

Denny Borsboom

Faculteit der Maatschappij- en Gedragwetenschappen
Universiteit van Amsterdam
d.borsboom@uva.nl

Maarten Marsman

Faculteit der Maatschappij- en Gedragwetenschappen
Universiteit van Amsterdam
m.marsman@uva.nl

Latente variabelen en netwerken: verschillende benaderingen van psychometrische data

Psychometrische gegevens spelen een steeds belangrijker rol in onze maatschappij. Van IQ-score tot Citotoets en van persoonlijkheidstest tot psychodiagnostiek; iedereen krijgt ergens in zijn of haar leven wel eens met een psychologische test te maken. De statistische analyse van testcores, die op deze manier verzameld worden, is het domein van de psychometrie. De centrale taakstelling van de psychometrie is het ontwikkelen van modellen die hiervoor optimaal geschikt zijn. In dit artikel geven Denny Borsboom en Maarten Marsman een overzicht van de belangrijkste ideeën en wiskundige representaties die geassocieerd zijn met netwerkbenaderingen van de psychometrie, en zij illustreren de hieruit voortkomende technieken met empirische data. Daarnaast bespreken zij openstaande kwesties en te verwachten ontwikkelingen.

Hoe men de psychometrische analyse van testgegevens aanpakt, en welk model men daarvoor kiest, hangt deels af van de vraag hoe de relatie tussen die gegevens en de eigenschap die men wil bestuderen in elkaar steekt. Daar komen naast statistische ook inhoudelijke en wetenschapsfilosofische vragen bij kijken. Wie denkt dat depressie een meetbare eigenschap is zal de data anders analyseren dan iemand die denkt dat het maar een naam is voor een groepje symptomen. In dit stuk zullen we twee manieren bespreken om over deze kwestie na te denken. In de klassieke opvatting zijn psychometrische gegevens (bijvoorbeeld IQ-scores, symptomen van depressie, observaties van menselijk gedrag in verschillende situaties) metingen van een ‘onderliggende’ eigenschap (intelligentie, depressie, persoonlijkheid). Gedurende de afgelopen vijftien jaar is daar ook een nieuwe opvatting tegenover geplaatst, namelijk dat de in psychologische tests uitgevraagde eigenschappen beter gezien

kunnen worden als complexe systemen waarin allerlei verschillende gevoelens, gedachten, en gedragingen interacteren. Deze gedachte heeft aanleiding gegeven tot het ontstaan van de netwerkbenadering van de psychometrie.

Latente variabelen

In de psychometrie was lange tijd het belangrijkste statistische model het latente-variabelenmodel. Dit model is gebaseerd op de constatering dat we de eigenschappen waarin we in de psychologie geïnteresseerd zijn (zoals bijvoorbeeld depressie) niet direct kunnen waarnemen. Voor het bepalen van de mate waarin iemand depressief is zijn we aangewezen op variabelen die wel min of meer direct waarneembaar zijn; bijvoorbeeld of iemand slaapgebrek heeft, zich niet kan concentreren, zegt zich neerslachtig te voelen, of zelfs geprobeerd heeft zich te suïcideren. In de psychometrie wordt dit type variabelen gezien als *manifest*, terwijl de theoretische

eigenschap die men via zulke variabelen wil bepalen als *latent* [3, 8, 15].

De hoofdstrategie in de psychometrie was gedurende de twintigste eeuw om een wiskundig model aan te leggen dat het functionele verband tussen manifeste en latente variabelen specificceert, en met dit model manieren te vinden om op basis van de waargenomen scores op manifeste variabelen tot een schatting van iemands waarde op de latente variabele te komen. Zo’n schatting kan worden gezien als een (indirecte) meting van depressie, en zodoende wordt het latente-variabelenmodel vaak een *meetmodel* genoemd. (Het idee dat via dit type modellen inderdaad — kwantitatieve — metingen kunnen worden gerealiseerd is overigens niet onomstreden. Zie Michell [28, 29] voor een aanval op dit idee, en Borsboom [6] voor een verdediging ervan.) Wiskundig gezien komt dit latente-variabelenmodel neer op een factorisatie van de gemeenschappelijke kansverdeling van de manifeste variabelen. In het geval van een aantal binaire manifeste variabelen en een continue latente variabele is deze factorisatie bijvoorbeeld van de vorm

$$P(\mathbf{x}) = \int_{-\infty}^{\infty} P(\mathbf{x} | \theta) f(\theta) d\theta,$$

waarin \mathbf{x} een vector van realisaties is van een set binaire random variabelen \mathbf{X} , θ een continue latente variabele weergeeft, $P(\mathbf{x} | \theta)$ de kansfunctie geeft die de

kans op de waargenomen data \mathbf{x} aan θ relateert, en $f(\theta)$ de dichtheid van de latente variabele weergeeft. De functie $f(\theta)$ representeert de verdeling van de gemeten latente variabele in de populatie. De functie $P(\mathbf{x}|\theta)$ kent voor iedere waarde van de gemeten eigenschap θ een specifieke kans toe aan alle mogelijke responspatronen \mathbf{x} . In het geval van depressie zou deze functie bijvoorbeeld voor ieder niveau van depressie de kans bepalen dat iemand de symptomen neerslachtigheid, vermoeidheid, en concentratieproblemen al dan niet zou hebben.

De vorm van de kansfunctie $P(\mathbf{x}|\theta)$ wordt bepaald door de zogenaamde Item Respons Functie (IRF). De IRF kan in principe alle vormen aannemen [27], maar meestal wordt de voorkeur gegeven aan een logistische functie. Deze functie specificeert bijvoorbeeld dat de kans dat persoon i met waarde θ_i op de latente variabele een item goed maakt (in het geval van vaardigheidstests), het betreffende gedrag laat zien (in het geval van persoonlijkheidstest), of het symptoom in kwestie heeft (in het geval van psychopathologie), de volgende vorm heeft:

$$P(X_{ij} = 1|\theta_i) = \frac{e^{\alpha_j(\theta_i - \beta_j)}}{1 + e^{\alpha_j(\theta_i - \beta_j)}}.$$

In deze kansfunctie wordt de kans dat de manifeste variabele X_{ij} (de respons van persoon j op item i) de waarde '1' aanneemt dus gemodelleerd als een functie van een latente eigenschap van de persoon (θ_i) en twee eigenschappen van het item (een interviewvraag, opgave, of vragenlijst-item): een locatieparameter β_j en een discriminatieparameter α_j . De locatieparameter geeft aan hoe hoog de waarde op de latente trek moet zijn om een kans van $\frac{1}{2}$ op een waarde 1 te genereren. Deze wordt bij het meten van cognitieve vaardigheden bijvoorbeeld geïnterpreteerd als de moeilijkheid van het item, en bij het meten van psychopathologie als de ernst van het symptoom. De parameter α_j is de richtingscoëfficiënt van de IRF geëvalueerd op het punt $\theta_i = \beta_j$. Deze parameter geeft aan hoe sterk het item discrimineert tussen verschillende waarden van θ . In de meeste testtheoretische toepassingen streeft men naar een goede spreiding van de waardes van β en een zo hoog mogelijke waarde van α .

Aan de keuze voor de logistische functie liggen principiële en praktische redenen

ten grondslag. Een belangrijke principiële reden is dat bepaalde specifieke invullingen van het zo ontstane kansmodel zeer aantrekkelijke meeteigenschappen opleveren. Indien de discriminatieparameters allen gelijk zijn ontstaat het zogenaamde Rasch-model, vernoemd naar de Deense wiskundige Georg Rasch [33]. Onder dit model is de totaalscore van een voldoende statistiek voor de latente variabele en daardoor kunnen de itemparameters bepaald worden onafhankelijk van de vorm van de dichtheid $f(\theta)$. Omdat de latente variabele latent is, hebben we zelden sterke a priori redenen om verdelingsaannames te motiveren, en daarom is dit een aantrekkelijke eigenschap.

In het bovenstaande voorbeeld zijn de manifeste variabelen dichotoom, en is de latente variabele continu, maar er bestaan soortgelijke modellen voor allerlei soorten en maten manifeste en latente variabelen [27]. Hoewel het latente-variabelenmodel een psychologische uitvinding is [41] wordt het tegenwoordig ook steeds belangrijker in velden als *machine learning* en bij modellen die gebruik maken van neurale netwerken als in *deep learning* (ook het neurale netwerkmodel vindt overigens zijn oorsprong in de psychologie [21]).

Tegenwoordig is het gemakkelijk het latente-variabelenmodel te schatten op basis van data. Dit levert voor iedere manifeste variabele een discriminatie- en locatieparameter op, en voor iedere persoon een schatting van de positie van die persoon op de latente variabele. Een voorbeeld van een model, geschat met het *ltm*-pakket [35] in het *open source* statistisch softwarepakket R [32] is weergegeven in Figuur 1. De analyse is uitgevoerd op een dataset van 8973 personen in de zogenaamde VATSPSUD-studie [23].

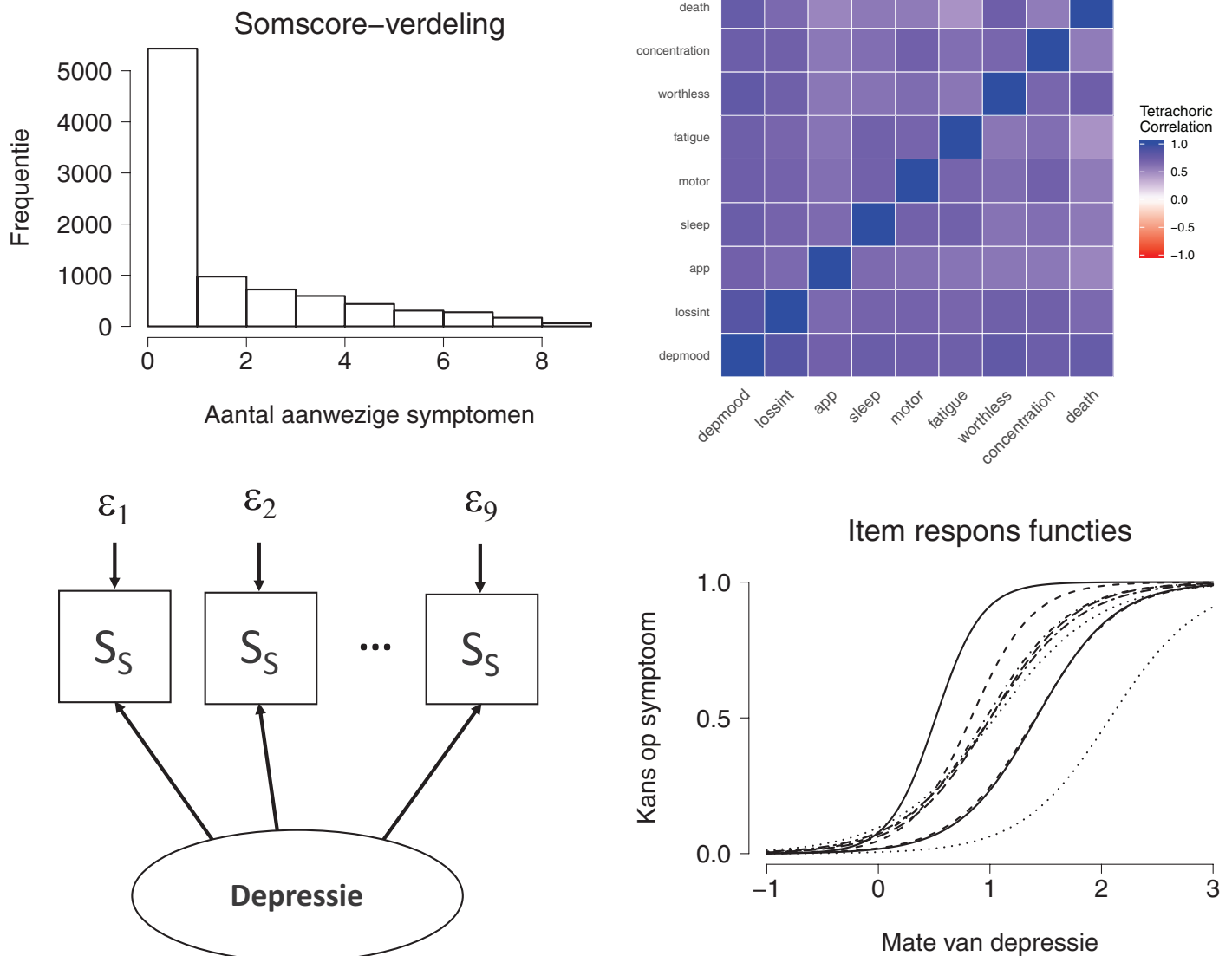
De verdeling van depressiesymptomen is scheef naar rechts, en de items discrimineren dus alleen aan de bovenkant van de populatie. Dit is ook wat men verwacht van een diagnostisch instrument. De symptomen vertonen verder de *positive manifold*: alle correlaties zijn positief. De *positive manifold* is een interessante eigenschap van psychometrische data die zich ook voordoet bij cognitieve vaardigheden (alle vaardigheidstests correleren in de algemene populatie positief), persoonlijkheidstests (vrijwel alle persoonlijkheidseigenschappen correleren positief als men deze zo codeert, dat een hogere score een

wenselijker gedragsprofiel indiceert) en psychopathologie in het algemeen (symptomen correleren bijna altijd positief, met uitzondering van enkele symptomen die met manie te maken hebben — deze correleren negatief met de inactiviteit die bij typische depressies hoort). Dit robuuste fenomeen, dat honderden keren gerepliceerd is, vormt een van de belangrijkste ontdekkingen van de psychometrie.

Netwerken

Het meetmodel dat hierboven is uiteengelegd is van ontzagwekkend belang voor de organisatie van onderzoek in de psychopathologie. Als het model klopt, dan kunnen we namelijk de telling van symptomen gebruiken als *meting* van een theoretisch relevante variabele als depressie, net zoals we andere tellingen kunnen gebruiken als metingen van angststoornissen, fobieën, schizofrenie, et cetera. Inderdaad is het onderzoek langs deze psychometrische lijnen georganiseerd. Als men bijvoorbeeld leest over de neurale grondslag voor depressie, dan gaat het om vergelijkingen van hersendata tussen mensen die hoog of laag op deze telling scoren. Therapeutisch effectonderzoek richt zich op de relatieve afname van de gemiddelde totaalscore als functie van een controle dan wel experimentele conditie. De erfelijkheid van stoornissen (die overigens zeer substantieel is) wordt bepaald door sterkte van de correlaties tussen symptoomtellingen te vergelijken over paren van één- en twee-eiige tweelingen. De zoektocht naar de genetische basis van stoornissen is gebaseerd op het regresseren van symptoomtellingen op variaties in basenparen op het DNA.

Maar de conclusie dat wij de mate van de depressie kunnen meten door de toepassing van een statistisch model is niet gratis. Wij moeten haar 'kopen' met zware aannames over de processen die ten grondslag liggen aan de waargenomen symptomatologie. De belangrijkste daarvan is de aanname dat de symptomen bepaald worden door één en de dezelfde onderliggende grootte — de latente variabele. Zoals weergegeven in Figuur 1 betekent dit dat aan stoornis als depressie de *gemeenschappelijke oorzaak* van de symptomen is, net zoals bijvoorbeeld een tumor in de longen de gemeenschappelijke oorzaak van symptomen als pijn op de borst, hoesten, en bloed opgeven kan zijn.



Figuur 1 Het latente-variabelenmodel toegepast op depressie. Linksboven de verdeling van de somscore (het aantal aanwezige symptomen in de data), rechtsboven een heatmap van de tetrachorische correlaties (allen positief), linksonder de structuur van het latente-variabelenmodel (waarin symptomen S_1 – S_9 gerelateerd worden aan de latente variabele ‘depressie’), en tot slot, rechtsonder, de Item Respons Functies die de bijbehorende modelschatting oplevert.

Deze structuur verraadt zich door de wijze waarop de kansverdeling gefactoriseerd wordt. De factorisatie heeft de vorm van een product en berust dus op de aanname dat de manifeste variabelen statistisch onafhankelijk zijn, gegeven de latente variabele. Deze aanname staat bekend als *lokale onafhankelijkheid*. Het model zegt in wezen dat het statistisch verband tussen symptomen als vermoeidheid en insomnie in essentie ‘spurieuus’ is. Net zoals de correlatie tussen het aantal ooiervaars en het aantal geboortes in Macedonische dorpjes berust op hun gemeenschappelijke afhankelijkheid van het inwonertal (meer babies ← meer stelletjes ← meer inwoners → meer schoorstenen → meer ooiervaars), zo berust de correlatie tussen vermoeidheid en

insomnia op hun gemeenschappelijke afhankelijkheid van de mate van depressie. Wanneer men voor deze gemeenschappelijke oorzaak controleert, dan vallen de correlaties tussen de manifeste variabelen weg, net zoals de correlatie tussen het aantal ooiervaars en het aantal geboortes wegvalt als men voor de populatiegrootte van de dorpjes controleert. Samen met de aanname dat de kansverdeling van iedere manifeste variabele monotoon is in de latente variabele — hoe depressiever, hoe groter de kans op ieder van de symptomen — levert dit een model op waarin de manifeste variabelen tot op zekere hoogte uitwisselbaar zijn [22]

Dit is weliswaar een mooi statistisch model, maar inhoudelijk gezien is dit mo-

del niet altijd even voor de hand liggend. Een goed voorbeeld van waar het misgaat is het domein van de psychopathologie. De symptomen van psychologische stoornissen lijken op het eerste gezicht immers geen plausibele kandidaten voor uitwisselbaarheid (bij depressie is het bijvoorbeeld lastig vol te houden dat symptomen als suicidaliteit en gebrek aan eetlust uitwisselbaar zijn). Bovendien zijn die symptomen niet zuiver en alleen gecorreleerd vanwege een gemeenschappelijke afhankelijkheid van latente variabelen. De correlatie tussen insomnie en vermoeidheid berust bijvoorbeeld naar alle waarschijnlijkheid wel op een direct verband tussen deze symptomen: wie niet slaapt, wordt moe. Hetzelfde geldt voor allerlei andere symptomen die

we aantreffen in almanakken als de DSM (het belangrijkste diagnostisch handboek in de psychiatrie [1]). Hallucinaties leiden tot angsten, drugsgebruik tot tolerantie en onthoudingsverschijnselen, gebrekkige energie tot zelfverwilt, de overtuiging dik te zijn tot excessief afvallen, angst tot vermijding van datgene waar men bang voor is, enzovoort. Symptomen van psychische stoornissen zijn dus hoogstwaarschijnlijk betrokken in allerlei directe causale relaties. Vaak worden symptomen daarbij ook nog gekenmerkt door *feedback*-relaties: angst voor spinnen leidt tot vermijding van spinnen, wat weer leidt tot instandhouding van de angst; wie overtuigd is dat de CIA achter hem aan zit ziet overal camera's, wat de betreffende waan in stand houdt; de gokker die geld verliest aan de gokautoomaat zal vaak dat geld terug proberen te verdienen met gokken, wat weer tot meer schulden leidt.

Het beeld dat hieruit oprijst is dat van symptomen die samenklonteren in allerlei patronen van al dan niet wederzijdse bekrachtiging. Dat staat op gespannen voet met de causale interpretatie van het latente-variabelenmodel. In het onderzoeksprogramma, dat we in 2007 gestart zijn aan de Universiteit van Amsterdam, besloten we daarom de zaak om te draaien. In plaats van aan te nemen dat alle associaties tussen symptomen van stoornissen ontstaan doordat die symptomen veroor-

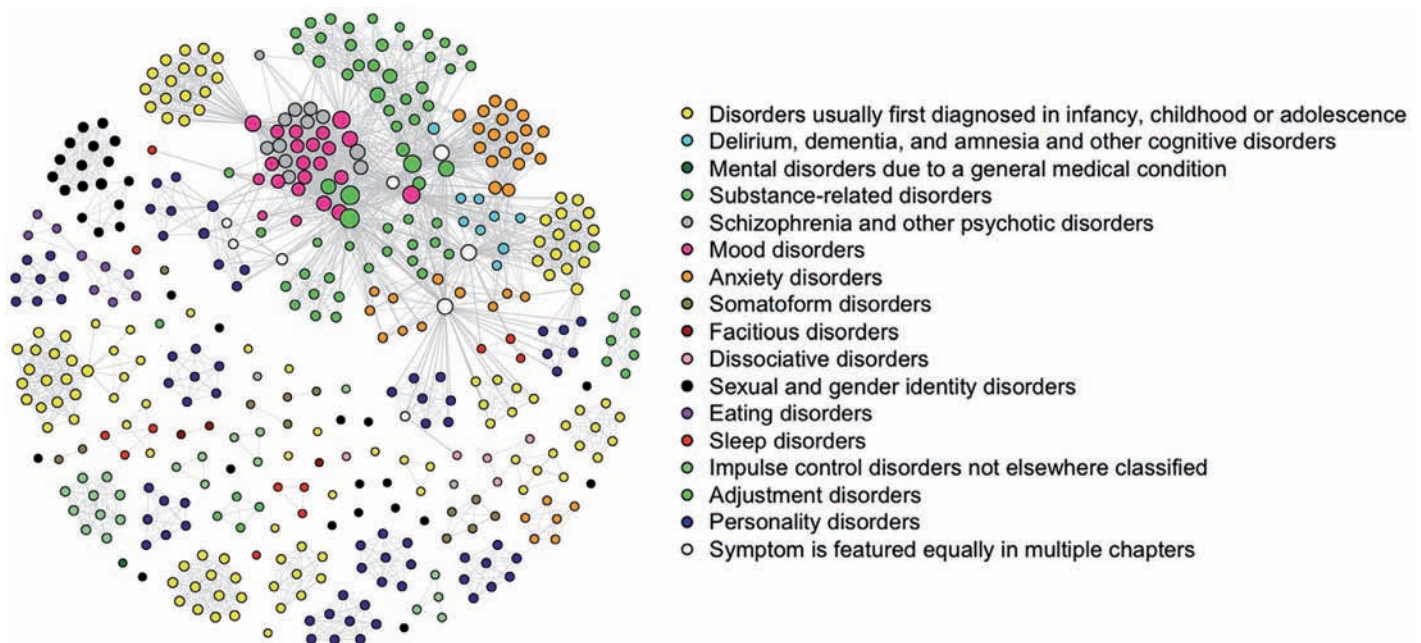
zaakt worden door dezelfde stoornis, gingen wij uit van een situatie waarin symptomen eigenstandige causale entiteiten zijn, die netwerken vormen van interacties.

In een van onze eerste studies [7] besloten we als vertrekpunt de *Diagnostic and Statistical Manual of Mental Disorders* [1] te nemen. Het leek ons plausibel dat symptomen, die in de loop van de eeuw in dezelfde stoornis waren gecategoriseerd, sterkere verbindingen in het netwerk zouden moeten hebben. Om een eerste zicht te krijgen op de structuur van het netwerk van psychopathologie als geheel produceerden we een afbeelding van de DSM, waarin ieder symptoom is weergegeven als een knoop, en twee symptomen verbonden zijn als ze als diagnostische criteria voor dezelfde stoornis in de DSM figureren (dit is dus de projectie van een *bipartite* netwerk). Dit netwerk is weergegeven in Figuur 2. Interessant is dat de graaf een zogenaamde *giant component* laat zien (bijna 50% van de symptomen is verbonden).

Gegeven de netwerkhypothese levert dit patroon eigenlijk een 'gratis' verklaring voor de enorme comorbiditeit tussen mentale stoornissen. Een groot deel van de mensen die aan de diagnostische criteria voldoen voor een stoornis (bijvoorbeeld depressie), voldoen óók aan de criteria voor een andere stoornis (bijvoorbeeld de paniekstoornis), en dit patroon

doet zich voor bij vrijwel alle stoornissen in de DSM. Dit wordt vaak gezien als een meetprobleem, maar in de netwerktheorie is dat een verkeerde opvatting. Als stoornissen ontstaan door interacties tussen symptomen, dan houden die interacties natuurlijk niet op bij de grenzen van een diagnostische categorie. Wanneer iemand bijvoorbeeld niet van de drank af kan blijven, en daardoor ontslagen wordt en ruzie krijgt met zijn familie (symptomen van alcoholmisbruik), dan kan die persoon zich daardoor waardeloos gaan voelen (symptoom van depressie) of zich excessief zorgen gaan maken (symptoom van gegeneraliseerde angststoornis). Zo kan comorbiditeit ontstaan via causale paden die symptomen van verschillende stoornissen verbinden. De gemiddelde kortste padlengte tussen de symptomen van twee stoornissen in het netwerk correleren dan ook sterk negatief ($r = -0,7$) met de empirisch bepaalde mate van comorbiditeit tussen de betreffende stoornissen, zelfs als daarbij gemeenschappelijke symptomen niet worden meegenomen.

Met behulp van een netwerkstructuur als die in Figuur 2 kan men simulatiemodellen opstellen om het ontstaan en verloop van stoornissen na te bootsen. In deze modellen, die verwant zijn aan epidemiologische modellen voor virussen, kunnen actieve symptomen hun burens 'aansteken'. Zo kan de symptoomactivatie



Figuur 2 Het DSM-IV-netwerk op basis van de structuur van het diagnostische systeem, waarin twee symptomen verbonden zijn als ze bij dezelfde stoornis zijn ingedeeld (boven) en voor een deel van de symptomen op basis van empirische data, waarin twee symptomen verbonden zijn als ze conditioneel geassocieerd zijn, gegeven de andere symptomen.

zich verspreiden over het netwerk. Zulke simulaties laten zien dat dit proces geloofwaardige eigenschappen heeft: het leidt ertoe dat symptoomactiviteit zich zal concentreren rond hecht verbonden groepjes symptomen ('stoornissen') die evenwel niet goed afgescheiden zijn, waardoor symptoomactiviteit van de ene naar de andere stoornis over kan springen ('comorbiditeit') via symptomen die met beide groepjes symptomen verbonden zijn (deze noemen we 'brugsymptomen' [14]). Bovendien kan zulke symptoomactiviteit er bij hecht verbonden netwerken toe leiden dat het netwerk blijft plakken in een patroon van zichzelf in stand houdende symptomen, omdat er *hysteresis* optreedt: het netwerk heeft niet-lineaire eigenschappen, in de zin dat een kleine perturbatie kan leiden tot een systematische symptoomactivatie, die dan niet uit zichzelf meer kan verdwijnen. Het netwerk heeft dus een *tipping point*, zoals dat ook bekend is bij allerlei andere complexe dynamische systemen [38, 39]

Onder de versimpelde aanname dat symptomen zich inderdaad uitsluitend organiseren langs de lijnen van paarsgewijze interacties in een netwerk, kan men de structuur en parameters van zo'n netwerk ook schatten. Technisch gezien is een systeem van binaire variabelen die paarsgewijs geassocieerd zijn statistisch equivalent met het bekende Ising-model uit de natuurkunde [19, 26]. De kans dat we het systeem in de configuratie \mathbf{x} aantreffen wordt in het Ising-model gekarakteriseerd door de formule

$$p(\mathbf{x}) = \frac{\exp(\mathbf{x}^T \boldsymbol{\mu} + \mathbf{x}^T \boldsymbol{\Sigma} \mathbf{x})}{\sum_{\mathbf{x}} \exp(\mathbf{x}^T \boldsymbol{\mu} + \mathbf{x}^T \boldsymbol{\Sigma} \mathbf{x})},$$

waarbij de som in de noemer over alle mogelijke configuraties van de vector \mathbf{x} loopt. De hoofdeffecten $\boldsymbol{\mu}$ bepalen de kans dat knopen in het netwerk geactiveerd zijn, buiten de invloed van de andere knopen in het netwerk om; hogere waarden leiden tot een hogere kans tot activatie. De invloed van de andere knopen worden uitgedrukt in de matrix met paarsgewijze interacties, $\boldsymbol{\Sigma}$, waarbij een hogere waarde van een interactie-parameter σ_{ij} leidt tot een hogere kans dat een paar variabelen X_i en X_j beide geactiveerd zijn. Deze interacties worden in de psychometrie dus gebruikt om, bijvoorbeeld, de wederkerige interacties tussen symptomen te kenmerken.

De structuur van het Ising-model in termen van hoofdeffecten en paarsgewijze interacties lijkt in veel opzichten op dat van een andere bekende multivariate verdeling, de multivariate normale verdeling. Maar waar het doorgaans eenvoudig rekenen is met de normale verdeling, is het Ising-model notoir lastig. De som in de noemer, bijvoorbeeld, kan slechts in uitzonderlijke gevallen worden versimpeld, wat er in de praktijk op neer komt dat alle 2^p configuraties geëvalueerd dienen te worden, en met slechts $p = 20$ variabelen zijn dit meer dan een miljoen termen. Dit is extra kostbaar voor de iteratieve procedures die gebruikt worden om de waarden van modelparameters te bepalen uit geobserveerde data. Om deze reden zijn er allerlei trucs bedacht om de parameters te schatten zonder het model in zijn volledigheid te hoeven evalueren.

Een belangrijke schattingsmethode gebruikt de zogeheten pseudo-likelihood. Hierbij optimaliseert men niet de parameters in de gezamenlijk verdeling, $P(\mathbf{X} = \mathbf{x})$, maar optimaliseert men de conditionele verdeling van één variabele, zeg X_i , gegeven de rest van de variabelen in het netwerk. De kans dat we variabele i in de toestand $x_i = 1$ observeren gegeven de overige variabelen $\mathbf{x}^{(i)}$ wordt gegeven door de formule

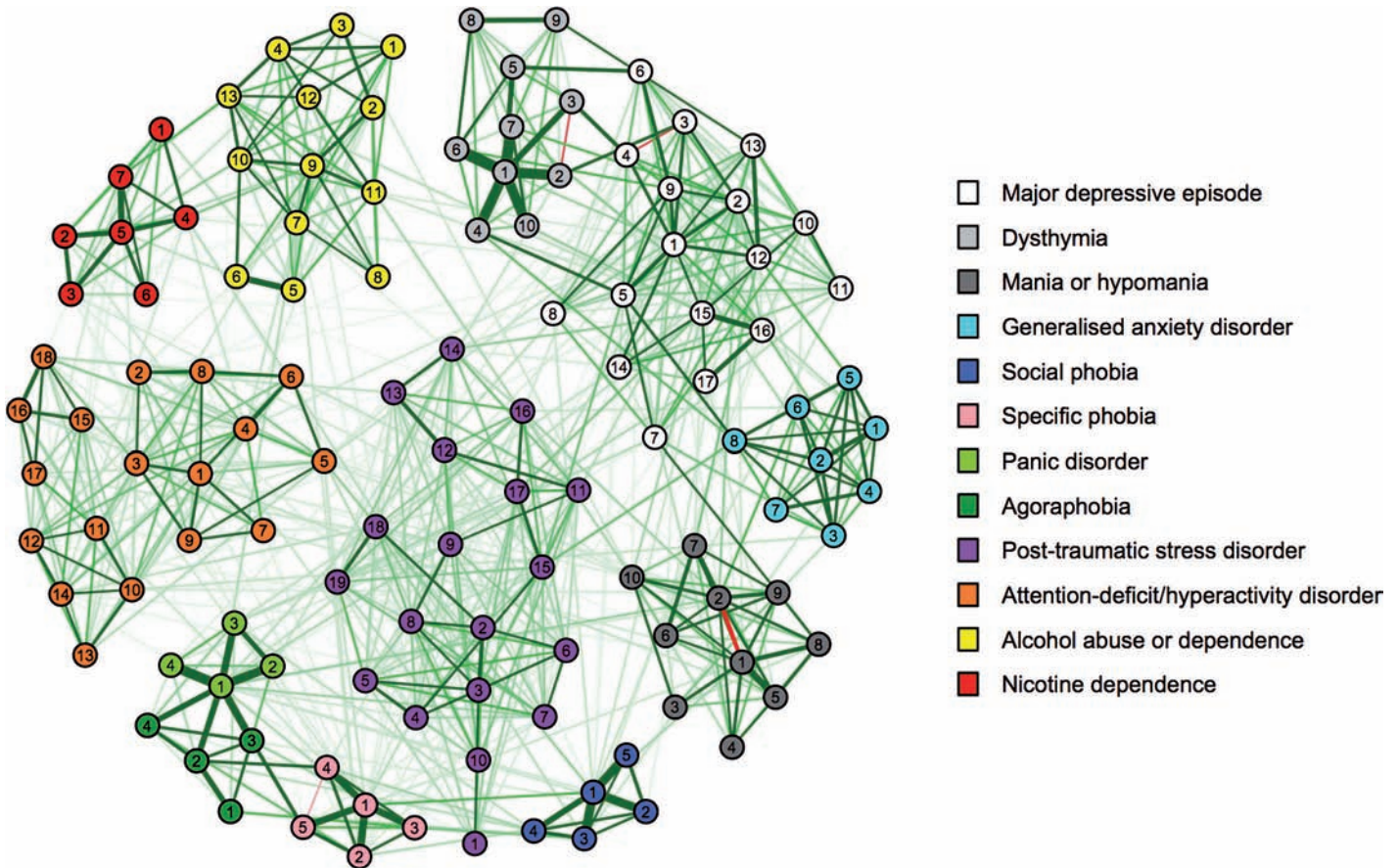
$$P(X_i = 1 | \mathbf{x}^{(i)}) = \frac{\exp(\mu_i + 2 \sum_{j \neq i} x_j \sigma_{ij})}{1 + \exp(\mu_i + 2 \sum_{j \neq i} x_j \sigma_{ij})},$$

ofwel de regressie van X_i op $\mathbf{x}^{(i)}$. Hier zien we ook direct hoe de interacties gemodelleerd worden in het Ising-model; als $\sigma_{ij} = 0$ dan zijn de variabelen i en j onafhankelijk van elkaar gegeven de rest van het netwerk.

De afgelopen jaren zijn schattingsmethoden waarmee deze netwerken geschat kunnen worden [34] door onze groep in Amsterdam verder ontwikkeld en geïmplementeerd in het vrij beschikbare softwareprogramma R. Een voorbeeld van een netwerk dat daarmee is geschat op basis van empirische gegevens is weergegeven in Figuur 3. In dit netwerk wordt ieder symptoom weergegeven als een knoop, en zijn twee knopen verbonden als het model met die verbinding beter op de data fit, dan het model zonder de verbinding [4]. Dit betekent in de praktijk dat variabelen verbonden raken als het verband tussen die variabelen niet 'wegverklaard' kan worden

door voor de andere variabelen in het netwerk te controleren. De schattingsmethode benadert zo het Markov Random Field [24], waarin twee niet direct verbonden variabelen onafhankelijk zijn, gegeven hun gemeenschappelijke burens. Deze structuur van conditionele afhankelijkheden is nauw gerelateerd aan de moderne causaliteitsopvattingen van bijvoorbeeld Pearl [31]. Het is overigens belangrijk om op te merken dat een verbinding tussen variabelen in een netwerk niet betekent dat aangetoond is dat die variabelen elkaar beïnvloeden; een verbinding kan ook ontstaan door gemeenschappelijke afhankelijkheid van een latente variabele of door conditioneren op een gemeenschappelijk effect [16]. Netwerkmethodologie zoals in Figuur 3 weergegeven moet men dus in eerste instantie zien als een hypothese-genererende techniek waarmee inzicht in de structuur van de data kan worden gegeven, en niet zozeer als een test van de netwerkhypothese zelf.

Toen de eerste netwerkrepresentaties van psychometrische variabelen beschikbaar werden via het programma *qgraph* van Sacha Epskamp [18] en in het verlengde daarvan schattingsmethoden op basis van geregulariseerde regressie geïmplementeerd werden door Claudia van Borkulo [4] was dat eigenlijk meteen een enorm succes. Ook onafhankelijk van de onderliggende theorie geeft een netwerk als in Figuur 3 een inzichtelijke visualisatie van complexe multivariate structuren in datasets, die er ook nog aantrekkelijk uitziet. Toen wij deze methodologie introduceerden leidde dat dan ook tot een explosie van toepassingen in allerlei velden van de psychologie en daarbuiten. In het kielzog van die populariteit hebben we de netwerkmethodologie in allerlei richtingen verder ontwikkeld. De huidige programmatuur is in staat netwerken te schatten voor binaire [4], continue [17], en gemengde variabelen [20], zowel op basis van data die bij een groot aantal mensen eenmalig verzameld zijn (zogenaamde cross-sectionele data) als op basis van herhaalde metingen [13, 16]. Bovendien hebben we systematische methodologie ontwikkeld om de stabiliteit en robuustheid van de geschatte netwerken te bepalen [17]. Op dit moment zijn de meeste standaard datasets hierdoor eenvoudig om te zetten in een netwerkrepresentatie.



Figuur 3 Een empirisch geschat Ising-model voor de DSM-IV-symptomatie geschat op basis van een grote steekproef van mensen bij wie de aan- dan wel afwezigheid van de symptomen vastgesteld is [9]. Het model is geschat op basis van geregulariseerde regressiemodellen [4].

Conclusie

In dit artikel hebben we de belangrijkste uitgangspunten en eigenschappen uiteengezet van twee benaderingen in de psychometrie: de klassieke benadering, waarin manifeste variabelen worden gezien als metingen van een onderliggend theoretisch construct, en de netwerkbenadering, waarin manifeste variabelen worden gezien als onderdeel van een gecompliceerd systeem waarin gevoelens, gedachten, en gedragingen elkaar beïnvloeden. Het latente-variabelenmodel is de uitkomst van een eeuw aan psychometrisch onderzoek, en is statistisch gezien grotendeels af, terwijl de jeugdiger netwerkbenadering nog maar net op stoom is en stormachtige ontwikkelingen doormaakt.

Het is de vraag waar die ontwikkelingen ons zullen brengen. Nu het mogelijk is de topologie van psychometrische netwerken te bepalen op basis van data, ontstaat de vraag wat die topologie ons eigenlijk oplevert. Zijn meer centrale variabelen in een netwerk bijvoorbeeld op

een of andere manier belangrijker? Is de netwerkstructuur bij mensen met psychopathologie anders dan die van de algemene populatie? Kunnen netwerkstructuren helpen bij de keuze waar in het systeem moet worden geïntervenieerd? Er lijken inderdaad aanwijzingen te zijn dat de topologie van netwerken ertoe doet. De positie van knopen in psychopathologie-netwerken blijkt geassocieerd met hun voorspellend vermogen: hoe centraler de knoop, hoe beter deze de toekomstige ontwikkeling van de stoornis voorspelt [10,37]. Daarnaast zijn groepsverschillen in netwerkstructuren in verband gebracht met relevante externe variabelen. Zo lijken mensen die blijven hangen in een depressie een dichter verbonden structuur te hebben dan mensen die herstellen [5,40] en zijn de negatieve emoties van mensen met verhoogde kans op depressie sterker met elkaar verbonden [12,30]. Sommige experimentele interventies lijken op sommige variabelen in het netwerk sterker aan te grijpen dan op andere, wat zich kan geven op de nog grotendeels

onbekende werkingsmechanismen van psychotherapie [2]. Karakteristieke patronen die in verband zijn gebracht met fase-overgangen in netwerken en hysteresis, zoals het vertragen van het systeem voorafgaand aan een fase-overgang [39] worden ook gezien in psychopathologie-onderzoek [42]. En de eerste therapeutische interventietechnieken op basis van netwerkanalyse worden inmiddels onderzocht [25].

De ontwikkelingen gaan in feite zo snel, dat er inmiddels ook allerlei waarschuwingen in de literatuur opduiken die kanttekeningen plaatsen bij de validiteit van conclusies die op basis van netwerkanalyse worden getrokken. Zo zijn centraliteitsmaten heel populair, terwijl niet altijd even duidelijk is wat deze meten [11,16]. Daarnaast zijn sommige netwerk-schattingen gevoelig voor steekproeffluctuaties [17], en het effect daarvan is niet altijd even gemakkelijk te bepalen. Zulke vragen zijn natuurlijk zeer belangrijk en zullen de komende jaren ongetwijfeld veel aandacht krijgen.

Referenties

- 1 American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders*, 5th ed., APA, 2013.
- 2 T.F. Blanken, T. van der Zweerde, A. van Straten, E.J.W. van Someren, D. Borsboom en J. Lancee, Introducing network intervention analysis to investigate sequential, symptom-specific treatment effects: A demonstration in co-occurring insomnia and depression, *Psychotherapy and Psychosomatics* 88 (2019), 52–54.
- 3 K.A. Bollen, Latent variables in psychology and the social sciences, *Annual Review of Psychology* 53 (2002), 605–634.
- 4 C.D. van Borkulo, D. Borsboom, S. Epskamp, T.F. Blanken, L. Boschloo, R.A. Schoevers en L.J. Waldorp, A new method for constructing networks from binary data, *Scientific Reports* 4 (2014), 5918.
- 5 C.D. van Borkulo, L. Boschloo, D. Borsboom, B.W.J.H. Penninx, L.J. Waldorp en R.A. Schoevers, Association of symptom network structure with the course of longitudinal depression, *JAMA Psychiatry* 72 (2015), 1219–1226.
- 6 D. Borsboom, *Measuring the Mind: Conceptual Issues in Contemporary Psychometrics*, Cambridge University Press, 2005.
- 7 D. Borsboom, A.O.J. Cramer, V.D. Schmittmann, S. Epskamp en L.J. Waldorp, The small world of psychopathology, *PLoS ONE* 11 (2011), e27407.
- 8 D. Borsboom, G.J. Mellenbergh en J. van Heerden, The theoretical status of latent variables, *Psychological Review* 110 (2003), 203–219.
- 9 L. Boschloo, C.D. van Borkulo, M. Rhemtulla, K.M. Keyes, D. Borsboom en R.A. Schoevers, The network structure of symptoms of the Diagnostic and Statistical Manual of Mental Disorders, *PLoS One* 10 (2015), e0137621.
- 10 L. Boschloo, C.D. van Borkulo, D. Borsboom en R.A. Schoevers, Letter to the editor: A prospective study on how symptoms in a network predict the onset of depression, *Psychotherapy and Psychosomatics* 85 (2016), 183–184.
- 11 L.F. Bringmann, T. Elmer, S. Epskamp, R.W. Krause, D. Schoch, M. Wichers, J.T.W. Wigman en E. Snippe, What do centrality measures measure in psychological networks? *Journal of Abnormal Psychology* (2019), in press.
- 12 L.F. Bringmann, M.L. Pe, N. Vissers, E. Ceulemans, D. Borsboom, W. Vanpaemel, F. Tuerlinckx en P. Kuppens, Assessing temporal emotion dynamics using networks, *Assessment* 23 (2016), 425–435.
- 13 L.F. Bringmann, N. Vissers, M. Wichers, N. Geschwind, P. Kuppens, F. Peeters, D. Borsboom en F. Tuerlinckx, A network approach to psychopathology: New insights into clinical longitudinal data, *PLoS ONE* 8 (2013), e60188.
- 14 A.O.J. Cramer, L.J. Waldorp, H.L.J. van der Maas en D. Borsboom, Comorbidity: A network perspective, *Behavioral and Brain Sciences* 33 (2010), 137–150.
- 15 L.J. Cronbach en P.E. Meehl, Construct validity in psychological tests, *Psychological Bulletin* 52 (1955), 281–302.
- 16 S. Epskamp, *Network Psychometrics*, proefschrift Universiteit van Amsterdam, 2017, <http://sachaepskamp.com/dissertation/EpskampDissertation.pdf>
- 17 S. Epskamp, D. Borsboom en E.I. Fried, Estimating psychological networks and their accuracy: A tutorial paper, *Behavior Research Methods* 50(1) (2018), 195–212.
- 18 S. Epskamp, A.O.J. Cramer, L.J. Waldorp, V.D. Schmittmann en D. Borsboom, Qgraph: Network visualizations of relationships in psychometric data, *Journal of Statistical Software* 48 (2012), 1–18.
- 19 S. Epskamp, G.K.J. Maris, L.J. Waldorp en D. Borsboom, Network psychometrics, in: P. Irwing, D. Hughes en T. Booth (Eds.), *Handbook of Psychometrics*, Wiley, 2018.
- 20 J.M.B. Haslbeck en L.J. Waldorp, mgm: Estimating time-varying mixed graphical models in high-dimensional data, *Journal of Statistical Software* (2019), in press.
- 21 D.O. Hebb, *The Organization of Behavior*, Wiley, 1949.
- 22 B.W. Junker en J.L. Ellis, A characterization of monotone unidimensional latent variable models, *The Annals of Statistics* 25 (1997), 1327–1343.
- 23 K.S. Kendler en C.A. Prescott, *Genes, Environment, and Psychopathology: Understanding the Causes of Psychiatric and Substance Use Disorders*, Guilford Press, 2006.
- 24 R. Kindermann en J.L. Snell, *Markov Random Fields and their Applications*, American Mathematics Society, 1980.
- 25 R. Kroeze, D.C. van der Veen, M.N. Servaas, J.A. Bastiaansen, R.C. Oude Voshaar, D. Borsboom, H.G. Ruhe, R.A. Schoevers H. Riese, Personalized feedback on symptom dynamics of psychopathology: A proof-of-principle study, *Journal for Person-Oriented Research* 3 (2017), 1–10.
- 26 M. Marsman, D. Borsboom, J. Kruijs, S. Epskamp, R. van Bork, L.J. Waldorp, H.L.J. van der Maas en G.K.J. Maris, An introduction to network psychometrics: Relating Ising network models to item response theory models, *Multivariate Behavioral Research* 53 (2018), 15–35.
- 27 G.J. Mellenbergh, Generalized linear item response theory, *Psychological Bulletin* 115 (1994), 300–307.
- 28 J. Michell, Quantitative science and the definition of measurement in psychology, *British Journal of Psychology* 88 (1997), 355–383.
- 29 J. Michell, *Measurement in Psychology: A Critical History of a Methodological Concept*, Cambridge University Press, 1999.
- 30 M.L. Pe, K. Kircanski, R.J. Thompson, L.F. Bringmann, F. Tuerlinckx, M. Mestdagh, J. Mata, S.M. Jaeggi, M. Buschkuhl, J. Jonides, P. Kuppens en I. H. Gotlib, Emotion-network density in major depressive disorder, *Clinical Psychological Science* 3 (2015), 292–300.
- 31 J. Pearl, *Causality: Models, Reasoning, and Inference*, 2nd ed., Cambridge University Press, 2009.
- 32 R Core Team R: A language and environment for statistical computing, R Foundation for Statistical Computing, 2014, www.R-project.org.
- 33 G. Rasch, *Probabilistic Models for Some Intelligence and Attainment Tests*, Paedagogisk Institut Copenhagen, 1960.
- 34 P. Ravikumar, M.J. Wainwright en J.D. Lafferty, High-dimensional Ising model selection using l1-regularized logistic regression, *The Annals of Statistics* 38 (2010), 1287–1319.
- 35 D. Rizopoulos, ltm: An R package for latent variable modeling and item response theory analyses, *Journal of Statistical Software* 17 (2006), 1–25.
- 36 D.J. Robinaugh, N.J. Leblanc, H.A. Vuletich en R.J. McNally, Network analysis of persistent complex bereavement disorder in conjugally bereaved adults, *Journal of Abnormal Psychology* 123 (2014), 510–522.
- 37 T.L. Rodebaugh, N.A. Tonge, M.L. Piccirillo, E. Fried, A. Horenstein, A.S. Morrison, P. Goldin, J.J. Gross, M.H. Lim, K.C. Fernandez, C. Blanco, F.R. Schneier, R. Bogdan, R.J. Thompson en R.G. Heimberg, Does centrality in a cross-sectional network suggest intervention targets for social anxiety disorder? *Journal of Consulting and Clinical Psychology* 86(10) (2019), 831–844.
- 38 M. Scheffer, J.E. Bolhuis, D. Borsboom, T.G. Buchman, S.M.W. Gijzel, D. Goulson, J.E. Kammenga, B. Kemp, I.A. van de Leemput, S. Levin, C.M. Martin, R.J.F. Melis, E.H. van Nes, L.M. Romero en M.G.M. Olde Rikkert, Quantifying resilience of humans and other animals, *Proceedings of the National Academy of Sciences* 115 (2018), 11883–11890.
- 39 M. Scheffer, S.R. Carpenter, T.M. Lenton, J. Bascompte, W. Brock, V. Dakos, J. van de Koppel, I.A. van de Leemput, S.A. Levin, E.H. van Nes, M. Pascual en J. Vandermeer, Anticipating critical transitions, *Science* 338 (2012), 344–348.
- 40 L. Schwenen, C.D. van Borkulo, E. Fried en I.M. Goodyer, Assessment of symptom network density as prognostic marker of treatment response in adolescent depression, *JAMA Psychiatry* 75 (2018), 98–100.
- 41 C. Spearman, General intelligence, objectively determined and measured, *American Journal of Psychology* 15 (1904), 201–293.
- 42 M. Wichers, P. Groot, C.J.P. Simons, F. Peeters, Psychosystems Group, ESM Group en ESW Group, Critical slowing down as a personalized early warning signal for depression, *Psychotherapy and Psychosomatics* 85 (2016), 114–116.