

Diagnosis ‘threats’: Zo dreigend zijn ze niet

- ▶ Isabella J.M. Niesten
- ▶ Harald Merckelbach
- ▶ Brechje Dandachi-FitzGerald

Inleiding

Kan een diagnose zich tegen de patiënt keren? Kan een patiënt onbedoeld de negatieve connotaties van een diagnose op zichzelf van toepassing verklaren en daarmee van de regen in de drup raken? Daarover lopen de meningen uiteen. Discussies spitsen zich vooral toe op het risico van diagnostisch vocabulaire dat vage en ambigue klachten probeert te beschrijven (zie voor een bondige discussie Creed e.a., 2010). Want om patiënten het gevoel te geven dat ze serieus worden genomen, voorzien artsen en psychologen zulke klachten nogal eens van indrukwekkende labels. Te denken valt aan *fibromyalgie*, *whiplash*, *chronisch vermoeidheidssyndroom*, *postconcussiesyndroom* of *prikkelbare darmsyndroom*. Soms ook worden de klachten geschaard onder geaccepteerde en duidelijk klinkende diagnoses als depressie of zelfs ADHD.

Er zijn auteurs die beweren dat het diagnostisch labelen van vage klachten averechts werkt. Het zou bevorderen dat mensen de rol van patiënt aannemen en zich conform het label gaan gedragen (Buitenhuis, 2012; Buitenhuis e.a., 2008; Page & Wessely, 2003). Buitenhuis en de Jong (2009) wijzen bijvoorbeeld op de prevalentiecijfers voor whiplash. Toen whiplash een in de media veelbesproken diagnose was, staken er steeds meer gevallen de kop op. Nu de populariteit van de term onder clinici tanende is, komt de aandoening minder vaak voor. Dit patroon suggereert – maar niet meer dan dat – dat patiënten hun vage klachten kunnen modelleren naar de ideaaltypische kenmerken van een diagnostisch label.

Er zijn verhitte discussies gevoerd over de vraag of artsen in het bijzijn van hun patiënten wel over bepaalde diagnoses moeten beginnen, of de voordelen van het noemen van een diagnostisch label opwegen tegen de nadelen (Huibers & Wessely, 2006), en of sommige labels wellicht schadelijker zijn dan anderen (Hamilton e.a., 2005; Kempe e.a., 2013). Het zijn

belangrijke vragen en ze raken direct aan de klinische praktijk. Woodward en collega's (1995) vonden bijvoorbeeld dat 70% van de Australische artsen afkerig was van het gebruik van de diagnose chronisch vermoeidheidssyndroom, omdat deze diagnose patiënten zou bestendigen in hun passieve rol van zieke. Aan de andere kant doen huisartsen, (neuro)psychologen en andere zorgverleners op ruime schaal aan diagnostiek en kunnen vaak niet anders dan de resultaten daarvan delen met hun patiënten. Bezorgdheid over de mogelijk iatrogene effecten die kunnen optreden als clinici diagnostische labels opperen voor vage klachten is daarom begrijpelijk.

Diagnostiek in het lab

Het idee dat diagnostische labels iatrogene effecten kunnen sorteren, is inmiddels aan empirische toetsing onderworpen. In de literatuur wordt in dat verband gesproken over *diagnosis threat* en soms ook over *stereotype threat* (zie onder). Omdat deze termen zo ingeburgerd zijn, zullen we ze hier onvertaald laten. *Diagnosis threat* verwijst naar de volgende (hypothetische) dynamiek: mensen op wie een diagnose van toepassing wordt verklaard, zullen geïntimideerd raken door zo'n diagnose en haar connotaties; ze gaan vervolgens slechter presteren en wel zodanig dat het past binnen de diagnose (Suhr & Gunstad, 2002). Dus bij een diagnose die sterke associaties heeft met concentratieproblemen – neem ADHD – zou de patiënt dat soort problemen gaan overrapporteren als hij wordt blootgesteld aan de diagnose. Het effect zou zich ook vertalen in lagere prestaties op neuropsychologische tests, zelfs als blootstelling aan de diagnose subtiel is (Kit e.a., 2008).

Suhr en Gunstad (2002) waren de eersten die *diagnosis threat* bestudeerden. Hun proefpersonen waren studenten die in het verleden een mild hersenletsel hadden opgelopen, waarbij er hooguit dertig minuten sprake was geweest van bewustzijnsverlies. Op willekeurige basis werden studenten toegewezen aan een conditie met *diagnosis threat* of aan een controleconditie (geen blootstelling aan het diagnostische label). Binnen de conditie met *diagnosis threat* werden de studenten eraan herinnerd dat zij in het verleden een hersenletsel hadden opgelopen. Ook kregen ze te horen dat onderzoek laat zien dat dit type letsel vaak gepaard gaat met verminderd cognitief functioneren. Studenten in de controleconditie werden niet herinnerd aan hun medische voorgeschiedenis. Vervolgens werden in beide condities cognitieve tests afgenomen. Studenten die waren blootgesteld aan *diagnosis threat* bleken significant slechter te scoren op tests voor intelligentie en geheugen. Het enkel op-

wekken van herinneringen aan een diagnose bleek dus voldoende om met die diagnose geassocieerde beperkingen uit te lokken. In een tweede studie met een overeenkomstige opzet replicateerden Suhr en Gunstad (2005) dit effect.

Een meer systematische studie werd beschreven door Ozen en Fernandes (2011). Studenten die ooit een hersenletsel hadden opgelopen werden weer toegewezen aan een threat-conditie ($n = 22$) of een controleconditie ($n = 21$). Maar dit keer werden er een threat-conditie ($n = 21$) en een controleconditie ($n = 23$) aan het experiment toegevoegd bestaande uit proefpersonen zonder voorgeschiedenis van hersenletsel. Wat hun verdere achtergrond betreft, waren deze deelnemers in alle opzichten vergelijkbaar met de proefpersonen in de eerste twee condities. Anders dan Suhr en Gunstad vonden deze auteurs niet dat diagnosis threat prestaties op cognitieve taken ondermijnt. Studenten die een voorgeschiedenis van hersenletsel hadden bleken na diagnosis threat echter wél een verslechterd cognitief functioneren te *rapporteren*. Studenten zonder zo'n voorgeschiedenis waren ongevoelig voor diagnosis threat. Dat is echter niet zo vreemd als je bedenkt dat zij, na te zijn geïnformeerd over de langetermijneffecten van mild hersenletsel, te horen kregen dat zij behoorden tot de gezonde groep wier gegevens vergeleken zouden worden met die van mensen met een geschiedenis van hersenletsel. Deze proefpersonen werden zodoende blootgesteld aan een diagnostisch label dat voor hen irrelevant was. Het is dan ook discutabel of je dit, zoals de auteurs herhaaldelijk doen, een threat-conditie kan noemen.

Onlangs keken Pavawalla en collega's (2013) naar de rol van zelfrelevantie – de mate waarin iemand zichzelf identificeert met een diagnose – in de ondermijnende effecten van diagnosis threat. Gebruikmakend van een onderzoeksopzet die vergelijkbaar was met die van voorgaande studies, vonden deze auteurs dat diagnosis threat de cognitieve prestaties verslechtert van mensen met een voorgeschiedenis van traumatisch hersenletsel. Maar het effect deed zich enkel gelden bij mannen en trad alleen dan op als zij zich met de diagnose konden identificeren.

Een aantal studies keek naar de averechtse effecten van (neuro)psychiatrische diagnoses (Farina e.a., 1971; Quinn e.a., 2004), de verwachtingen die ouderen hebben over ouderdom en cognitief functioneren (Haslam e.a., 2012), en de verwachtingen die gebruikers hebben over de langetermijneffecten van ecstasy (Cole e.a., 2006). Als sleutelterm hanteerden de onderzoekers daarbij niet diagnosis threat, maar de verwante termen 'negatieve stereotypering' en 'stereotype threat'. Ze verwijzen niet naar de intimiderende werking van diagnoses, maar naar die van stereotypen over bepaalde sociale groepen. Een voorbeeld is het stereotype dat oude-

ren een zwak geheugen hebben. Wanneer iemand zich bewust wordt van deze stereotypen, zou hij zich in lijn daarmee gaan gedragen (zie voor een uiteenzetting Kit e.a., 2008). Zo vonden Farina en collega's (1971) dat psychiatrische patiënten taken minder goed uitvoeren als hun wordt voorgehouden dat een andere deelnemer weet heeft van hun psychiatrische achtergrond. Volgens de auteurs zou dit komen door het stigma dat aan psychiatrische diagnoses kleeft. Een replicatie van deze bevinding werd beschreven door Quinn en collega's (2004), maar hun studie richtte zich op cognitieve prestaties van mensen met een depressie. De meest krasse effecten werden gerapporteerd door Haslam en collega's (2012). Zij gaven oudere proefpersonen het idee tot de oudste of tot de jongste deelnemers in de studie te behoren en stelden hen vervolgens bloot aan negatieve stereotypen die geassocieerd zijn met veroudering. Zo'n 70% van de proefpersonen in de 'oud'-conditie ging een dementieel patroon op cognitieve tests aan de dag leggen (tegen 14% van de 'jong'-conditie). Significante effecten van diagnosis threat werden ook gerapporteerd door Cole en collega's (2006) maar dan in een groep van ecstasy-gebruikers. De onderzoekers informeerden de helft van hun proefpersonen dat geheugenproblemen door ecstasy-gebruik een wetenschappelijk vaststaand feit zijn. De andere helft kreeg te horen dat een verband tussen geheugenproblemen en ecstasy wetenschappelijk gezien niet vaststaat. De deelnemers in de eerste conditie vertoonden vervolgens slechtere geheugenprestaties dan de deelnemers in de tweede conditie.

Een kritische blik

Allerlei auteurs beschouwen het ondermijnend effect van diagnosis threat als een krachtig verschijnsel. In die gedaante duikt het ook op in beschouwende overzichtsartikelen. Silver (2012; p. 838) schrijft bijvoorbeeld dat '[p]erformance under stereotype threat is significantly impaired; when the same test is administered under conditions that eliminate the threat, performance is normal.' Pollak (2013; p. 1) bouwt daarop voort door te concluderen dat '[p]articularly salient is the possibility that expectation, leading to *increased* attention to task demands, results in paradoxically poorer performance'. En Kit en collega's (2008; p. 144) waarschuwen zelfs dat 'stereotype threat poses a risk within a clinical environment'. Dat soort uitspraken veronderstelt niet alleen significante bevindingen, maar ook imposante effectgroottes. In Tabel 1 hebben we resultaten van het eerder besproken onderzoek op meer formalistische wijze samengevat.

TABEL 1 Effectgrootten per studie en het gewogen gemiddelde

Studie	N	Type test (aantal)	Effectgrootte (Cohen's d)	95% CI
Suhr en Gunstad (2002)	36	Cognitieve tests (10)	.38	[-0,28, 1,04]
Suhr en Gunstad (2005)	53	Cognitieve tests (13)	.43	[-0,12, 0,97]
Ozen en Fernandes (2011) ^a	43	Cognitieve tests (5) Symptoomlijsten (5)	.01	[-0,59, 0,61]
Ozen en Fernandes (2011) ^b	44	Cognitieve tests (5) Symptoomlijsten (5)	.001	[-0,59, 0,59]
Pavawalla e.a. (2013)	98	Cognitieve tests (1) Symptoomlijsten (1)	.07	[-0,33, 0,47]
			Gewogen gemiddelde	Gewogen 95% CI
			.16	[-0,08, 0,40]

Noot. Cohen's *d*: zwak effect = .20, gemiddeld effect = .50, groot effect = .80. Voor de meta-analyse maakten we gebruik van het random effects model. De heterogeniteit tussen de studies bedraagt $I^2 = 0\%$ ($\leq 25\%$ = lage heterogeniteit, 50% = matige heterogeniteit, $\geq 75\%$ = hoge heterogeniteit). De gewogen gemiddelden voor cognitieve tests en symptoomlijsten afzonderlijk bedragen respectievelijk .20, CI [-0,04, 0,44] en -.02, CI [-0,31, 0,27]. De gebruikte software is ESCI, zie Cumming (2012). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. New York: Routledge.

Ozen & Fernandes^a = De groep met geschiedenis van hersenletsel. Ozen & Fernandes^b = de groep zonder geschiedenis van hersenletsel.

Het patroon is duidelijk: op de bevindingen van Suhr en Gunstad na, zijn zowel de afzonderlijke effectgroottes als het gewogen gemiddelde van de effectgroottes ($d = .16$, 95% CI [-0,08, 0,40]) qua omvang bescheiden. Ook in dit domein van onderzoek geldt dat een significant effect niet gelijkstaat aan een klinisch imposant effect.

Er zijn bovendien kanttekeningen te plaatsen bij de definitie en de reikwijdte van diagnosis threat. Om met het eerste te beginnen: onderzoekers hanteren het begrip diagnosis threat nogal eens losjes. Zo definiëren Pavawalla en collega's (2013; p. 305) het als 'reduced cognitive performance due to preexisting beliefs', waarmee zij verwijzen naar de gevolgen (oftewel: de afhankelijke variabele) van blootstelling aan het diagnostisch label. Maar dezelfde auteurs spreken ook over diagnosis threat in de zin van de blootstelling zelf (oftewel: de onafhankelijke variabele). De verwarring van het diagnostisch label met het averechtse ef-

fect dat het label sorteert, maakt discussies vaag. Om dat te vermijden spreken wij hier consequent over het ondermijnend effect van diagnosis threat.

Dan de reikwijdte: hoe universeel is het effect en drijft het ook wel op *threat*, angst derhalve? Wie de studie van Suhr en Gunstad (2005) erop naslaat, zal zien dat het ondermijnend effect van diagnosis threat niet verklaard kan worden door hoe angstig proefpersonen zich voelen. Ozen en Fernandes (2011) vonden zelfs dat hun controlegroep meer angst ervoer dan de threat-groep, iets wat men intuïtief niet zou verwachten. Diagnosis *threat* lijkt derhalve geen passende term om te beschrijven wat er in het hoofd van proefpersonen gebeurt.

Dat Ozen en Fernandes, anders dan Suhr en Gunstad, het averechtse effect van labels alleen aantreffen op symptoomlijsten is veelzeggend. De kans bestaat immers dat het effect tot stand komt omdat proefpersonen de hypothese van de onderzoekers raden en graag behulpzaam willen zijn. Een studie in een iets andere context liet zien dat het makkelijk is om korte metten te maken met de ondermijnende effecten van diagnosis threat, namelijk door proefpersonen geld te bieden voor goede prestaties (Barber & Mather, 2013). Ook de bevinding van Pavawalla en collega's (2013) dat het averechtse effect van diagnosis threat slechts optreedt bij mannen en alléén dan als zij zich met het diagnostisch label identificeren, doet af aan de veronderstelde universaliteit van het fenomeen.

Al deze overwegingen suggereren dat het averechtse effect van diagnostische labels reuze meevalt. Dat zou betekenen dat de klinische relevantie van het fenomeen minder groot is dan sommige auteurs ons willen doen laten geloven. Om de proef op de som te nemen, voerden we zelf twee experimenten uit.

Verminderde aandacht en concentratie

In een eerste studie probeerden wij de resultaten van Suhr en Gunstad (2002; 2005) te repliceren door studenten bloot te stellen aan tekortkomingen die vaak worden geassocieerd met ADHD. We verdeelden studenten daarom over een threat-conditie ($n = 16$) en een controleconditie ($n = 15$). Studenten in de threat-conditie kregen te horen dat ze waren geselecteerd omdat ze tijdens een eerdere diagnostische screening onder eerstejaarsstudenten ondermaats hadden gescoord op concentratie en aandacht. Zo'n screening had daadwerkelijk plaatsgevonden, maar dat deze studenten afwijkend hadden gescoord, was niet het geval. Meer in het bijzonder kregen de studenten de volgende instructie: 'Je bent geselecteerd voor dit onderzoek omdat je tijdens de screening bij aanvang van

dit jaar lager dan gemiddeld scoorde op concentratie en aandacht. Wij willen je daarom verzoeken een aantal vragenlijsten in te vullen en een korte taak uit te voeren.' Studenten in de controleconditie ontvingen louter het verzoek om vragenlijsten in te vullen en een taak uit te voeren. Vervolgens vulden beide groepen een zelfrapportagelijst in over ADHD-symptomen. Een deel van de items ging erover in welke mate men als kind last had gehad van typische ADHD-symptomen. Een ander deel ging over ADHD-symptomen in het hier en nu (bijvoorbeeld: 'Ik heb een continu gevoel van rusteloosheid'). Proefpersonen kregen ook de d2-test, een veelgebruikte test die aandacht en concentratie meet. De crux ervan is dat proefpersonen een doelstimulus – een d met twee streepjes – moet aankruisen die voortdurend ligt ingebed in afleiders (p's, d's met een streep et cetera). De taak laat het toe om het aantal correct en foutief aangestreepte items te turven en af te zetten tegen normgegevens.

Tabel 2 laat de resultaten zien voor de ADHD-symptomen en de d2-testscores. De twee groepen verschilden niet in zelfgerapporteerde ADHD-symptomen. Dit gold zowel voor symptomen tijdens de kindertijd als de volwassen levensjaren. De groepen verschilden evenmin in hun prestaties op de d2-test. Een ondermijnd effect van diagnosis threat trad niet op bij het aantal correct verwerkte items, en evenmin bij het aantal gemaakte fouten. Een heranalyse van de data waarbij we de gegevens van een proefpersoon excludeerden vanwege zijn extreme scores op de ADHD-lijst veranderde niets aan deze stand van zaken. De gemiddelde effectgrootte van diagnosis threat in onze studie was $d = -.23$, 95% CI [-0,94, 0,48]. Dat wil zeggen dat het effect de verkeerde kant uitging. Zodoende vonden we eerder steun voor een boemerang-effect (proefpersonen proberen te laten zien dat de labels niet op hen van toepassing zijn; Hovland e.a., 1953) dan voor de ondermijnende werking van diagnosis threat.

Toegegeven: we stelden onze deelnemers niet bloot aan de diagnose van ADHD, maar aan tekortkomingen die vaak geassocieerd worden met zo'n diagnose. We bestudeerden dus 'symptom' en niet diagnosis threat. Het principe zou echter hetzelfde moeten zijn: wanneer mensen wordt aangepreft dat zij minder goed functioneren op een bepaald gebied, zou dat volgens de doctrine van diagnosis threat moeten resulteren in lagere prestaties op dat gebied. Dat vonden we dus niet. Onze bevindingen zijn strijdig met de gedachte dat alleen al het benoemen van diagnostische labels of kenmerken – in dit geval ADHD-symptomen – leidt tot het pathologiseren van normale, alledaagse ervaringen.

TABEL 2 Gemiddelde scores (SD) op de ADHD vragenlijst en D2-test van proefpersonen in de threat-conditie ($n = 16$) en de controleconditie ($n = 15$)

	Threat-conditie	Controleconditie	t (p)
Vragenlijst			
ADHD kindertijd	10,9 (8,0)	13,9 (6,1)	-1,14 (.26)
ADHD Volwassenheid	12,7 (9,3)	12,6 (7,4)	.03 (.98)
Cognitieve taak			
D2-test Aantal Verwerkt	205,4 (42,2)	195,0 (33,4)	.77 (.45)
D2-test Aantal Fout	96,5 (41,1)	106,3 (33,2)	-0,72 (.48)

Noot. D2-testscores ontbreken voor één deelnemer (threat-conditie). Deze gemiddelden zijn daarom gebaseerd op een kleinere steekproef.

Hersenletsel

In onze tweede studie hanteerden we een onderzoeksofzet die meer leek op die van Suhr en Gunstad. Aan deze studie deden enkel studenten mee die eerder tijdens een algemene screening hadden aangekruist in de afgelopen tien jaren een mild hersenletsel te hebben opgelopen. Deze studenten werden blootgesteld aan threat ($n = 15$) of geen threat (controleconditie; $n = 14$). De threat bestond eruit dat proefpersonen een stuk lezen waarin ze bewust werden gemaakt van het eerdere hersenletsel en de daar mogelijk mee gepaard gaande cognitieve gevolgen. De tekst luidde: 'Je hebt bij aanvang van dit studiejaar meegedaan aan de screening die bij alle eerstejaars psychologiestudenten plaatsvindt. Daarbij heb je een vragenlijst ingevuld die als volgt begon' – hier werd een gedeelte van de Head Injury Symptom Checklist getoond – 'Op deze vragenlijst heb je ingevuld dat je in het verleden (in de afgelopen tien jaar) een hersenschudding of een forse klap op je hoofd hebt gehad. Uit onderzoek blijkt dat zulke gebeurtenissen gepaard kunnen gaan, ook op langere termijn, met cognitieve beperkingen. Om inzicht te krijgen in de aard van deze beperkingen, doen we onderzoek en daarom verzoeken we je een aantal vragenlijsten in te vullen en taken uit te voeren.' De studenten in de controleconditie werden niet herinnerd aan hun vroegere hersenletsel. Ver-

volgens kregen de proefpersonen een korte testbatterij voorgelegd. Die bestond uit de eerste 45 items van de Symptom Checklist-90 (SCL-90), de Posttraumatic Stress Disorder Symptom Scale (PSS), de Stroop Kleur-Woord-test en de Trail Making Test (TMT).

Tabel 3 laat de resultaten zien. De groepen verschilden niet in hun scores op de SCL-90 of de PSS. De groepen verschilden evenmin in hun interferentie-scores op de Stroop of de TMT. De gemiddelde effectgrootte was dit keer $d = .06$, 95% CI [-0,67, 0,79]. Dus zelfs als proefpersonen worden geconfronteerd met een voor hen relevant label, nemen de ondermijnende effecten van diagnosis threat geen significante vorm aan. Je zou kunnen denken dat onze nulbevindingen haaks staan op eerdere resultaten, maar dat is slechts zeer ten dele het geval. De effectgroottes die we vonden stemmen heel wel overeen met eerder onderzoek. Dit maakt duidelijk dat het gaat om een fragiel fenomeen.

TABEL 3 Gemiddelde scores op de SCL-90, de PSS, TMT, en de Stroop Kleur-Woord-test van proefpersonen in de threat-conditie ($n = 15$) en de controleconditie ($n = 14$)

	<i>Threat-conditie</i>	<i>Controleconditie</i>	<i>t (p)</i>
Vragenlijsten			
SCL-90 totaalscore	66,5 (13,5)	63,4 (13,9)	-.61 (.55)
PSS totaalscore	1,6 (2,4)	2,5 (3,2)	.86 (.40)
Cognitieve taken			
TMT interferentiescore	10,0 (3,6)	8,6 (5,0)	-.78 (.44)
Stroop interferentiescore	24,8 (9,7)	24,5 (9,4)	-.09 (.93)

Noot. De SCL-90 totaalscore is gebaseerd op 45 items. De cognitieve taken bestonden elk uit drie onderdelen. Interferentiescores zijn berekend door de uitvoeringstijd van onderdeel 3 (interferentie) te delen door de som van de gemiddelde uitvoeringstijd van onderdelen 1 en 2 (geen interferentie). TMT interferentiescores ontbreken voor drie deelnemers (twee uit de threat-conditie en één uit de controleconditie). Deze gemiddelden zijn daarom gebaseerd op een kleinere steekproef.

Conclusie

Fragiel of niet, auteurs dichten het fenomeen een verklarende rol toe in uiteenlopende verbanden. Rohlings en medewerkers (2012) schrijven bijvoorbeeld dat patiënten met medisch onverklaarbare klachten na een mild hersenletsel vals-positieven zijn die hun beslag krijgen door diagnosis threats. Andere auteurs gaan nog een stap verder en stellen dat test-resultaten die gewoonlijk worden toegeschreven aan het veinzen van symptomen beter te verklaren zijn door de ondermijnende effecten van diagnosis threat (Bigler 2012; Silver; 2012). Falen op tests die onderprestatie meten zou simpelweg optreden omdat mensen zich inbeelden een aandoening te hebben en anderen en zichzelf daarvan proberen te overtuigen: diagnosis threat zou hen er onbewust toe verleiden om ondermaats te gaan presteren. Het is een hypothese die op enige schaal de ronde doet, maar die zich slecht verdraagt met de fragiliteit van het fenomeen.

Want inderdaad: als wij de effectgrootten van onze studies samenvoegen met de effectgrootten van eerdere threat-studies valt onmogelijk aan de conclusie te ontsnappen dat het iatrogene effect van diagnosis threat ronduit zwak is (gewogen $d = .11$, 95% CI [-0,10, 0,33]). Op zichzelf genomen zijn diagnostische labels niet zo riskant en dreigend als veel auteurs ons willen doen laten geloven. De negatieve effecten van diagnosis threat als robuuste fenomenen voorstellen en ze als hoekstenen aandragen voor de verklaring van onderprestatie komt ons daarom als simplistisch voor.

De diagnostiek van vage klachten is een complexe aangelegenheid die in de praktijk verder reikt dan enkel het aandragen van een label. Er zijn andere factoren die bij het ontstaan van iatrogene effecten een rol spelen (Bender & Matuszewicz, 2013). Zo is het niet ongewoon dat een arts – met de beste bedoelingen – herhaaldelijk tests aanvraagt, symptomen probeert te behandelen en de patiënt naar allerlei specialisten stuurt. Zulke interventies kunnen als misinformatie over de ernst van symptomen gaan fungeren, waardoor de persoon met vage klachten meer en meer gevangen raakt in zijn rol als patiënt (Bender & Matuszewicz, 2013; Hatcher & Arroll, 2008; Page & Wessely, 2003).

Eerder beschreven we (Merckelbach e.a., 2011a; 2011b) de uitwerking van misinformatie als die wordt aangereikt door een ogenschijnlijk kundig persoon. In een studie gaven we bijvoorbeeld proefpersonen misleidende feedback over symptomen die ze nooit hadden gerapporteerd. Zo'n 60% van de proefpersonen had de discrepantie tussen wat hen werd voorgehouden en wat ze eerder de facto hadden gerapporteerd niet door.

Uiteindelijk maakten de proefpersonen de aangereikte symptomen tot de hunne.

Vaak roepen vage klachten wisselende reacties op bij de omgeving (inclusief de arts), zoals frustratie, onbegrip, en twijfels over de authenticiteit van de door de patiënt gepresenteerde klachten (Hatcher & Arroll, 2008; Huibers & Wessely, 2006). De neiging om de klachten te rechtvaardigen en de aanwezigheid van beperkingen te bewijzen, kan in zulke gevallen sterk zijn en uitmonden in symptoomoverdrijving. Zo'n route is vooral riskant als er secundaire belangen op het spel staan – bijvoorbeeld het verwerven van een uitkering of medicatie – die vereisen dat de patiënt herhaaldelijk bewijs levert voor zijn ziekte. Zulk soort belangen is niet zeldzaam. Van Egmond en Kummeling (2002) vonden dat maar liefst 42% van de psychiatrische patiënten nog andere belangen (uitkering, huisvesting et cetera) had dan alleen maar beter worden. De behandelresultaten voor deze groep waren beduidend somberder dan die voor patiënten zonder secundaire belangen.

Secundaire belangen kunnen aanzetten tot het willens en wetens overdrijven van klachten. Patiënten doen dat ook: de mate waarin financiële belangen testprestaties kunnen dempen, bereikt eenvoudig een effectgrootte van $d = 0,9$ (Binder e.a., 1997). Ondertussen ervaren de meeste mensen het opzettelijk overdrijven van symptomen als iets onaangenaams (Merckelbach & Merten, 2012; Niesten e.a., 2014). Dat heeft te maken met het conflict dat simulanten voelen tussen hun zelfdefinitie ('ik ben een eerlijk persoon') en hun misleidend gedrag. De technische term daarvoor is cognitieve dissonantie. Mensen hebben de natuurlijke neiging om cognitieve dissonantie op te heffen door hun interpretaties te veranderen en zo'n herinterpretatie zal zich bij simulanten richten op de gerapporteerde symptomen. Dissonantie zal simulanten er dus toe aanzetten om geveinsde symptomen te herinterpreteren als echte symptomen en dat kan uitmonden in symptoomescalatie (zie voor empirische ondersteuning Merckelbach e.a., 2011b; Niesten e.a., 2014). Het is langs deze weg dat sommige patiënten zichzelf misinformereren en daarmee een pathologische dynamiek in werking zetten.

De boodschap van het voorgaande is dat het iatrogene effect van diagnostische labels wordt overschat. De literatuur overziend luidt onze conclusie dat zulke labels pas een iatrogene lading krijgen in combinatie met interventies, secundaire belangen, en klachtenoverdrijving. Het is deze constellatie die onbedoeld misinformatie genereert en wel zo dat – om Freuds metafoer te gebruiken – wat eerst slechts de pantomime van een ziekte was uiteindelijk een echte aandoening wordt. Als we iatrogene effecten willen begrijpen en willen voorkomen, dan kan onderzoek naar

deze constellatie ons verder helpen. Daarom: klinici zouden zich minder zorgen moeten maken over het fnuikende effect van diagnostische termen en meer moeten kijken naar hoe deze termen het risico van overbodige interventies oproepen. Daarnaast zou bij vage klachten een vorm van diagnostiek mogelijk moeten zijn waarbij de patiënt wordt uitgedaagd om de diagnose te weerleggen. Maar alvorens we toe komen aan de vraag hoe zo'n boemerangeffect valt op te roepen en therapeutisch kan worden geëxploiteerd, moeten we eerst af van die overmatige aandacht voor diagnosis threat. Daaraan probeerde dit artikel een bijdrage te leveren.

Samenvatting Het idee dat diagnostische labels averechts kunnen werken, is wijdverbreid. Sommige studies lijken inderdaad te hebben aangetoond dat mensen slechter gaan presteren op allerlei tests zodra ze horen dat labels als ADHD of hersenletsel op hen van toepassing zijn. Dat effect heet 'diagnosis threat'. Maar is het wel zo krachtig en universeel? In dit artikel nemen we het onderzoek naar diagnosis threat kritisch onder de loep. Ook bespreken we eigen onderzoek dat laat zien dat het averechtse effect van labels lang niet zo sterk is als wel wordt verondersteld. Onze conclusie is dat diagnostische labels pas problematisch worden als zij hand in hand gaan met voorbarige interventies, de aanwezigheid van secundaire belangen en klachtenoverdrijving. Clinici doen er wijs aan daaraan aandacht te besteden en zich minder zorgen te maken over diagnosis threat.

Isabella J.M. Niesten Faculteit der Psychologie en Neurowetenschappen, sectie Forensische Psychologie, Universiteit Maastricht. **Correspondentieadres:** I.J.M. Niesten, MSc., Universiteit Maastricht, faculteit der Psychologie en Neurowetenschappen, sectie Forensische Psychologie, postbus 616, 6200 MD Maastricht, elly.niesten@maastrichtuniversity.nl.

Harald Merckelbach Faculteit der Psychologie en Neurowetenschappen, sectie Forensische Psychologie, Universiteit Maastricht.

Brechje Dandachi-FitzGerald School for Mental Health and Neuroscience, Faculty of Health, Medicine & Life Sciences, Universiteit Maastricht; Mondriaan voor geestelijke gezondheid, Maastricht.

Literatuur

Barber, S.J. & Mather, M. (2013). Stereotype threat can both enhance and impair older adults' memory. *Psychological*

Science, 24, 2522-2529.

doi:10.117/0956797613497023.

Bender, S.D. & Matuszewicz, M. (2013). PCS, iatrogenic symptoms, and malingering following concussion. *Psychological In-*

- jury and Law, 6, 113-121. doi:10.1007/s12207-013-9156-9.
- Bigler, E.D. (2012). Symptom validity testing, effort, and neuropsychological assessment. *Journal of the International Neuropsychological Society*, 18, 632-642. doi:10.1017/S1355617712000252.
- Binder, L.M., Rohling, M.L. & Larrabee, G. (1997). A review of mild head trauma. Part I: Meta-analytic review of neuropsychological studies. *Journal of Clinical and Experimental Neuropsychology*, 19, 421-431.
- Buitenhuis, J. (2012). Posttraumatische nekpijn: Whiplash, imago en verwachtingen. In: P. Calders, J.J.X.R. Geraets, J. Nijs, C. Veenhof, E.E.H. van Wegen & C.P. van Wilgen (red.), *Jaarboek Fysiotherapie Kinesitherapie* (pp. 135-145). Bohn Stafleu van Loghum.
- Buitenhuis, J. & De Jong, P.J. (2009). De term 'whiplash' vermijden: Commentaar op de multidisciplinaire richtlijn voor ongecompliceerde whiplash. *Nederlands Tijdschrift voor de Geneeskunde*, 15.
- Buitenhuis, J., De Jong, P.J., Jaspers, J.P.C. & Groothoff, J.W. (2008). Catastrophizing and causal beliefs in whiplash. *Spine*, 33, 2427-2433. doi:10.1097/BRS.0b013e318183c6ca.
- Cole, J.C., Michailidou, K., Jerome, L. & Sumnall, H.R. (2006). The effects of stereotype threat on cognitive function in ecstasy users. *Journal of Psychopharmacology*, 20, 518-525. doi:10.1177/0269881105058572.
- Creed, F., Guthrie, E., Fink, P., Henningsen, P., Rief, W., Sharpe, M. & White, P. (2010). Is there a better term than 'medically unexplained symptoms'? *Journal of Psychosomatic Research*, 68, 5-8. doi:10.1016/j.jpsychores.2009.09.004.
- Egmond, J.J., van & Kummeling, I. (2002). A blind spot for secondary gain affecting treatment outcome. *European Psychiatry*, 17, 46-54.
- Farina, A., Gliha, D., Boudreau, L.A., Allen, J.G. & Sherman, M. (1971). Mental illness and the impact of believing others know about it. *Journal of Abnormal Psychology*, 77, 1-5. doi:10.1037/h0030496.
- Hamilton, W.T., Gallagher, A.M., Thomas, J.M. & White, P.D. (2005). The prognosis of different fatigue diagnostic labels: A longitudinal survey. *Family Practice*, 22, 383-388. doi: 10.1093/fampra/cmio21.
- Haslam, C., Morton, T.A., Haslam, S.A., Varner, L., Graham, R. & Gamaz, L. (2012). 'When the age is in, the wit is out': Age-related self-categorization and deficit expectations reduce performance on clinical tests used in dementia assessment. *Psychology and Aging*, 27, 778-784. doi: 10.1037/a0027754.
- Hatcher, S. & Arroll, B. (2008). Assessment and management of medically unexplained symptoms. *British Medical Journal*, 336, 1124-1128. doi: 10.1136/bmj.39554.592014.BE.
- Hovland, C.I., Janis, I.L. & Kelley, H.H. (1953). *Communication and persuasion*. New Haven: Yale University Press.
- Huibers, M.J.H. & Wessely, S. (2006). The act of diagnosis: Pros and cons of labeling chronic fatigue syndrome. *Psychological Medicine*, 36, 895-900. doi:10.1017/S0033291705006926.
- Kempe, C.B., Sullivan, K.A. & Edmed, S.L. (2013). The effect of varying diagnostic terminology within patient discharge information on expected mild traumatic brain injury outcome. *The Clinical Neuropsychologist*, 27, 762-778. doi:10.1080/13854046.2013.795245.
- Kit, K.A., Tuokko, H.A. & Mateer, C.A. (2008). A review of the stereotype threat literature and its application in a neurological population. *Neuropsychology Review*, 18, 132-148. doi:10.1007/s11065-008-9059-9.
- Merckelbach, H., Jelicic, M. & Jonker, C. (2011a). Planting a misdiagnosis of Alzheimer's disease in a person's mind. *Acta Neuropsychiatrica*, 24, 60-62. doi:10.1111/j.1601-5215.2011.00586.x.
- Merckelbach, H., Jelicic, M. & Pieters, M. (2011b). The residual effect of feigning: How intentional faking may evolve into a less conscious form of symptom reporting. *Journal of Clinical and Experimental Neuropsychology*, 33, 131-139. doi:10.1080/13803395.2010.495055.
- Merckelbach, H. & Merten, T. (2012). A note on cognitive dissonance and malingering. *The Clinical Neuropsychologist*, 26, 1217-1229.

- Niesten, I.J.M., Nentjes, L., Merckelbach, H. & Bernstein D. (ongepubliceerd manuscript). *Faking in a forensic context: DSM's conceptualization falls short.*
- Ozen, L.J. & Fernandes, M.A. (2011). Effects of 'diagnosis threat' on cognitive and affective functioning long after mild head injury. *Journal of the International Neuropsychological Society*, 17, 219-229. doi:10.1017/S135561771000144X.
- Page, L.A. & Wessely, S. (2003). Medically unexplained symptoms: Exacerbating factors in the doctor-patient encounter. *Journal of the Royal Society of Medicine*, 96, 223-226. doi:10.1258/jrsm.96.5.223.
- Pavawalla, S.P., Salazar, R., Cimino, C., Belanger, H.G. & Vanderploeg, R.D. (2013). An exploration of diagnosis threat and group identification following concussion injury. *Journal of the International Neuropsychological Society*, 19, 305-313. doi:10.1017/S135561771200135X.
- Pollak, T.A. (2013). What a jerk: Perils in the assessment of psychogenic movement disorders. *Journal of Neurology, Neurosurgery, and Psychiatry*, 0, 1. doi:10.1136/jnnp-2012-304682.
- Quinn, D.M., Kahng, S.K. & Crocker, J. (2004). Discreditable: Stigma effects of revealing a mental illness history on test performance. *Personality and Social Psychology Bulletin*, 30, 803-814. doi:10.1177/0146167204264088.
- Rohling, M.L., Larrabee, G.J. & Millis, S.R. (2012). The 'miserable minority' following mild traumatic brain injury: Who are they and do meta-analyses hide them? *The Clinical Neuropsychologist*, 26, 197-213. doi:10.1080/13854046.2011.647085.
- Silver, M. (2012). Effort, exaggeration and malingering after concussion. *Journal of Neurology, Neurosurgery & Psychiatry*, 83, 836-841. doi:10.1136/jnnp-2011-302078.
- Suhr, J.A. & Gunstad, J. (2002). 'Diagnosis threat': The effect of negative expectations on cognitive performance in head injury. *Journal of Clinical and Experimental Neuropsychology*, 24, 448-457. doi:10.1076/jcen.24.4.448.1039.
- Suhr, J.A. & Gunstad, J. (2005). Further exploration of the effect of 'diagnosis threat' on cognitive performance in individuals with mild head injury. *Journal of the International Neuropsychological Society*, 11, 23-29. doi:10.1017/S1355617705050010.
- Woodward, R.V., Broom, D.H. & Legge, D.G. (1995). Diagnosis in chronic illness: Disabling and enabling – the case of chronic fatigue syndrome. *Journal of the Royal Society of Medicine*, 88, 325-329.