

## Bewering: 'Neurowetenschap, dus nieuws'

### Twijfel

HANS VAN MAANEN



Experiment met een muis. Foto Colourbox

### De power van onderzoek

**N**eurowetenschap is een 'hot' onderzoeksgebied – zowel voor wetenschappers als voor journalisten. 'Overproductie van interleukine-6 verhoogt kans op Alzheimer.' 'Migrainelijders hebben lange variant van het dopaminereceptor-gen.' 'Prematuren vaker linkshandig' en 'Volwassenen met ADHD mogelijk gebaat bij antidepressiva.' De kranten staan er vol mee, maar hoeveel waarde moeten wij hechten aan al die neurowetenschappelijke nieuwtjes?

Weinig, is het advies van een groep Britse en Amerikaanse onderzoekers vorige maand in, nota bene, *Nature Reviews Neuroscience*. De onderzoekers, onder aanvoering van Marcus Munafò, vinden er geen doekjes om: 'Wij tonen aan dat het gemiddelde statistisch vermogen van studies in de neurowetenschap zeer laag is' – het statistisch vermogen zegt iets over de kans dat een onderzoek een werkelijk bestaand effect (of verschil, of wat dan ook) boven water krijgt. Wetenschappers spreken van power: een power van 80 procent betekent dat we acht van de tien keer het bestaande effect zullen vinden; twee keer komt de proef toevallig zo uit dat we het effect missen. Bijvoorbeeld: om met een power van 80 procent aan te tonen dat mannen gemiddeld langer zijn dan vrouwen – een groot, duidelijk effect – zijn zes mannen en zes vrouwen nodig. Twee op de tien keer zullen de mannen en vrouwen in de steekproef toch geen duidelijk verschil vertonen. Hoe kleiner het effect, des te groter de steekproef moet zijn: om 80 procent kans te hebben aan te tonen dat vrouwen meer schoenen hebben, zijn 16 mannen en 16 vrouwen nodig; om zo aan te tonen dat mannen zwaarder zijn, moeten we 47 mannen en 47 vrouwen vergelijken.

**Het probleem van de neurowetenschap**, zeggen Munafò en consorten, is dat ze op zoek zijn naar subtiele effecten maar daar evident te kleine steekproeven voor nemen. En ze kunnen het uiteraard niet laten ook andere wetenschapsgebieden – MRI-studies, experimenten met muizen en mensen – in hun beschouwing te betrekken.

De onderzoekers bedachten een mooie manier om te laten zien hoe 'underpowered' de meeste studies in de neurowetenschappen zijn. Naarmate meer onderzoek naar een verschijnsel wordt gedaan, komt langzamerhand het 'ware' effect er wel uit rollen – als er eindelijk een keer een echt grote studie wordt gedaan, of als iemand alle kleine studies bij elkaar optelt en analyseert. Terugkijkend kun je dan zien of de eerste studies groot genoeg waren om dit vastgestelde effect ooit te vinden. Dat blijkt dus maar zelden het geval te zijn, zegt Munafò. 'Onze resultaten wijzen erop dat de statistische power van studies in de neurowetenschappen waarschijnlijk niet veel hoger is dan 8 tot 31 procent.' En dat is vermoedelijk nog een overschatting.

'Het betekent,' vervolgt hij, 'dat de kans dat een significant resultaat op een echt effect berust, klein is.' Met name kleine studies met grote, opzienbarende effecten – pasmunt van de wetenschapspagina's – moeten met wantrouwen worden bekeken. Van de tien zullen er negen onzin wezen.

**Het betekent ook** dat er veel tijd, geld, moeite, proefdieren en proefpersonen worden verspild in al die te kleine studies. De onderzoekers geven het voorbeeld van ratjes die moeten leren door een doolhof te zwemmen. Leren vrouwtjes dat sneller dan mannetjes? Om dat aan te tonen, beperken onderzoekers zich veelal tot niet veel meer dan twintig ratten, maar om het bestaande effect met een power van 80 procent aan te tonen, zouden er in feite 134 nodig zijn geweest. Als twintig ratjes voldoende waren geweest, zou het verschil zo evident zijn dat het iedereen allang was opgevallen – en al die ratten dus evenmin hadden hoeven te zwemmen voor hun leven. 'De ethische implicaties zijn duidelijk,' aldus de onderzoekers.

Zij bepleiten, onder veel meer, een veel nauwkeuriger schatting van de nodige proefdieren en proefpersonen vooraf. Dat zijn er meestal beduidend meer dan onderzoekers zouden wensen, want grote experimenten zijn moeilijker gefinancierd te krijgen dan kleine.

Het streven naar snel leuke, nieuwe en onverwachte resultaten – van onderzoekers, geldschietters en tijdschriften – verspilt niet alleen de tijd van krantenlezers, het kost levens.

**Oordeel:**  
Onzin (negen van de tien keer)

### Digitale recherche

REPORTAGE HET PROFIEL VAN EEN MAILER

Veel reacties op het Amerikaanse spionageprogramma Prism komen neer op: ik heb niks te verbergen. Echt niet? Ook niet als digitale experts eens goed snuffelen?

Door Bard van de Weijer Illustratie Thijs Balder

# Wie ben ik? TNO ging op zoek in mijn 32.272 mails

**H**et voelt alsof ik weer een jongetje van 6 ben dat bij Sinterklaas op schoot zit, mijn hersenen pijnigend met de vraag of ik het afgelopen jaar wel goed geluisterd heb naar mijn ouders, mijn best gedaan heb op school en het hok van de cavia vaak genoeg heb verschoond. Het is de vraag die elk kind zich stelt als de Goedheiligman nabij is: ben ik wel braaf geweest?

Maar ik zit niet bij Sinterklaas, ik zit naast Wessel Kraaij, gespecialiseerd in het verzamelen en analyseren van grote hoeveelheden ongestructureerde data. Tegenover hem zit Mark van Staalduinen, expert digital profiling. Hij adviseert namens TNO onder meer politie over het maken en analyseren van digitale profielen. Hoewel ik de twee mannen nog nooit heb ontmoet, weten ze alles van me. Ze weten met wie ik de afgelopen vijf jaar heb gecommuniceerd en hoe vaak en wat ik met al die mensen heb besproken. Het zou zomaar kunnen dat ze de krochten van mijn liefdesleven kennen en weten met wie ik zoal in onmin leef.

Ik was dus een beetje nerveus toen ik mijn paspoort toonde aan de portier van het anonieme TNO-gebouw in Delft en ik door het poortje liep waarachter Van Staalduinen me stond op te wachten. 'Hallo, ik ben Bard. Ik hoef me denk ik niet verder te introduceren.'

Ik heb dit zelf veroorzaakt. Ik heb Kraaij, behalve onderzoeker bij TNO ook bijzonder hoogleraar aan de Radboud Universiteit, enkele dagen eerder het wachtwoord gegeven van mijn privémailbox. Daarin staan alle inkomende berichten van de afgelopen vijf jaar: 32.272 mails, waarvan trouwens 28.076 ongezien – kennelijk ben ik niet zo'n zorgvuldige lezer.

De aanleiding dat Kraaij en Van Staalduinen mijn mail bekeken is Prism, het omstreden Amerikaanse af luisterprogramma. En de reactie daarop van veel Nederlanders die zeggen dat 'ze' alles van me mogen weten omdat ik toch niks te verbergen heb. Ik denk dat ik ook niks te verbergen heb (al weet ik dat van de afgelopen vijf jaar niet zo zeker), maar heb geen idee wat 'ze' dan allemaal te weten komen als mijn mailbox wordt geanalyseerd. Als Facebook op basis van een paar likes al weet wat mijn seksuele geaardheid is, wat zal er dan wel niet uit vijf jaar mailverkeer

#### Er ontstaan clusters van afzenders die iets met elkaar hebben

geconcludeerd kunnen worden? Welk beeld zal er van mij opdoemen bij iemand die me niet kent? Wat voor persoonlijkheid komt daar uit? Wellicht zaken die minder rooskleurig zijn, of zelfs iets waarop een overheid zou 'aanslaan'?

Met deze vragen als uitgangspunt had ik contact gelegd met Kraaij en Van Staalduinen, die het een aardige casus vonden. Suzan Verberne, postdoctoraal onderzoeker aan de Radboud Universiteit, en pro-

movendus Maya Sappelli trokken mijn mailbox leeg en analyseerden de ruim dertigduizend berichten: 86 megabyte tekst, exclusief attachments, die niet zijn bekeken. 'Als iemand deze hoeveelheid tekst 'handmatig' moet lezen', zegt Van Staalduinen, 'is hij ongeveer 500 uur kwijt.' Handmatig 'af luisteren' is alleen al om die reden onbegonnen werk. Daarom wordt de hulp ingezet van algoritmen die teksten kunnen analyseren. De computer van Verberne en Sappelli deed er een half uur over en stelde op basis van mijn inbox een netwerk samen van mensen met wie ik contact heb gehad.

Hoe gaat zoiets in zijn werk? Een belangrijke aanwijzing vormen de ontvangers die in de mails worden ge-cc'd, zegt Sappelli. Kennelijk zijn al de ontvangers binnen deze groep aan elkaar gelinkt. Als hetzelfde e-mailadres in meerdere cc's van verschillende afzenders voorkomt, is de bijbehorende persoon belangrijk en wordt zijn cirkel in het netwerkoverzicht groter. Zo iemand kan de spil zijn in een subnetwerk.

Ook wordt gekeken hoe vaak iemand mij mailt. Wie vaak contact legt, staat vermoedelijk dichterbij bij me en verschijnt daarom ook letterlijk dichterbij in het netwerk. Mensen die niet met elkaar communiceren zullen elkaar waarschijnlijk niet kennen. Zij worden daarom ondergebracht in een andere subgroep, waarvan opnieuw gekeken wordt wie met wie contact heeft. Op basis van deze vergelijkingen ontstaan clusters van afzenders die iets met elkaar hebben. Deze clusters hebben hun eigen kleur.

Met deze analyse is mijn sociale netwerk in kaart gebracht. Alleen zegt een overzicht van al deze mensen nog niet zo veel, zegt Van Staalduinen. 'Je ziet dat je kennelijk vrij veel communiceert met de Vereniging van Eigenaren. Als je weet dat dat niet belangrijk is, kun je stukken uitsluiten. Dan hoef je verder niet meer naar die VvE te kijken en kan worden ingezoomd op andere clusters.' Dat inzoomen betekent letterlijk kijken waarover wordt gemaaid.

Van mensen uit een cluster die dicht bij me staan of die als belangrijk worden beschouwd, kan per afzender een termenwolk worden gemaakt. Zo'n wolk, bekend van sommige zoekmachines op internet, toont allerlei veelgebruikte woorden. Hoe vaker een woord wordt gebruikt, hoe groter het wordt afgebeeld. 'Met een termen-

#### JE EIGEN NSA

**Wie zelf NSA'tje wil spelen: op internet staan allerlei hulpmiddelen die je berichtenverkeer op sociale netwerken als Facebook en Twitter analyseren.**

**Tweetgenie.nl**  
**Probeer een onderzoeksproject van Dong Nguyen van de Universiteit Twente, probeert op basis van Nederlandstalige tweets het geslacht en de leeftijd te raden van de twitteraar. Blijkt heel aardig te werken.**

**youarewhatyoulike.com**  
**Probeer persoonlijkheidskenmerken vast te stellen op basis van 'likes' op Facebook.**

**www.mapmyfollowers.com**  
**Geografische weergave van volgers.**

**twitonomy.com**  
**Geeft een overzicht van meestgenoemde twitteraars, vaakst gebruikte hashtags, vaakst geretweette gebruikers en hoe vaak je getweet bent. En hoe laat je meestal tuitert. En op welke dagen van de week.**