

*

Auteurs: Marc Meulemans en Jan Verplaetse
Bron: Skepter 11(1), maart 1998



*

De valkuilen van ons denken



De mens is niet zo rationeel als we graag zouden geloven. Studies hebben aangetoond dat we vaak op een systematische en voorspelbare wijze denkfouten maken. Onze redeneringen steunen niet zozeer op de regels van de logica, als wel op eenvoudige denkstrategieën die leiden tot vooroordelen en illusies.

Door de geschiedenis heen hebben tal van filosofen en logici zich over de geïdealiseerde wereld van de logica gebogen. Zij stelden zich de vraag hoe de mens 'zou moeten denken', en probeerden het correcte denken en redeneren in regels uit te drukken. Pas sinds de jaren '60 getroosten wetenschappers zich de moeite met experimenten na te gaan hoe het in werkelijkheid met ons redeneren is gesteld. Toonaangevend voor dit empirisch onderzoek is het werk van de Amerikaanse psychologen Amos Tversky en Daniel Kahneman. Hun publicaties uit de jaren '70 en '80 hebben inmiddels niet alleen in de cognitieve psychologie, maar ook in andere disciplines zoals rechtspraak, geneeskunde en economie het nodige stof doen opwaaien.

Tversky en Kahneman onderzochten hoe mensen in situaties van onzekerheid oordelen en beslissen. Het soort situaties dus waarmee we in het dagelijkse leven geconfronteerd worden. Of we nu de kwaliteit van een product inschatten, het talent van een sollicitant beoordelen, of de evolutie van de markt voorspellen, meestal beschikken we over onvoldoende informatie, en kunnen we aan gebeurtenissen slechts een bepaalde graad van waarschijnlijkheid toekennen. Tversky en Kahneman constateerden dat mensen bij dit soort problemen niet de regels van de logica en de waarschijnlijkheidsleer toepassen, maar wel vuistregels of strategieën die het hen mogelijk maken op een eenvoudige manier een oplossing te vinden. Deze vuistregels - 'heuristieken' - zijn soms efficiënt, maar kunnen ons vaak ook op een dwaalspoor zetten, doordat we ze niet op de juiste manier of niet bij het juiste soort probleem toepassen.

De concepten of categorieën die we in het dagelijks leven hanteren steunen in belangrijke mate op typische voorbeelden, prototypes of stereotypes genoemd. Als we bijvoorbeeld aan een hond denken, denken we niet aan een chihuahua, maar veeleer aan een Duitse herdershond, een terriër, of een ander ras dat er opvallend 'hondachtig' uitziet. Of we geneigd zijn een object in een bepaalde categorie onder te brengen, hangt dan ook af van de mate waarin het gelijkenis vertoont met - of representatief is voor - een typisch object uit die categorie. Tversky en Kahneman noemen dit de heuristiek van representativiteit.

Met het redeneren aan de hand van typische voorbeelden is niets mis, zolang we ook

rekening houden met andere vormen van relevante informatie. En daar loopt het mank. Zo hebben we niet of onvoldoende oog voor de statistische verdeling van een eigenschap binnen een populatie. Tversky en Kahneman illustreerden dit met het volgende vraagstukje: 'Ik heb een vriend die professor is. Hij houdt van het schrijven van gedichten, is nogal verlegen en is klein van gestalte. Op welk terrein is hij werkzaam: Chinese literatuur of psychologie?'

De meerderheid antwoordt hier 'Chinese literatuur', vanuit de veronderstelling dat het schrijven van poëzie eerder een typisch kenmerk is voor een professor literatuur dan voor een professor psychologie. Deze redenering is niet fout, maar wel onvolledig. Want je dient ook aandacht te hebben voor de statistische verhouding binnen de populatie, met name het relatieve aantal professoren in beide disciplines. Aangezien er veel meer professoren werkzaam zijn in de psychologie dan in de Chinese literatuur (en wellicht ook sommige professoren psychologie poëzie schrijven) is het waarschijnlijker dat het hier om een hoogleraar psychologie gaat.

Het denken in termen van representativiteit en stereotypes leidt soms tot intuïties die elke logica tarten. Ziehier een persoonsbeschrijving die door Tversky en Kahneman aan een aantal studenten werd voorgelegd. 'Linda is 31 jaar oud, alleenstaand, openhartig en bijzonder schrand. Ze is afgestudeerd in de filosofie. Als studente was ze enorm begaan met rassendiscriminatie en sociale onrechtvaardigheid, en nam ze actief deel aan anti-nucleaire demonstraties.' De studenten kregen de opdracht acht beweringen over mogelijke hobby's en beroepen van Linda te rangschikken, van zeer waarschijnlijk naar zeer onwaarschijnlijk. Tot deze lijst behoorden onder meer:

A: Linda werkt op een bank.

B: Linda werkt op een bank en is actief in de vrouwenbeweging.

De meeste studenten vonden bewering B waarschijnlijker dan bewering A. De reden daarvoor moet niet ver worden gezocht. De persoonsbeschrijving van Linda stemt niet overeen met het profiel van een typische bankbediende, zodat de eigenschap 'bankbediende' onwaarschijnlijk wordt geacht. De mogelijkheid 'feministische bankbediende' daarentegen voldoet al heel wat meer aan de persoonsbeschrijving, en dus wordt daar een grotere kans aan toegekend. (1)

Nochtans is deze redenering een zinsbegoocheling. Uitspraak B kan onmogelijk waarschijnlijker zijn dan uitspraak A. Want de vrouwen die bankbediende én lid van de vrouwenbeweging zijn, vormen een deelgroep van de vrouwen die bankbediende zijn. Algemeener geformuleerd is het dus zo dat de kansen dat X zich voordoet, altijd hoger liggen dan de kansen dat X en Y zich samen voordoen (als X en Y bijvoorbeeld elk een kans van 1 op 10 hebben en onafhankelijk zijn, dan is de kans dat beide zich voordoen slechts 1 op 100). Hier laten de meesten zich evenwel misleiden door de representativiteit.

Opmerkelijk in deze studie was dat studenten die een hogere opleiding in statistiek en kansrekening hadden genoten, het amper beter deden dan anderen: 85 procent van de 'experts' en 90 procent van de 'naïeve' studenten was hier uit de bocht gegaan. Eenzelfde soort uitglijder werd overigens ook bij concrete problemen uit diverse kennisdomeinen waargenomen. Zo toonde een vergelijkbare test aan dat heel wat artsen de aanwezigheid van twee symptomen (bijvoorbeeld migraine en braken) waarschijnlijker achten dan de aanwezigheid van een van de twee afzonderlijk.

Fata morgana

Ook de omvang van een steekproef - cruciaal in statistiek en kansrekening - zorgt voor valkuilen in onze redeneringen. Neem de volgende denkpuzzel: 'In een groot hospitaal worden gemiddeld 45 baby's per dag geboren, en in een klein hospitaal bedraagt dat gemiddelde 15 per dag. Het gemiddelde aantal jongetjes is uiteraard ongeveer 50 procent,

maar dat percentage varieert per dag: soms is dat meer, soms is dat minder. In beide hospitalen wordt het aantal dagen geteld waarop het percentage jongetjes 60 procent bedraagt. Welk hospitaal heeft meer van die dagen: het kleine, het grote, of tellen ze er beide evenveel?'

Niet weinigen zijn van oordeel dat er in beide hospitalen evenveel van die dagen zijn. Helaas fout. In het kleine hospitaal zijn er meer van die dagen, omdat je in een kleine steekproef (15 geboorten per dag) meer kans hebt op een afwijking van het gemiddelde - 60 procent in plaats van 50 procent - dan in een grote steekproef (45 geboorten per dag).

Dit soort vergissing wordt wel eens de wet van de kleine getallen genoemd, met name de neiging om een zeer beperkte steekproef uit een populatie als representatief voor de hele populatie te beschouwen. Toegepast op kansspelen betekent dit dat we een korte willekeurige reeks representatief achten voor een lange reeks die door hetzelfde toevalsmatige proces wordt gegenereerd. Bij een willekeurige reeks kruis (K) of munt (M) bijvoorbeeld, denken velen dat de kenmerken van een lange reeks - willekeurige volgorde en gelijke kansen voor kruis of munt - ook tot uiting komen in een klein onderdeel van die reeks. Daarom vinden ze de korte reeks KMKMMK waarschijnlijker dan de reeks KKKMMM, die er te geordend uitziet, en ook waarschijnlijker dan de reeks KKKKMK, die te onevenwichtig lijkt. De wet van de kleine getallen is een fata morgana: ofschoon een lange reeks ongeveer evenveel kruis als munt zal bevatten, kan een klein onderdeel van die reeks een sterke afwijking van het statistisch gemiddelde vertonen. De drie zojuist vermelde korte reeksen zijn dan ook even waarschijnlijk, en hebben elk een kans van 1 op 64 (2 tot de 6de macht, aangezien zes keer werd opgegooid).

De wet van de kleine getallen verklaart ook de 'drogredenering van de gokker'. Menig gokker zal in een roulettespel op zwart spelen na een reeks rood, vanuit de overtuiging dat de kansen op zwart dan zijn toegenomen. Maar de kans blijft bij elke inzet uiteraard 50 procent (de nul niet meegerekend). Het bolletje heeft geen geheugen, net zoals de gokker blijkbaar, die er maar niet in slaagt lessen te trekken uit z'n in het verleden opgestapelde verliezen. De illusoire wet van de kleine getallen manifesteert zich overigens in uiteenlopende contexten. Zo durven wetenschappers al eens de representativiteit van een kleine steekproef te overschatten.

Of neem dit voorbeeld: toen Londen tijdens de Tweede Wereldoorlog werd gebombardeerd, dachten de inwoners dat er op welbepaalde plaatsen werd gemikt, omdat sommige delen van de stad meer werden geraakt dan andere. Een statistische studie wees evenwel uit dat de bombardementen willekeurig verliepen. Alleen was de bommenreeks 'te kort' om een gelijkmatige spreiding van de schade te verkrijgen. Ook pseudo-wetenschappers zijn tuk op dit soort drogredenering. Wanneer een helderziende of andere paranormaal begaafde zich aan een beperkte proef waagt, behaalt hij niet zelden een score die boven het toevalsgemiddelde ligt. Een statisticus kijkt daar niet van op, maar een pseudo-wetenschapper meent hier al vlug magische krachten aan het werk te zien.

Hoe beperkter de steekproef, hoe meer de representativiteit ervan in het gedrang komt. Dit statistische gegeven mag al evident en bovendien door de meesten gekend zijn, in onze dagelijkse redeneringen liggen we er niet wakker van. Zelfs één voorbeeld lijkt ons te volstaan om algemene conclusies te trekken. Deze onverkwikkelijke denkfout werd op een griezelijk nauwkeurige wijze door de Amerikaanse psycholoog Ruth Hamill en z'n collegae geregistreerd. Hamill liet twee groepen proefpersonen een video-opname van een geënceneerd interview met een zogenaamde cipier zien. De eerste groep zag een cipier die in het interview als een zachtaardig persoon naar voor kwam, terwijl de tweede groep kennis maakte met een onbetamelijke bruto.

Beide groepen werden vervolgens nog eens in drie subgroepen verdeeld: een eerste groep werd verteld dat de cipier op de video een typische cipier was, de tweede groep kreeg zeggegd een atypisch of extreem geval te zien, de derde groep werd daarover in het ongewisse gelaten. Nadien moesten ze een vragenlijst invullen die onder meer enkele vragen over de attitudes van Amerikaanse cipiers bevatte. De resultaten waren schokkend. Er bleek dat de drie groepen zich op basis van het interview eenzelfde beeld over cipiers in het algemeen hadden gevormd, ongeacht de informatie die ze over de representativiteit van de geïnterviewde hadden gekregen. Als de geïnterviewde cipier bijvoorbeeld hardvochtig was, concludeerde elke groep dat alle cipiers zo waren. Zelfs diegenen die gewaarschuwd waren dat het voorbeeld atypisch was.

Lof voor landing

Een ander statistisch principe dat we maar al te vaak over het hoofd zien, is de regressie naar het gemiddelde. Dit concept klinkt wellicht obscuur, maar een illustratie kan het verhelderen. Als een individu op verschillende dagen gelijkwaardige IQ-tests aflegt, zullen de resultaten variëren. De score op zo'n test wordt immers ook beïnvloed door factoren als vermoeidheid, motivatie, externe omstandigheden, en niet te vergeten stom toeval dat men de oplossing van een puzzeltje wel of niet meteen ziet zitten. Daardoor zijn de testresultaten niet zo betrouwbaar. Dit heeft tot gevolg dat als we een groep mensen selecteren met dezelfde hoge of lage score, die groep bij een volgende (of eerdere!) test gemiddeld wat dichter bij het algemeen gemiddelde blijkt te scoren. Hoe zwakker het verband tussen meting en hermeting is, des te sterker is dit effect. Enige skepsis is dus op zijn plaats als we horen dat een of ander bijspijkerprogramma de scores van een groep zwakke leerlingen heeft verbeterd.

Ofschoon regressie naar het gemiddelde al eind vorige eeuw ontdekt werd, laten we er ons telkens door om de tuin leiden. Zo bracht een onderzoek aan het licht dat Israëlische vlieginstructeurs uit hun ervaringen een eigen theorie over beloning en straf hadden gedistilleerd. Ze waren er namelijk van overtuigd geraakt dat lof bij een extreem goede landing nadien tot slechtere landingen, en harde kritiek bij een extreem slechte landing nadien tot betere landingen leidt. Hun conclusie was fout, omdat een schitterende landing doorgaans door een minder goede landing wordt gevolgd, zelfs wanneer de instructeur zich van alle commentaar onthoudt. De instructeurs begingen dus de vergissing dat ze de daaropvolgende slechtere prestatie aan hun lovend commentaar toeschreven, in plaats van aan het evidente gegeven dat ook een leerling-piloot zichzelf niet elke dag overtreft.

Rad van fortuin

Een andere heuristiek waar we ons door laten leiden is de beschikbaarheid van informatie in het geheugen. 'Konden wij weten wat de mensen het meest geneigd zijn zich te herinneren, dan zouden wij weten wat ze het meest geneigd zijn te doen,' zei Lord Halifax, die misschien niet de term 'heuristiek', maar des te beter de menselijke psyche begreep. In termen van waarschijnlijkheden komt de heuristiek van beschikbaarheid erop neer dat we een gebeurtenis waarschijnlijker achten naarmate we ons makkelijker een voorbeeld van die gebeurtenis voor de geest kunnen halen. Tot op zekere hoogte een gerechtvaardigde denkstrategie, want hoe frequenter een gebeurtenis, hoe makkelijker we ons een voorbeeld ervan kunnen herinneren. Maar herinneringen hangen uiteraard ook van andere factoren af. Dat bijvoorbeeld Romeo Julia niet uit zijn gedachten kon bannen, had meer met verliefdheid dan met objectieve frequenties te maken. Door de invloed van die andere factoren wordt de heuristiek van beschikbaarheid hoogst onbetrouwbaar.

Tversky en Kahneman demonstreerden dit met een experiment waarin een aantal proefpersonen een lijst werd voorgelezen. Deze lijst bevatte meer vrouwen- dan

mannennamen, maar de meeste mannen die erin voorkwamen waren bekende figuren. Wanneer een tijd later werd gevraagd of de lijst meer vrouwen- dan wel mannennamen bevatte, was het antwoord overwegend 'mannennamen'. Bekende namen kunnen we ons nu eenmaal makkelijker herinneren dan onbekende. Niet alleen bekende, maar ook opvallende of emotioneel geladen fenomenen kunnen we ons sneller voor de geest halen. Getuige zijn van een brand, maakt meer indruk dan het lezen van een krantenartikel over die brand. En een bericht over een vliegtuigcrash laat in ons geheugen diepere sporen na dan een bericht over een auto-ongeval. Dat is een van de redenen waarom we ons in de auto veiliger voelen dan in het vliegtuig, ofschoon de risico's op de weg veel groter zijn dan in de lucht. Niet minder irrationeel is onze gewoonte tijdelijk voorzichtiger te gaan rijden nadat we van nabij getuige zijn geweest van een tragisch auto-ongeluk. Net alsof de kansen op een ongeval dan plots zienderogen zijn toegenomen.

Rationeel is het niet, maar spectaculaire en emotionele evenementen vinden we als bewijsmateriaal overtuigender dan statistische frequenties. Een dramatisch beeld op televisie heeft dan ook meer impact dan een uitvoerig betoog met objectieve cijfers en redeneringen. De invloed van de media kan hier moeilijk worden overschat. De feiten die de krantenkoppen en het televisiejournaal halen, vinden we waarschijnlijker dan sommige minder ophefmakende, maar frequenter voorkomende gebeurtenissen. In de VS vroegen onderzoekers een aantal mensen welke van beide doodsoorzaken relatief het meest voorkwamen: moord of zelfmoord, zelfmoord of diabetes, borstkanker of diabetes. De meerderheid stipte telkens de eerste doodsoorzaak aan, ofschoon het in werkelijkheid telkens om de tweede ging. Maar de eerste doodsoorzaak is inderdaad meer emotioneel geladen, en krijgt meer aandacht in de media.

De manier waarop we een probleem oplossen, is vaak afhankelijk van de wijze waarop dat probleem wordt geformuleerd of omkaderd. Psychologen spreken hier van het 'omkaderen van keuzes'. Dat ook hooggeschoolden zich door deze heuristiek laten beetnemen, mocht blijken toen Amerikaanse artsen werd gevraagd of ze in een specifiek geval al dan niet een chirurgische ingreep zouden aanbevelen. Als hen werd verteld dat de ingreep tot een sterftecijfer van 7 procent (binnen de vijf jaar na de operatie) leidde, dan bevalen ze die ingreep slechts met enige aarzeling aan. Wanneer hen echter werd gezegd dat de overlevingskans 93 procent bedroeg, dan waren ze veel meer geneigd de ingreep warm aan te bevelen. Of het glas als halfvol of als halfleeg wordt beschreven, maakt voor onze hersenen dus wel degelijk een verschil uit.

Niet minder verraderlijk is de zogenaamde verankering, het vasthouden aan een beginindruk. Bij het schatten van kwantitatieve grootheden hebben we de kwalijke gewoonte een eerste schatting onvoldoende in het licht van nieuwe informatie te corrigeren. Dit soort mentale inertie kwam onder meer tot uiting in een experiment van Tversky en Kahneman. Een eerste groep deelnemers kreeg vijf seconden de tijd om het product $2 \times 3 \times 4 \times 5 \times 6 \times 7 \times 8$ te schatten, terwijl een tweede groep als opdracht $8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2$ kreeg. Aangezien alleen de volgorde van de getallen verschilt, is de uitkomst in beide gevallen dezelfde, namelijk 40.320. De schattingen van de eerste groep bedroegen evenwel gemiddeld 512, die van de tweede groep 2.250.

De verklaring hiervoor is eenvoudig: bij deze opdracht worden alleen de eerste twee of drie producten berekend (in het eerste geval $2 \times 3 \times 4$), en doet men vervolgens, bij gebrek aan tijd, een gok naar het eindresultaat. Het beginresultaat wordt onvoldoende naar boven toe gecorrigeerd. Vandaar ook dat de tweede groep met hogere getallen (maar wel veel lager dan het echte resultaat) komt aandragen dan de eerste groep. Het beginresultaat van de dalende reeks ($8 \times 7 \times 6$) ligt immers hoger dan dat van de stijgende reeks ($2 \times 3 \times 4$). Deze denkfout valt misschien nog te begrijpen, maar helemaal grotesk wordt het wanneer een volslagen irrelevante beginwaarde ons in haar greep houdt. Zo werd in een experiment in de

aanwezigheid van proefpersonen een willekeurige waarde tussen 1 en 100 vastgesteld door aan een soort 'rad van fortuin' te draaien. Vervolgens werd de proefpersonen gevraagd hoeveel procent van de Afrikaanse landen lid is van de Verenigde Naties. Ofschoon de proefpersonen er uitdrukkelijk op gewezen werden dat er geen enkel verband bestond met het willekeurig aangeduide getal, lieten ze zich toch door dit getal beïnvloeden. Wanneer het eerste getal bijvoorbeeld 12 was, kwamen ze tot lagere schattingen dan wanneer het getal 92 was.

Dodenaantal

Verwant met de verankering is de neiging aan opinies en opvattingen vast te houden. Psychologen spreken hier van de confirmation-bias ('confirmatie-vooordeel'), omdat we doorgaans op zoek gaan naar een bevestiging, en niet naar een ontkenning van onze vermoedens of overtuigingen. De drang tot confirmatie zit diep in onze hersenen geworteld. Ook wanneer we ten aanzien van een te testen hypothese geen voorkeuren of verwachtingen koesteren, zullen we vooral naar feiten zoeken die de hypothese bevestigen.

Dit werd onder meer aan het licht gebracht door de Amerikaanse psychologen Mark Snyder en William Swann. Ze gaven een eerste groep proefpersonen de opdracht met een interview uit te zoeken in welke mate een bepaald persoon introvert was. Een tweede groep daarentegen moest nagaan in welke mate de geïnterviewde extravert was. De vragen die zij mochten stellen, dienden te worden gekozen uit een lijst met drie soorten vragen, namelijk 'introverte' vragen (zoals: 'Waarom is het moeilijk om spontaan met mensen om te gaan?'), 'extraverte' vragen (zoals: 'Hoe breng je wat meer sfeer op een party?'), en een aantal neutrale vragen. Zij die de introversie-hypothese testten, selecteerden vooral introverte vragen, terwijl de anderen duidelijk meer extraverte vragen stelden. Beide groepen zochten dus bewijsmateriaal dat hun hypothese bevestigde, en vonden dat ook. Hoe kan het ook anders? De geïnterviewde kan bezwaarlijk antwoorden dat hij een feestje opvrolijkt door afwezig in een hoekje te zitten kniezen. Met andere woorden, extraverte vragen leiden onvermijdelijk tot extraverte antwoorden. Door hoofdzakelijk voor één type vragen te kiezen, creëerden de vragenstellers een realiteit die in overeenstemming was met hun hypothese.

Wat gebeurt er als we over een discussiethema wél al een uitgesproken mening hebben? Dat we in zo'n situatie ijverig naar bevestiging zoeken, lag in de lijn der verwachtingen. Niettemin hadden onze hersenkronkels hier voor de psychologen nog een verrassing in petto. In een studie van de Amerikaanse psycholoog S. Plous kregen fervente voor- en tegenstanders van kernenergie dezelfde krantenartikelen te lezen over het ongeluk in de kerncentrale van Three Mile Island, dat een balans opleverde van een oversterfte van circa 200 (een statistisch verwachte totale toename van sterfte in de gehele wereld over een lange periode) en nog eens 1000 extra mensen wereldwijd met ziekteverschijnselen ten gevolge van bestraling. Aangezien beide partijen met hetzelfde feitenmateriaal werden geconfronteerd, zouden we kunnen verwachten dat hun visies na het lezen van de krantenartikelen dicht bij elkaar kwamen te liggen. Maar niets was minder waar. Integendeel, de meerderheid van zowel voor- als tegenstanders nam na het lezen van de artikelen nog extremere standpunten in. Gevraagd naar hun argumenten, wezen de contra's op de gevaren van kernenergie. De pro's daarentegen vonden dat de schade en het dodenaantal tengevolge van het ongeluk op Three Mile Island al bij al meevielen, en benadrukten de aanwezigheid van veiligheidssystemen.

Dubbelzinnige gebeurtenissen hebben veel weg van inktvlekken waarin we alleen zien wat we wensen of verwachten te zien. Maar wat met informatie die ontegensprekelijk met onze opvattingen in strijd is? Met deze vraag voor ogen verstrekke de Amerikaanse psycholoog Craig Anderson een aantal proefpersonen valse informatie. Een eerste groep kreeg

beschrijvingen van individuele gevallen waaruit bleek dat mensen die risico's nemen betere brandweperlui zijn dan anderen. De tweede groep las case studies waaruit het omgekeerde bleek. Vervolgens kregen beide groepen de opdracht een theorie of een verklaring te vinden voor het gesuggereerde verband tussen het nemen van risico's enerzijds en het geschikt zijn als brandweerman anderzijds. De eerste groep kwam met verklaringen als: 'Een brandweerman die te voorzichtig is, aarzelt en is verloren', terwijl de tweede groep uitspraken deed in de trant van: 'Een brandweerman moet nadenken voor hij iets onderneemt'.

Nadien onthulde de leider van het experiment dat de verstrekte informatie verzonnen en fout was. Dat maakte evenwel weinig indruk. Gevraagd naar hun opvattingen over brandweperlui, bleven beide groepen koppig bij hun theorie. Iemand een theorie laten construeren is kennelijk heel wat makkelijker dan hem er nadien van overtuigen dat die theorie niet klopt. Maar er is ook goed nieuws. Als je - zo blijkt uit onderzoek - mensen ertoe aanspoort een alternatieve theorie te evalueren, klampen ze zich minder weerbarstig aan hun oorspronkelijke visie vast.

Attributiefout

We zijn geen experts in de logica of de waarschijnlijkheidsleer, en evenmin wetenschappers die nauwgezet en kritisch hypothesen testen. Maar hoe zit het met onze zelfkennis? Ook hier lijken we aan enige bijscholing toe. Mensen hebben namelijk een weinig realistisch beeld van zichzelf. Om niet te zeggen dat ze zichzelf behoorlijk overschatten. In een klassiek onderzoek kregen mensen de opdracht een reeks kennisvragen te beantwoorden, en om bij elk antwoord aan te geven hoe zeker ze van hun stuk waren (ze konden dus voor 100 procent of bijvoorbeeld voor slechts 80 of 50 procent zeker zijn). Van de antwoorden waarvan de deelnemer absoluut (voor 100 procent) zeker wist dat ze juist waren, bleek uiteindelijk 17 procent niet correct te zijn. Verontrustend in deze en andere studies is ook de constatering dat de zelfoverschatting toeneemt naarmate iemand meer over het onderwerp in kwestie weet. Hoe kundiger we op een bepaald terrein zijn, hoe groter de discrepantie tussen wat we weten en wat we denken te weten, wordt.

Onze blik op de toekomst wordt evenmin gehinderd door realiteitszin. Studies rapporteren dat studenten hun kansen op succes in een job, een hoog salaris, een geslaagd huwelijk, het hebben van begaafde kinderen, enz. sterk overschatten. De kansen op alcoholverslaving, ziekten en depressies daarentegen worden miniem geacht. Welke dimensie ook ter sprake komt - gezondheid, geluk, succes, materiële welvaart - de meeste mensen lijken er niet aan te twijfelen dat ze een rooskleuriger toekomst tegemoet treden dan de gemiddelde medemens. Een onderzoeker kan uiteraard moeilijk nagaan in hoeverre de verwachtingen van één bepaald persoon onrealistisch zijn. Maar het is statistisch bekeken enigszins verdacht wanneer de meerderheid erop vertrouwt tot de 'happy few' te zullen behoren. Psychologen hebben uit een en ander dan ook geconcludeerd dat zelfoverschatting en onrealistisch optimisme niet alleen normaal, maar voor het psychisch evenwicht ook noodzakelijk zijn. Een conclusie die overigens aansluit bij de observatie dat depressieve mensen doorgaans een realistischer zelfbeeld hebben dan diegenen die gezwind door het leven gaan.

Ook wanneer we anderen beoordelen, kunnen we de plank behoorlijk misslaan. Een veel voorkomende vergissing is wat in vaktermen de fundamentele attributiefout wordt genoemd. Deze bestaat erin dat we de oorzaken van andermans gedrag aan de persoon zelf toeschrijven, en niet - of onvoldoende - aan de situatie. Met andere woorden, we stellen mensen verantwoordelijk voor hun daden, en onderschatten de invloed van de omgeving op die daden. Wanneer we ons eigen gedrag beoordelen, doet zich veeleer het omgekeerde voor. Dan hebben we wel meer oog voor de situatie, en minimaliseren we onze eigen

inbreng of verantwoordelijkheid. Ofschoon ook hier onze behoefte aan een positief zelfbeeld een rol speelt: positief of verdienstelijk gedrag schrijven we graag toe aan onze talenten, terwijl we bedenkelijke prestaties doorgaans aan de omstandigheden wijten.

Hoezeer de attributiefout onze kijk op de medemens kan vertroebelen, bleek onder meer toen de Amerikaanse psycholoog Lee Ross en zijn collegae onder studenten een quiz organiseerden. Een eerste groep moest tien moeilijke, maar niet onmogelijke vragen stellen. Een tweede groep moest die vragen beantwoorden, waarop de eerste groep zei of het antwoord juist of fout was. Een derde groep bestond uit waarnemers die nadien de parate kennis van de beide groepen dienden te evalueren. Zoals verwacht onttaardde de quiz al gauw in een demonstratie van de eigen kennis en kunde. De vragenstellers kozen onderwerpen uit het eigen interessegebied, en kwamen voor de dag met naar het einde toe alsmaar esoterischer vragen, zoals 'Wat is de naam van de langste ijstijd?' of 'Wie was de eerste gouverneur van Idaho?'

Weinig verwonderlijk dat diegenen die het vuur aan de schenen werd gelegd, gemiddeld maar vier van de tien vragen juist konden beantwoorden. Des te opmerkelijker was dan ook hoe de waarnemers de algemene kennis van beide partijen inschatten. Ze bedachten de vragenstellers met een puntenscore van acht op tien, en vonden dat de beantwoorders niet meer dan vijf op tien verdienden. Wat uiteraard absurd is. Bij dit soort spel kan de vragensteller ongehinderd pronken met zijn kennis, en slaat diegene die antwoordt onvermijdelijk een modderfiguur. De waarnemers hielden evenwel geen rekening met deze onfaire situatie, en schreven de prestaties toe aan de kennis van beide partijen.

Dat we bij het beoordelen van anderen minder aandacht besteden aan situationele factoren dan bij het beoordelen van onszelf, heeft onder meer te maken met het gezichtspunt. We observeren namelijk, als we iets ondernemen, niet onszelf maar besteden vooral aandacht aan de omgeving. Wanneer we daarentegen anderen gadeslaan, focussen we op de persoon. De impact van het gezichtspunt op de beoordeling kwam onder meer aan het licht in een - overigens ook in juridisch opzicht opzienbarend - experiment van de Amerikaanse psychologen Daniel Lassiter en Audrey Irvine. Lassiter en Irvine onderzochten hoe het oordeel van een jury door video-opnames van een ondervraging wordt beïnvloed. Wat geen louter theoretische aangelegenheid was. Advocaten van verdachten protesteerden tegen het afdwingen van bekentenissen en eisten controle van het verhoor, zodat in de praktijk ondervragingen vaak op videoband werden vastgelegd.

De twee psychologen registreerden een geacteerde bekentenis vanuit drie verschillende cameraposities, zo gekozen dat ofwel het gezicht van de ondervrager, ofwel dat van de ondervraagde, ofwel dat van beiden in profiel in beeld kwam. Deze opnamen werden vervolgens getoond aan drie verschillende groepen die als jury optraden. Zij moesten uitmaken of de tijdens het verhoor verkregen bekentenis vrijwillig dan wel onder druk van de ondervrager tot stand was gekomen. De resultaten lieten weinig aan de verbeelding over. Ofschoon iedereen hetzelfde gesprek had gevolgd, liepen de conclusies van de drie groepen sterk uiteen. Diegenen die op video de verdachte hadden gezien, vonden de bekentenis veeleer vrijwillig. Zij die daarentegen de ondervrager in beeld hadden gehad, meenden dat er dwang in het spel was. De derde groep, die beide gezichten in profiel had gezien, nam een tussenpositie in.

De meeste leden van onze soort mogen dan al uitblinken in optimisme, de sociale en cognitieve psychologen zien daar weinig reden toe. Vooroordelen en illusies maken in onze redeneringen de dienst uit, en dat zullen we geweten hebben. Propaganda, pseudo-wetenschap, sektevorming, stereotypering en discriminatie van rassen, relationele problemen, enz., de lijst van fenomenen die op irrationele denkmechanismen gedijen, is lang en vrijwel onuitputtelijk. Niettemin ontwaren de psychologen ook licht in de duisternis.

Naarmate we meer contact hebben met anderen, zo melden ze, wordt ons oordeel over hen accurater. En hoewel ons algemene beeld van mensen uit onze omgeving weinig nauwkeurig is, kunnen we doorgaans goed voorspellen hoe ze zich in onze aanwezigheid - bijvoorbeeld op het werk - zullen gedragen. Sommige studies signaleren bovendien dat mensen rationeler denken wanneer hun eigen belangen op het spel staan. *Last but not least* is er ook de weliswaar bescheiden, maar niettemin positieve impact van leerprogramma's die het accent leggen op kritisch denken in concrete situaties. Sommigen pleiten er dan ook voor om in het onderwijs meer aandacht te besteden aan het ontmaskeren van cognitieve illusies en drogredeneringen. Rationeel denken is immers geen spontane gave, maar een capaciteit die moeizaam op het duistere continent van de irrationaliteit moet worden veroverd.

Noot

1. Bij besprekingen van het voorbeeld van 'Linda de bankbediende' protesteren briefschrijvers vaak dat de betekenis van mededelingen contextafhankelijk is, en dat de context van mogelijkheid B de interpretatie bij A 'is niet actief in de vrouwenbeweging' in de hand werkt. Zo ook in *Skepter*. Tversky en Kahneman wisten dit echter, en hadden met aanvullende proeven duidelijk gemaakt dat het eerder zo is dat het stereotype 'bank dus niet bij vrouwenbeweging' al aanwezig was bij hun proefpersonen, en dat dit stereotype alleen in de context van B wordt uitgeschakeld. (noot toegevoegd voor de Skepsis-site).

[Terug](#).

Een eerdere versie van dit artikel verscheen in *Eos*, januari 1998. In het boek *Inevitable Illusions* van Massimo Piatelli-Palmarini (Wiley, 1994) wordt een groot deel van het hier besprokene uitvoerig behandeld. Ook in *How we know what isn't so*, van Thomas Gilovich (Prometheus, 1991) worden een aantal van deze thema's behandeld, met toepassingen op geloof in het wonderbaarlijke.

Enkele concrete illusies en drogredenen

De drogreden van de jury

Kansen en waarschijnlijkheden schatten we dagelijks in, maar oefening lijkt hier weinig kunst te baren. Een veel voorkomende denkfout is bijvoorbeeld dat we onvoldoende rekening houden met de statistische verdeling van een eigenschap binnen de populatie. Zelfs wanneer we over exacte cijfers beschikken, lijken we niet goed te weten wat ermee aan te vangen. De Amerikaanse psychologen Amos Tversky en Daniel Kahneman illustreerden dit met de 'drogreden van de jury':

In een stad zijn er twee taximaatschappijen. De taxi's van de ene maatschappij zijn blauw, die van de andere groen. In de nacht dat een taxichauffeur na een ongeval doorreed, was 85 procent van de rijdende taxi's blauw, en dus 15 procent groen. Volgens een ooggetuige reed de dader met een groene taxi. Onderzoek wijst uit dat een getuige in gelijkaardige omstandigheden in 80 procent van de gevallen een groene van een blauwe taxi weet te onderscheiden, namelijk 20 procent van de blauwe voor groen en 20 procent van de groene voor blauw. Zijn getuigenis is met andere woorden voor 80 procent betrouwbaar. Wat is nu, geacht jurylid, het meest waarschijnlijk: dat de schuldige met een groene, dan wel met een blauwe taxi reed?

De spontane reactie van velen luidt dat de taxi wellicht groen was. Dat betekent dat ze de hoge betrouwbaarheid (80 procent) van de getuige sterker laten doorwegen dan het relatieve aantal blauwe taxi's (85 procent). Een eenvoudige berekening toont nochtans dat het correcte

antwoord 'blauw' is. Als de getuige in die nacht honderd taxi's heeft gezien, dan zijn daarvan naar verwachting 15 groen en 85 blauw. Van deze zal hij er 12 respectievelijk 17 groen noemen. Dus van de 29 keer dat de getuige zegt 'het was een groene' zou het 17 maal een blauwe zijn geweest. De kans dat de taxi blauw was, gegeven het feit dat de getuige 'groen' zegt, was $17/29$, dus 0,59 oftewel 59 procent.

Dit voorbeeld is wat gestroomlijnd, en we mogen hopen dat we nooit door jury's zullen worden beoordeeld die vinden dat een foutkans van 20 procent te klein is om 'redelijke twijfel' te rechtvaardigen. Een soortgelijk voorbeeld gaat over een medische test voor het een of ander (het doet er niet toe of het baarmoederhalskanker, aids, zwangerschap of de ziekte van Pfeiffer is). Laten we uitgaan van de feitelijke getallen uit 1966 voor borstkanker en onderzoek met een mammogram. De kans op een positieve uitslag als de ziekte aanwezig is, bedroeg toen 0,79. De kans op een foute uitslag als de ziekte afwezig is bedroeg 0,10. De verhouding tussen aantallen correct-positieve en fout-positieve uitslagen is dus 79:10. Wat kunnen we zeggen als de test 'positief' uitvalt? Dat hangt erg af van hoe groot de kans daarop was voor de test gedaan werd. Laten we aannemen dat oorspronkelijk de kansen tussen 'heeft het' en 'heeft het niet' zich verhouden als P staat tot Q . In de situatie met de taxi's was dat 15:85 en de 'test' was de getuigenis, en bij borstkanker is dat ongeveer 1:99. We kunnen nu uitrekenen wat gegeven een positieve testuitslag de nieuwe kans is. Er komt in het geval van borstkanker uit dat de nieuwe verhouding $79 \times P$ staat tot $10 \times Q$ is, dus 79:990, hetgeen met een kans van 7,4 procent correspondeert.

Als je echter om wat voor reden al van oordeel bent dat de kansen op wel of geen kanker zich verhouden als 1:1, dan maakt de testuitslag daar 79:10 van. Zo'n gedachte zal bij borstkanker niet vlug opkomen, maar bij aids zijn er verschillende risicogroepen met sterk uiteenlopende a priori risico's, en mensen die een test ondergaan doen dat ook niet zomaar. Dit geeft ook aan waar de problemen liggen bij de toepassing van deze kennis: om de uitslag te beoordelen moet je weten hoe groot je risico was voor de test, en dat is heel vaak een intuïtieve schatting die sterk door emoties vertekend kan zijn.

Deze berekeningsmethode heet naar Thomas Bayes (1702-1761), en ze wordt ook vaak in situaties toegepast waarin de 'kansen vóór de test' helemaal geen kansen zijn maar geloven. In feite stelde Bayes voor om als je helemaal niets weet, de 'kansen' van tevoren maar allemaal gelijk te veronderstellen, dus in dit geval de 'kans' op borstkanker op 0,5 te stellen (je hebt het of je hebt het niet).

De zojuist geschetste rekenregel werkt ook vlot voor het geval van de taxi's: de kansverhouding groen staat tot blauw was 15:85; de verhouding goed staat tot fout was 4:1, dus na de getuigenis verhouden de kansen op groen en blauw zich als 60:85.

De illusie van de historicus

Konden de verantwoordelijken op basis van de informatie waarover zij beschikten voorspellen wat er zou gebeuren? Het is een vraag die tegenwoordig verscheidene parlementaire onderzoekscommissies bezighoudt. Psychologen hebben zich afgevraagd hoe we dit soort vragen beantwoorden. En ook hier konden ze onze intuïtie op vooroordelen betrapten. Historici worden er wel eens van beschuldigd gebeurtenissen uit het verleden als onvermijdelijke - en dus voorspelbare - ontwikkelingen te zien. Wat dit betreft zit in elk van ons een kleine historicus verborgen. Diverse studies tonen namelijk aan dat we na de feiten de voorspelbaarheid ervan in niet geringe mate overschatten.

In een reeks experimenten werd aan twee groepen een aantal beschrijvingen gegeven van echt gebeurde feiten, zoals het uitbarsten van een conflict, het faillissement van een bedrijf, het succes van een product, enz. Daarnaast kregen ze ook informatie over wat daaraan

voorafging. Maar er zat, zoals wel vaker in psychologische experimenten, een addertje onder het gras. Ofschoon de beginsituatie voor beide groepen gelijk was, was de afloop verschillend. De tweede groep kreeg niet de ware eindsituatie, maar een fictieve, daaraan tegengestelde gebeurtenis. Toen aan de deelnemers werd gevraagd hoe voorspelbaar de eindsituatie was, kenden beide groepen aan de hen gepresenteerde versie eenzelfde graad van voorspelbaarheid toe. Kortom, na de feiten klinkt het in koor: 'We hebben het de hele tijd geweten!'

De onverbeterlijke Hamlet

Dat het Hamlet-personage in ons nooit ver weg is, bleek toen studenten in de volgende fictieve situatie een keuze moesten maken: 'Je hebt net een examen achter de rug, en je krijgt een uitzonderlijk voordelige aanbieding voor een vakantie. Je moet meteen beslissen en een voorschot betalen, want in geval van uitstel kost de vakantie je heel wat meer. Wat zou je doen als je wist (a) dat je geslaagd bent voor je examen, en (b) dat je gezakt bent?'

Een meerderheid van de studenten zou naar eigen zeggen in beide situaties dadelijk op het aanbod ingaan. Maar van die meerderheid beweerden de meesten ook dat ze de beslissing zouden uitstellen als ze nog niet wisten of ze nu wel of niet geslaagd waren. Intuïtief klinkt dat aannemelijk, maar in wezen is het een contradictie van formaat. Er zijn maar twee situaties mogelijk (geslaagd of gezakt) en in elk van beide nemen de studenten hetzelfde besluit. Maar als ze nog niet weten welke van beide zich zal voordoen, houden ze het voorlopig voor bekeken. Ook al beseffen ze dat die vakantie dan een stuk meer gaat kosten. Psychologen hebben het kind een naam gegeven: irrationele voorzichtigheid.

De hersenkraker van Monty Hall

Een nominatie voor de meest verwarrende hersenkraker verdient wellicht de 'paradox van Monty Hall', een vraagstukje dat in 1990-1991 de Amerikaanse media in beroering bracht en naar de presentator van een televisieshow werd genoemd die mensen een auto of een geit bood. Een eerdere versie ervan (het 'dilemma van de drie gevangenen') werd al in 1959 geformuleerd door Martin Gardner, bekend om zijn wiskundige puzzels in *Scientific American*. De makkelijkst te aanschouwen variant van deze hersenpijniger is eigenlijk een spel. Dat gaat als volgt. Er zijn drie identieke doosjes, elk voorzien van een deksel. In één ervan heb ik geld gestopt, en je mag raden in welk doosje het geld zit (als je juist raadt win je het geld, bij Gardner ging het niet om geld maar een executiebevel). Bijzonder nu aan het spel is dat het zich in twee fases afspeelt. Eerste fase: je maakt je eerste keuze. Het gekozen doosje blijft voorlopig gesloten. Ik open echter wel één van de andere twee doosjes, namelijk een doosje waarvan ik weet dat het leeg is (als je een leeg doosje hebt gekozen, kies ik het resterende lege doosje; als je een 'vol' doosje hebt gekozen, blijven er twee lege doosjes over, waarvan ik er één - om het even welk - openmaak).

Zo zijn we in de tweede fase beland: er blijven twee gesloten doosjes over en je mag opnieuw een (dit keer definitieve) keuze maken. Ofwel blijf je bij je eerste keuze, ofwel kies je het andere doosje. Hamvraag is nu welke strategie je het best kunt volgen in de tweede fase: bij je eerste keuze blijven, of het andere doosje kiezen? Of maakt het niet uit welk doosje je kiest? Met andere woorden, hoe liggen de kansen op winst in het geval dat je bij je eerste keuze blijft en welke kansen heb je in het geval dat je voor het andere doosje kiest?

De antwoorden lopen hier nogal uiteen, ofschoon de meesten geloven dat het om het even is welk van beide doosjes je kiest. Want - zo redeneren ze - je hebt maar twee keuzemogelijkheden, en dus bedraagt de kans in beide gevallen 50 procent. Velen zullen dan ook verrast zijn te horen dat je voor het andere doosje moet opteren. Want dan heb je tweederde kans te winnen. Als je bij je eerste keuze blijft daarentegen heb je maar eenderde

kans. Hoezo? De verklaring is eenvoudig: elk van de drie doosjes had aanvankelijk éénderde kans om het geld te bevatten. Het doosje dat je in de eerste fase kiest, heeft dus éénderde kans. De andere twee doosjes hebben samen tweederde kans. Deze situatie verandert niet doordat ik een doosje open maak waarvan ik weet dat het leeg is en jij weet dat ik het weet (als ik *op goed geluk*, bijvoorbeeld door een munt op te gooien, een doosje had opengemaakt dat toevallig leeg bleek, is het antwoord wel 50 procent, maar dat is een totaal ander spel, evenals de situatie dat ik na jouw keuze mag besluiten je geen tweede kans te geven). De twee doosjes (een open en een dicht) hebben nog steeds samen een kans van tweederde om het geld te bevatten, en dat is dan ook de kans dat het geld in dat dichte doosje zit.

Voor sommige mensen klinkt deze uitleg mysterieus. Een andere manier om de zaken te bekijken is als volgt. De persoon die moet raden, zal in de eerste fase in een van de drie gevallen het juiste doosje hebben aangeduid. In twee van de drie gevallen zal hij een verkeerde keuze hebben gemaakt. Ook hier is het weer wezenlijk van belang te begrijpen dat het openen van het lege doosje geen enkele informatie verstrekt die de kans verandert dat in de eerste fase het gelddoosje is aangewezen. Als ik geen doosjes zou openen en ik zou hem toestaan de inhoud van die beide andere doosjes samen te nemen, zou dat op hetzelfde neerkomen. Bijgevolg zal de strategie 'keuze veranderen' in twee van de drie gevallen succes hebben. De ongelovigen onder u raden we aan het spel herhaalde keren in werkelijkheid te spelen. Binnen tien tot twintig keer spelen tekent zich het dramatische verschil in winstkans tussen de strategieën 'vasthouden' of 'wisselen' af.

Brengt het onderwijs redding?

Neen, zeggen pessimisten, en ze refereren aan de experimenten van Tversky en Kahneman. Zelfs experts in kansrekening en statistiek scoren daarin niet veel beter dan anderen. We mogen dan al in staat zijn nauwgezet logische en statistische regels te volgen, aldus de doemdenkers, in eenvoudige, dagelijkse situaties is dat veeleer uitzonderlijk. In die situaties vallen we immers terug op vuistregels, en die volgen het principe van de hoogste opbrengst tegen de laagste kost. De opbrengst die we verwachten is een wereldbeeld dat veeleer eenvoudig, geruststellend en rooskleurig dan rationeel is. Zolang denkfouten en vooroordelen onze eigen belangen niet schaden, geven we de voorkeur aan eenvoudige denkoperaties en gezonde illusies. Onze biologische natuur, zo stellen de pessimisten, heeft geen baat bij een realistische kijk en kritische zin.

Optimisten daarentegen vinden deze uitspraken over onze natuur voorbarig. Zij verwachten wél redding van het onderwijs, maar dan niet het onderwijs zoals we dat vandaag kennen. Hun bezwaar luidt dat de klassieke cursussen logica en statistiek een te sterke nadruk leggen op wiskundige bewijsvoering en abstracte, formele structuren. Ook al beloven vele handboeken de lezer een rationeler kijk, de modellen en technieken zijn wereldvreemd. Dit is het standpunt van Critical Thinking, een groep Amerikaanse onderwijsvernieuwers die hun ideeën verspreiden in tijdschriften als *Argumentation* en *Informal Logic*. Zij ontkennen de waarde van formele logica en statistiek niet, maar beschouwen ze als aanvullend. Redeneren en argumenteren is volgens hen aan contexten gebonden. Een kind bezit geen uniform en universeel denkmodel dat het in alle situaties toepast. Boodschappen doen en sommetjes maken in de rekenklas gebeuren met behulp van aparte denkkaders (frames) waarin regels en feiten onderling sterk met elkaar verweven zijn. Zolang een scholier ongevoelig is voor gelijkenissen tussen denkkaders, wordt die informatie apart opgeslagen. En dat is net wat hij met de vakken wiskunde, logica en statistiek doet. Hij kent de regels en feiten om te betalen en om een goed cijfer voor rekenen te halen, maar ziet geen verbanden. Er treedt geen transfer op, zeggen pedagogen. Voor de didacticus is dit een nachtmerrie.

Het motto van Critical Thinking is eenvoudig. Keer terug naar contexten die leerlingen herkennen, leer ze daarin problemen formuleren en laat hen de behoefte aan oplossingen

voelen. De taak van de leerkracht beperkt zich tot het aanreiken van diverse methodes die door middel van dialoog en discussie worden beoordeeld. Zo wordt de kloof tussen abstracte regels en hun toepassingsgebieden overbrugd. Misschien lijkt dit vanzelfsprekend, maar critici wijzen op twee moeilijkheden. Vooreerst is er weinig empirisch bewijs dat leerlingen die dit programma doorliepen ook beter gaan redeneren en argumenteren. Wel zijn ze meer gemotiveerd, maar wellicht ligt dit aan de originaliteit van de oefeningen. Ook vragen sommigen zich af of het ene radicale idee niet door het andere is vervangen. Leerlingen beschikken in het beste geval over rijkere denkkaders, maar zien zij ook meer gelijkenissen? Het spook van de transfer is terug. Voor de didacticus is de nachtmerrie nog lang niet over.

[HOMEPAGE SKEPSIS](http://www.skepsis.nl/homepage_skepsis)

