waarbij zij horen. De figuur is gebaseerd op: Boschloo et

al. (2015) The Network Structure of Symptoms of the Diagnostic and Statistical Manual of Mental Disorders. *PLoS ONE*, 10 (9): e0137621.



Wie de diagnostische handboeken doorwerkt, die kan er eigenlijk niet omheen dat dit soort interacties tussen symptomen bestaan. Vaak zijn ze zelfs met zoveel woorden in de diagnostische criteria opgenomen. Overmatig handenvassen telt bijvoorbeeld alleen als symptoom van obsessief compulsieve stoornis indien het veroorzaakt wordt door een ander symptoom, namelijk het gevoel vies te zijn; roekeloos gedrag in het verkeer telt alleen als symptoom van alcoholverslaving als het veroorzaakt wordt overmatig alcoholgebruik – een variabele die zelf ook weer als symptoom wordt opgevoerd. Dat er hier sprake is van systemen van variabelen die elkaar beïnvloeden kan dus eigenlijk niet serieus ter discussie staan. De ecosysteem-analogie past daardoor *a priori* al vele malen beter dan het latente-variabelenmodel, waarin de correlatie tussen bijvoorbeeld slaapgebrek en vermoeidheid verklaard moet worden door hun gemeenschappelijke afhankelijkheid van een latente stoornis – een tamelijk vergezocht idee.

Met de technieken die wij ontwikkeld hebben kan de netwerkstructuur van de interacties tussen symptomen uit onderzoeksgegevens geschat worden. Een voorbeeld van zo'n netwerkstructuur staat in Figuur 1. Deze figuur is afkomstig uit een recent artikel van Lynn Boschloo en collega's (2015). Het netwerk geeft ieder symptoom uit de vierde editie

van de *Diagnostic and Statistical Manual of Mental Disorders* weer als een bolletje. Een groene verbinding tussen twee bolletjes betekent dat de symptomen in kwestie een positieve correlatie vertonen, die niet weg te krijgen is door rekening te houden met de aanwezigheid van andere symptomen. Dit is een aanwijzing dat er tussen de twee symptomen een causaal verband zou kunnen bestaan, hoewel dit type netwerk niet aangeeft wat de aard of richting van dat verband is. Een rode verbinding geeft een negatieve relatie aan. De meeste verbindingen in de figuur zijn groen, en dat verklaart de positieve manifold: problemen veroorzaken vaak andere problemen, maar heffen die zelden op. Ook zijn in het netwerk de mentale stoornissen uit de DSM-IV duidelijk terug te zien als eilandjes van wat sterker verbonden symptomen. Dit is een vrij robuuste bevinding in ons onderzoek, die mij wel een beetje verbaasde gezien het ad

hoc karakter van de DSM, en gegeven alle kritiek die er op het systeem bestaat. Blijkbaar correspondeert het boek toch wel redelijk met empirische regelmatigheden.

Wat heb je nu aan zo'n netwerk? Ten eerste biedt het een nieuw perspectief om na te denken over de fundamentele vraag wat mentale stoornissen als depressie eigenlijk zijn. In de huidige literatuur zijn er eigenlijk maar twee kandidaat-antwoorden op die vraag: ofwel stoornissen worden gezien als onderliggende oorzaken van symptomen, ofwel als arbitraire groepen van symptomen: het zijn dan willekeurige etiketten, sociale constructies die geen aanspraak op de realiteit kunnen maken. Het netwerkmodel biedt een derde alternatief: stoornissen zijn groepen symptomen die elkaar zodanig sterk beïnvloeden, dat er vicieuze en zelfversterkende cirkels ontstaan. Dat is een nieuw antwoord op één van de grondslagen-vragen: bestaat zoiets als depressie echt? Ja, maar niet als onderliggende ziekte of biologische abnormaliteit. Depressie is, in de netwerkbenadering, een alternatief evenwicht in een probleemnetwerk, dat ontstaat

omdat teveel problemen elkaar te sterk blijven beïnvloeden.

Ook diagnostiek zelf kan terugvoeren in de structuur van het systeem, zoals mij duidelijk werd in de rij bij de apotheek: omdat de man uit dat verhaal zijn stoornis opvoerde als oorzaak van zijn gedrag, ontsloeg hij zichzelf van de plicht om iets aan dat gedrag te doen. Zo werd zijn diagnose

onderdeel van zijn probleemnetwerk. In de netwerkbenadering kunnen stoornissen dus tegelijk te maken hebben met de biologie van ons menselijk systeem, met psychologische factoren en de interactie daartussen, én met de verhalen die wij over onszelf vertellen en waarmee we ons gedrag rationaliseren – inclusief het verhaal van de wetenschap zelf. We kunnen met netwerken dus de jacht op de slang Ouroboros openen. Dat is voor het eerst.

Ten tweede stelt de netwerkbenadering ons in staat een nieuw arsenaal aan verklaringsmechanismen op te stellen. Met netwerkmodellen, zoals hier afgebeeld, kun je de ontwikkeling van stoornissen te simuleren. In zulke computersimulaties kunnen we nagaan welke ons bekende fenomenen door het netwerkmodel kunnen worden verklaard, en welke niet. Dit is een enorme rijkdom voor een vakgebied als de klinische psychologie, dat tot nog toe vooral

De somscore is de Popeye van de psychometrie

aangewezen was op verbale, niet-geformaliseerde theorieën. Simulaties met het netwerkmodel tonen bijvoorbeeld aan dat stoornissen zich als continua of als categorieën kunnen gedragen, afhankelijk van de sterkte van de interacties tussen symptomen. Dat zeer verrassende resultaat lost een oeroude vraag uit de klinische psychologie op, namelijk of stoornissen nu uitlopers zijn van continue dimensies of discrete ziektebeelden. Dat kan dus allebei: sterk verbonden netwerken vertonen plotselinge, schoksgewijze overgangen, terwijl zwak verbonden netwerken zich continu en lineair gedragen. Ik denk zelf dat de schoksgewijze variant, waarin problemen zo krachtig zijn dat ze elkaar systematisch in stand houden, de echte stoornis verraadt, en dat het meer continue gedrag hoort bij onze gewone reacties op de ellende die wij meemaken in het leven.

Ook tonen onze simulaties aan dat, als er in netwerken van symptomen plotselinge, niet-lineaire overgangen optreden, deze worden aangekondigd door dezelfde karakteristieke waarschuwingssignalen die men ziet in andere complexe systemen; bijvoorbeeld vlak voordat een regenwoud verandert in een woestijn. Specifiek worden zulke transities vooraf gegaan door een afnemende weerbaarheid, die zich verraadt door een vertraging in het herstel van een systeem na toevallige schokken. In het geval van depressie zou dat betekenen dat iemand, die op het punt staat in een depressie te schieten, er merkbaar langer over doet om te herstellen van kleine dagelijkse narigheden. In samenwerking met Marieke Wichers uit Groningen en de groepen van Marten Scheffer in Wageningen en Ken Kendler in Virginia hebben we nu in een aantal studies aannemelijk gemaakt dat zulke waarschuwingssignalen inderdaad te zien zijn, voorafgaand aan het optreden van een depressie (Van de Leemput et al., 2014; Wichers et al., *in press*). Dat zou kunnen leiden tot nieuwe technologie om bijvoorbeeld terugval tijdig te detecteren.

Het netwerkmodel biedt tenslotte uitzicht op nieuwe manieren om stoornissen te behandelen. In samenwerking met onderzoekers in de groep van Robert Schoevers in Groningen werken we aan nieuwe methoden om voor individuele cliënten de netwerkstructuur van hun problematiek in kaart te brengen: welke problemen leiden, voor een specifiek individu, tot welke andere problemen (Kroeze et al., *submitted*). Dat doen we door cliënten langere tijd te volgen, om vervolgens patronen van samenhang tussen symptomen te analyseren zoals die zich in de tijd voordoen; een techniek die we samen met Francis Tuerlinckx en Laura Bringmann uit Leuven hebben ontwikkeld (Bringmann et al., 2013). Het

is nog te vroeg om te zeggen of zulke benaderingen ook echt nuttig zijn om behandelingen op te baseren, maar door deze nieuwe technologie zijn de vooruitzichten in mijn ogen nu wel beter dan ze lange tijd geweest zijn.

De ontwikkeling van netwerkmodellen maakt het onderzoekers mogelijk om systematisch aan theorievorming te gaan doen, omdat netwerkmodellen voldoende geformaliseerd kunnen worden om aan te rekenen en aan te simuleren. Maar het maakt ook systematische behandelplanning mogelijk, omdat met een netwerkmodel de effecten van verschillende interventies kunnen worden doorerekend. Zoals de sociaal psycholoog Kurt Lewin zei: er is niets zo praktisch als een goede theorie.

CONCLUSIE

Er is een groot aantal onderwerpen de revue gepasseerd. Via psychologische tests en psychometrie ben ik uitgekomen bij ecosystemen en psychopathologie. Wat hebben deze onderwerpen nu met elkaar te maken?

Het overkoepelende thema is theorievorming. Theorievorming begint heel vaak met een verbetering in ons waarnemingsvermogen. In ons geval is dat de uitvinding van de psychologische test. Zo'n uitvinding stelt ons in staat om nieuwe empirische fenomenen te detecteren. In ons geval is dat de *positive manifold*. Om zo'n fenomeen te verklaren, zoekt de wetenschappelijk onderzoeker naar analogieën uit andere gebieden, die hij of zij al begrijpt. In ons geval is zo'n analogie de basis voor het psychometrische latente variabelenmodel. Daarin worden geobserveerde scores immers gemodelleerd alsof zij metingen van dezelfde grootheid zijn. De stap van latente variabelenmodellen naar netwerkmodellen was mogelijk omdat wij inzagen dat ook een andere analogie past op de *positive manifold*; namelijk die van een ecosysteem. De daaruit volgende kanteling van ons begrip-apparaat heeft enorme consequenties voor de manier waarop wij denken over bijvoorbeeld stoornissen, en daarmee voor de manier waarop wij die onderzoeken.

In het ontdekken van empirische regelmatigheden, zoals de *positive manifold*, is de psychologie heel erg goed. Maar in het opstellen van theoretische verklaringsmodellen niet. Theorievorming in de psychologie is een informele, verbale aangelegenheid die vooral ondernomen wordt door uitgeleerde empirisch onderzoekers op leeftijd, die voor zij met pensioen gaan nog even snel een ei willen leggen. Dat moet veranderen. Theorievorming moet, net als in andere vakgebieden, het domein worden van de allerslimste geesten, in de kracht van hun leven, die waar nodig de meest

geavanceerde wiskundige technieken aanwenden voor het opstellen van theoretische geïnformeerde psychometrische modellen.

Daarvoor moeten we mensen anders gaan opleiden. En dat kan. De strategie van theorievorming door gebruik van analogieën, die is bijvoorbeeld heel goed te leren. Je kunt mensen trainen in het opstellen van goede theorieën, en je kunt ze feedback geven door te laten zien waar hun analogieën mank gaan. Je kunt er ook onderwijs op inrichten. Je kunt over theorievorming leerboeken schrijven. Waarom bestaan die niet? Dat moet ook veranderen. Zo kunnen we in de toekomst niet alleen mensen opleiden die nieuwe feiten produceren, maar ook mensen die het empirisch materiaal aan elkaar kunnen lijmen.

De netwerkbenadering van psychopathologie illustreert de krachten die kunnen loskomen als inhoudelijke theorie

succesvol gekoppeld wordt aan formele psychometrische modellen. Het is daarbij van belang op te merken dat het geheim van die benadering, ondanks alle statistische en data-analytische toeters en bellen, ligt in een puur grondslagenidee: een heroverweging van wat wij eigenlijk aan het meten zijn en hoe we daarover na moeten denken. Dat illustreert hoe het wroeten in de klamme grondslagenaarde de prachtigste bodemschatten kan opleveren, juist omdat het gaat om zulke fundamentele zaken. Wie iets verandert aan de fundamentele van een gebouw, beïnvloedt alle verdiepingen tegelijk.

OVER DE AUTEUR

Denny Borsboom is hoogleraar 'Grondslagen van de psychologie en de psychometrie' aan de Universiteit van Amsterdam. E-mail: d.borsboom@uva.nl.

Literatuur

- Borkulo, C.D. van, Borsboom, D., Epskamp, S., Blanken, T.F., Boschloo, L. et al. Schoevers, R.A. & Waldorp, L.J. (2014). A new method for constructing networks from binary data. *Scientific Reports*, 4:5918. doi: 10.1038/srep05918
- Boschloo, L., Borkulo, C.D. van, Rhemtulla, M., Keyes, K.M., Borsboom, D. & Schoevers, R.A. (2015). The network structure of symptoms of the Diagnostic and Statistical Manual of Mental Disorders. *PLoS One*, 10: e0137621. doi: 10.1371/journal.pone.0137621
- Bringmann, L.F., Vissers, N., Wichers, M., Geschwind, N., Kuppens, P., Peeters, F., Borsboom, D. & Tuerlinckx, F. (2013). A network approach to psychopathology: New insights into clinical longitudinal data. *PLoS ONE*, 8, e60188, 1-13. doi:10.1371/journal.pone.0060188
- Brysbaert, M. & Rastle, K. (2013). *Historical and conceptual issues in psychology*. Harlow: Pearson.
- Cai, N. et al. (2015). Sparse whole-genome sequencing identifies two loci for major depressive disorder. *Nature*, 523, 588-91.
- Cramer, A.O.J., Waldorp, L.J., van der Maas, H. & Borsboom, D. (2010). Comorbidity: A network perspective. *Behavioral and Brain Sciences*, 33, 137-150.
- Cramer, A.O.J., Borsboom, D., Aggen, S.H. & Kendler, K.S. (2012). The pathoplasticity of major depression. *Psychological Medicine*, 42, 957-965. doi:10.1017/S003329171100211X
- Cramer, A.O.J., van der Sluis, S., Noordhof, A., Wicherts, M., Geschwind, N., Aggen, S.H., Kendler, K.S. & Borsboom, D. (2012). Dimensions of normal personality as networks in search of equilibrium: You can't like parties if you don't like people. *European Journal of Personality*, 26, 414-431.
- Dalege, J., Borsboom, D., Harreveld, F. van, Berg, H. van den, Conner, M. & Maas, H.L.J. van der (2015). Toward a formalized account of attitudes: the Causal Attitude Network (CAN) model. *Psychological Review*. doi: 10.1037/a0039802
- Epskamp, S., Cramer, A.O.J., Waldorp, L.J., Schmittmann, V.D. & Borsboom, D. (2012). Qgraph: Network visualizations of relationships in psychometric data. *Journal of Statistical Software*, 48, 1-18.
- Gould, S. J. (1980). *The panda's thumb*. New York: Norton.
- Kossakowski, J.J., Epskamp, S., Kieffer, J.M., Borkulo, C.D. van, Rhemtulla, M. & Borsboom, D. (2015). The application of a network approach to health-related quality of life (HRQoL): Introducing a new method for assessing HRQoL in healthy adults and cancer patients. *Quality of Life Research*. doi: 10.1007/s11136-015-1127-z
- Kroeze, R., Van der Veen, D. C., Servaas, M. N., Bastiaansen, J. A., Oude Voshaar, R., Borsboom, D., Ruhe, H. G., Schoevers, R. A. & Riese, H. (submitted). Personalized feedback on daily dynamics of psychopathology: A proof-of-principle study.
- McNally, R. J., Robinaugh, D. J., Wu, Gwyneth W. Y., Wang, L., Deserno, M. & Borsboom, D. (2015). Mental disorders as causal systems: A network approach to Posttraumatic Stress Disorder. *Clinical Psychological Science*, 3, 1-14.
- Robinaugh, D. J., LeBlanc, N.J., Vulitech, H.A. & McNally, R.J. (2014). Network analysis of persistent complex bereavement disorder in conjugally bereaved adults. *Journal of Abnormal Psychology*, 123, 510-522.
- Ruzzano, L., Borsboom, D. & Geurts, H.M. (2014). A network analysis of repetitive behaviors in autism and obsessive compulsive disorder: a new perspective. *Journal of Autism and Developmental Disorders*. DOI: 10.1007/s10803-014-2204-9
- Skinner, B.F. (1987). Whatever happened to psychology as the science of behavior? *American Psychologist*, 42, 780-186.
- Spearman, C. (1904). "General intelligence", objectively determined and measured. *The American Journal of Psychology*, 15, 201-292.
- Van de Leemput, I.A., Wichers, M., Cramer, A.O.J., Borsboom, D., Tuerlinckx, F., Kuppens, P., Nes, E.H. van, Viechtbauer, W., Giltay, E.J., Aggen, S.H., Derom, C., Jacobs, N., Kendler, K.S., Maas, H.L.J. van der, Neale, M.C., Peeters, F., Thiery, E., Zachar, P. & Scheffer, M. (2014). Critical slowing down as early warning for the onset and termination of depression. *National Academy of Sciences. Proceedings*, 111 (1), 87-92.
- Wichers, M., Groot, P. C., Psychosystems, ESM Group & EWS Group (in press). Critical Slowing Down as a personalized Early Warning Signal for depression. *Psychotherapy and Psychosomatics*.