



Borreltijd met Robert Trivers

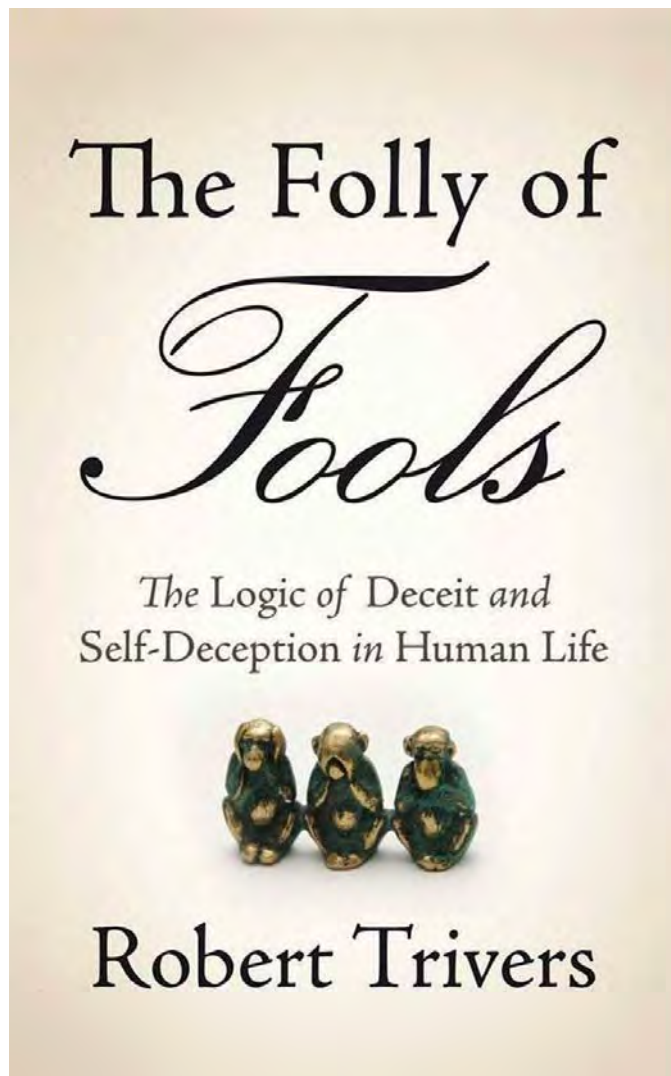
De sociobiologie van zelfbedrog

Harald Merckelbach

Zelfbedrog? Janki, hoofdpersoon in Lewinsky's roman *Het lot van de familie Meyer*, is een schoolvoorbeeld. De Joodse jongeman gaat als rekrut onder de wapenen in het Franse leger. Maar nog voordat zijn regiment slag kan leveren in de Frans-Pruisische oorlog van 1870, capituleert Frankrijk. Janki vestigt zich vervolgens in een Zwitserse provincieplaats en opent er een kledingzaak. Vanwege antisemitische vooroordelen wil de klandizie maar niet op gang komen. Een familielid schiet te hulp en schrijft in de lokale krant een artikel waarin Janki wordt geportretteerd als een oorlogsheld. Janki zou aan het front een kogel in zijn rechterbeen hebben gekregen en desondanks een andere soldaat het leven hebben gered. Janki speelt het spel mee en bedient zijn klanten hinkend en met een stok in de hand. Vanaf dan stromen de klanten toe. Ze willen de oorlogsheld met eigen

ogen zien. Een paar honderd pagina's later mist Janki zijn stok als hij hem niet bij de hand heeft, want hij voelt echt een doffe pijn in zijn rechterbeen.

Het is fictie, maar Janki's geval zou prima hebben gepast in *The Folly of Fools*. In dat boek betoogt Robert Trivers – evolutiebioloog en hoogleraar antropologie aan *Rutgers University* – dat zelfbedrog een functie heeft. Anderen beliegen gaat doorgaans gepaard met allerlei begeleidingsverschijnselen: je stem gaat omhoog, je taal wordt omslachtig en je lichaamshouding verstart. Het zijn uiterlijk waarneembare tekens die de leugenaar door de mand laten vallen. Maar de leugenaar die zichzelf om de tuin leidt, vertoont zulke verschijnselen niet langer. Dat maakt de kans op succes – in onderhandelingen, bij het vinden van een geschikte partner, tijdens conflicten – groter. Net als Janki, die zijn rol als oorlogsheld zo virtuoos speelt dat hij er zelf in gaat geloven. Zelfbedrog is dus een effectieve vorm van liegen en daarom heeft de evolutie er een premie op gezet. Aldus Trivers.



Evolutionaire wedloop

We zijn geneigd om liegen te beschouwen als iets dat thuishoort in het menselijke gedragsrepertoire. Maar list en bedrog komen wel degelijk ook op lagere treden van de evolutieladder voor. En hoe. Trivers geeft het voorbeeld van de koekoek die zijn eieren probeert te laten uitbroeden door andere vogels zoals de gele kwikstaart. De kans op uitstoting is geringer als het ei van de koekoek een kleur aanneemt die past bij die van het pleeggezin. Maar pleegouders nemen tegenmaatregelen, gaan eieren tellen en gooien elk exemplaar dat extra opduikt uit het nest. De koekoek past op zijn beurt de tactiek aan en gooit eerst een ei van de pleegmoeder weg voordat hij zijn parasitaire ei in het nest legt. Deze wedloop illustreert dat er evolutionaire druk staat op misleiding, waardoor de tactieken waarvan dieren zich bedienen steeds geslepen worden. Als je er zo naar kijkt, kun je je inderdaad voorstellen dat aan de top van de evolutionaire ladder misleiding culmineert in zelfbedrog.

Trivers beweert dat bij sociale dieren misleiding een belangrijke aanjager van intelligentie is. Kijk maar naar apensoorten, zegt Trivers: hoe bedrevener ze zijn in het misleiden, hoe groter hun neocortex en dus hoe intelligenter. Op andere plekken in zijn boek lijkt Trivers de causaliteit om te keren. Dat gebeurt bijvoorbeeld als hij te spreken komt over het klassieke experiment van kinderpsycholoog Michael Lewis, een experiment dat in allerlei varianten vaak is herhaald. De crux ervan is dat onderzoekers jonge kinderen met hun rug naar een ingepakt speelgoedje laten zitten. De proefleider pakt het speelgoed uit. De kinderen worden geïnstrueerd om vooral niet achterom te kijken. Dan verlaat de proefleider de ruimte om

even iets te halen. Sommige kinderen kijken vervolgens toch achterom en liegen daar later over ('nee, ik heb niet gekeken'). Zij scoren hoger op een IQ-test dan kinderen die ook kijken, maar dat eerlijk opbiechten. De boodschap is – althans wat Trivers betreft – dat intelligentie en liegen innig samenhangen en dat die samenhang zelfbedrog pas mogelijk maakt.

Illusies

Hoe gaat dat, intelligent liegen? Trivers haalt er de psychologie van de cognitieve illusies – in jargon *biases* – bij om dat te verhelderen. We neigen ertoe om te denken dat we intelligenter en aardiger zijn dan anderen. We onthouden de feiten die stroken met onze opvattingen en negeren tegenspraak. We reviseren onze autobiografische herinneringen zodanig dat ons leven er terugblikkend zonniger uitziet. Het zijn zulke vervormingen die ons in staat stellen om onszelf te beliegen. Maar ergens in een ander deel van ons brein kennen we volgens Trivers de waarheid. De essentie van zelfbedrog is dat de waarheid wordt onderdrukt en dat het een leugenachtige versie van de werkelijkheid is, die ons gedrag stuurt. De genetica biedt volgens Trivers de biologische vrijheidsgraden voor zo'n gespleten persoonlijkheid. Want, betoogt hij, onze moederlijke genen zetten aan tot sociaal gedrag, waarbij de hogere breinregionen een sleutelrol spelen. Maar onze vaderlijke genen zetten aan tot primitief gedrag gericht op onmiddellijke bevrediging en daarvoor zijn de lagere hersenregio's van belang. De kernboodschap van Trivers' boek wordt in deze zin mooi samengevat: 'in important parts of our family lives, we are two separable people (not one) with partly divergent aims, theories of reality, and degrees of deceit and self-deception – two people who are also tempted to deceive each other. We call these people our maternal and our paternal selves.' (p. 87).

Met zelfbedrog is op de korte termijn en in een kleinschalig sociaal verband winst te behalen. Maar, schrijft Trivers, op de lange termijn en in groter verband is zelfbedrog riskant. Catastrofes zoals die met de *Challenger* (1986) en de *Columbia* (2003) vinden hun oorzaak in zelfbedrog dat gecultiveerd werd binnen instellingen. In de genoemde voorbeelden gaat het om de luchtvaartorganisatie NASA en hoe haar topbestuurders de technische mankementen wegwuifden die de ingenieurs bij de *space shuttles* hadden geconstateerd. Het akelige aan dit type bedrog is dat de prijs ervoor door anderen wordt betaald. Hetzelfde is, aldus Trivers, aan de orde bij de Amerikaanse inval in Irak, die gerechtvaardigd werd door de leugen dat in dat land *weapons of mass destruction* zouden zijn. En institutioneel zelfbedrog is ook symptomatisch voor de Israëlische politiek in de Palestijnse gebieden, die geïnspireerd wordt door de fictie dat het Joodse volk een historisch recht kan doen laten gelden op deze gebieden. Al deze kwesties moeten hem behoorlijk dwarszitten, want Trivers besteedt er hele hoofdstukken aan. Daarbij zakt de argumentatie soms naar een bedenkelijk niveau, zoals daar waar Trivers zegt dat Joden een sterke verbale traditie hebben en dat dit hen in staat stelt om goed te liegen tegen anderen en zichzelf (p. 244).

Spektakel

Het zijn dat soort politieke excursies die de manco's van Trivers' benadering blootleggen. Als de schaal waarop je denkt de evolutionaire tijdslijn is, kan het makkelijk gebeuren dat je overmoedig raakt en denkt dat complexe historische problemen te herleiden zijn tot darwiniaanse dynamiek. Het is het soort intellectueel spektakel dat sociobiologen een slechte reputatie bezorgde. Dat die verdiend is, zie je bij Trivers geïllustreerd. Maar toch. Over Trivers gaat het verhaal dat hij als student zo geboeid was door de lectuur

van de filosoof Ludwig Wittgenstein, dat hij er niet meer van kon slapen. Met zo'n achtergrond zou je verwachten dat een auteur gevoelig is voor het inzicht dat er verschillende niveaus van analyse zijn – taalspelen voor mijn part – en dat je die niet met elkaar moet verwarren. Dat inzicht is ver te zoeken bij Trivers. Op allerlei plekken in het boek spreekt hij over NASA, over de Verenigde Staten en over Israël alsof het personen zijn, die zich inderdaad te buiten kunnen gaan aan misleiding en zelfbedrog. Categoriefouten derhalve, zouden de filosofen zeggen.

Er is nog iets wat Trivers gemeenschappelijk heeft met menig sociobioloog: een uitgesproken minachting voor de psychologie. In de karikatuur die Trivers van de psychologie maakt, is het een wetenschap die geen oog heeft voor de biologische wortels van gedrag. Het is een discipline van triviale en weinig realistische laboratoriumexperimenten. Het psychologisch laboratorium is 'a near-hopeless place in which to investigate deception and its consequences' (p. 336), schrijft Trivers. Het is een merkwaardig, maar ook onverstandig standpunt voor iemand die geïnteresseerd is in zelfbedrog. Merkwaardig omdat, als het hem zo uitkomt, Trivers wel nadrukkelijk verwijst naar bevindingen uit het psychologisch laboratorium. Het voorbeeld van Michael Lewis kwam hierboven al ter sprake. Veelzeggend zijn ook de colleges van Trivers op *YouTube* (zie bijvoorbeeld zijn voordracht tijdens *TedxJamaica*). Voor een belangrijk deel bestaan ze uit het etaleren van andermans psychologisch labonderzoek naar liegen.

Het dedain voor de psychologie is bovendien onverstandig voor een auteur die al zijn denkkraft wil spenderen aan het ophelderen van het raadsel dat zelfbedrog is. In een deel van zijn boek probeert Trivers de lezer ervan te overtuigen dat de mens uit een module van vaderlijke en een module van moederlijke genen bestaat. Zelfbedrog is daarmee het ene deel van onszelf dat het andere deel van onszelf besodemetert. Maar waarom maakt ons ge-

drag dan een redelijk gecoördineerde indruk? Hebben deze modules weet van elkaar? En hoe verwerven ze controle over het gedrag? Het zijn allemaal psychologische vragen, die de bioloog Trivers niet stelt en dat maakt zijn betoog er niet sterker op. Verderop in het boek horen we niets meer over modules in ons mentale apparaat en hoe die zelfbedrog in gang kunnen zetten. Zelfbedrog wordt dan iets waaraan politici en NASA-bestuurders zich te buiten gaan.

Trivers heeft zo'n modulaire opbouw van het mentale apparaat nodig. Het is zelfs een cruciale stap in zijn redenering. Wie die stap niet zet, eindigt al snel bij een logisch onmogelijke constellatie. Dan is immers zelfbedrog een en dezelfde persoon die weet dat A niet klopt en toch meent dat A klopt (zie voor een uitgebreider commentaar: Livingstone Smith, 2011). Aan deze paradox valt te ontkomen door losjes om te springen met het begrip zelfbedrog. Als je die weg bewandelt, kun je een rijkdom aan verschijnselen moeiteloos als uitingen van zelfbedrog duiden. Ook de bestuurders die A besluiten – zeg Irak binnenvallen of een *space shuttle* lanceren – ofschoon ze hadden kunnen weten dat A riskant is, kun je dan in de categorie van zelfbedrog plaatsen. De vraag rijst of we iets opschieten met zo'n ruimhartige aanpak. Ik betwijfel het. De bestuurders en politici die niet geïnteresseerd zijn in de waarheid en vanuit zo'n houding aan *bullshitting* doen, misleiden anderen. Dat zij ook zichzelf misleiden is lang geen uitgemaakte zaak.

Spieken

Trivers zou aanmerkelijk verder zijn gekomen als hij strenger was geweest. Streng ben je als je zelfbedrog alleen maar aan anderen toeschrijft als er hard valt maken dat zij zichzelf tegen beter weten in beliegen. Janki is daar alleen maar een literair voorbeeld van. Het existentiebewijs voor het fenomeen werd nog niet zo lang geleden in het psychologisch laboratorium geleverd. Ik doel op het werk



Daniel Day-Lewis in de film 'Left Foot'.

van Chance en collega's (2011), dat Trivers blijkbaar is ontgaan. Deze onderzoekers gaven proefpersonen een test met vragen van het type 'hoeveel Amerikaanse staten grenzen aan Mexico?' Er waren twee versies: een gewone versie en een versie waarbij de antwoordsleutel per ongeluk onder aan de bladzijde te vinden was. Veel proefpersonen die deze tweede versie invulden, spiekten. Niet verwonderlijk dus dat de spiekers beter presteerden dan degenen die niet hadden kunnen spieken. Toen de proefpersonen even daarna werd gevraagd hoe zij dachten het te gaan doen op een volgende test, werd een opmerkelijk patroon zichtbaar: de spiekers waren optimistischer over hun toekomstige prestaties dan degenen die niet hadden kunnen spieken. Maar op die tweede test deden beide groepen het even slecht (of goed). Zelfs als er een geldelijke beloning werd gezet op accurate voorspellingen van de eigen prestaties bleven de spiekers zichzelf overschatten. Kijk, dát is nu zelfbedrog. Want terwijl ze drommels goed weten dat ze hebben gespiekt, maken de spiekers zichzelf wijs dat hun fraaie prestaties een bewijs zijn van hun eruditie.

Betekent het dat de spiekers een modulair brein hebben? Een moederlijke module die weet dat spieken prestaties vertekent, en een vaderlijke module die neigt tot zelfoverschatting? Trivers zou het ongetwijfeld zo zien. Maar er is een betere verklaring en die valt te ontlenen aan wat, in psychologisch jargon, cognitieve dissonantie heet. Het komt er op neer dat mensen er sterk aan hechten dat hun opvattingen en gedragingen met elkaar in de pas lopen. Trots zijn op je fantastische test scores, maar ook weten dat je iets deed wat niet mag – spieken – is incoherent ('dissonant') en daarom onaangenaam. Mensen zijn geneigd om zo'n dissonantie op te lossen door gaandeweg hun opvattingen ingrijpend te amenderen: 'ja, ik heb gespiekt, maar ook als ik dat niet had gedaan, zou ik een hoge score hebben gehaald, want ik wist alle vragen.' Zo bezien is zelfbedrog iets dat zich over de tijd geleidelijk aan ontvouwt: het begint met een leugen die als vervelend wordt ervaren en eindigt met een geamendeerde interpretatie waaruit de leugen is verdwenen. Op tijdstip 1 weten we dat A niet klopt. Maar op tijdstip 2 hebben we onze mening herzien en denken we te weten dat A bij nader inzien toch wel klopt. Het is als met Janki die zijn carrière start als simulant en uiteindelijk tot de conclusie komt dat hij toch ook echt pijn heeft aan dat rechterbeen. Deze psychologische interpretatie van zelfbedrog heeft aanmerkelijk meer empirisch krediet dan Trivers' speculaties over het modulaire brein.

Daniel Day-Lewis

In de wereld van de psychiatrie zijn de Janki's in allerlei soorten en maten te vinden: alcoholici die weten dat hun excessief drankgebruik hun vroeg of laat fataal zal worden, anorectische patiënten die tegen beter weten in volhouden dat ze gewicht moeten kwijtra-ken en, misschien wel het meest in het oog springend, conversiepatiënten die heilig geloven dat zij aan een van hun ledematen verlamd zijn, terwijl daar geen sprake van is. In al deze aandoeningen zit een element van zelfbedrog. Je zou dan ook verwachten dat een boek over het thema de psychiatrische gevallen van zelfbedrog prominent in de etalage zet. Maar niet bij Trivers. Waarom? Is het teveel psychologie? Of is het zelfdestructieve karakter van deze aandoeningen te strijdig met Trivers' idee dat zelfbedrog op de korte termijn voordeel oplevert?

Waar Trivers zijn boek wél mee verluchtigt, is onnozele anekdotiek uit de eigen doos. Zo vertelt hij dat hij van de kamers van zijn collega's onbewust pennen, aanstekers en sleutels meeneemt. En dat hij de neiging – ook weer 'onbewust' – heeft om aantrekkelijke vrouwen met zijn linkerhand te bepotelen. Trivers wil er mee zeggen dat wij allemaal ontoegankelijke modules in ons brein hebben,

maar ondertussen zijn het weinig overtuigende ontboezemingen. Als lezer houd je er vooral de indruk aan over dat de auteur een clownske figuur is.

Natuurlijk, als evolutiebioloog heeft Trivers niet de zegen van een rijk patiëntenbestand dat meer reliëf kan geven aan een betoog over zelfbedrog. Maar hij had wel op zoek kunnen gaan naar interessante casuïstiek die zijn theorie ondersteunt of juist tegenspreekt. Zulke casuïstiek ligt voor het oprapen. Neem de acteur Daniel Day-Lewis. Hij is een exponent van de school die *method-acting* propageert. Daarbij probeert de toneelspeler helemaal op te gaan in het personage dat hij moet verbeelden. Dus toen Daniel Day-Lewis de rol van de verlamde dichter Christy Brown in *My Left Foot* (1989) ging spelen, kwam hij niet meer uit zijn rolstoel en liet hij zich met een lepel voeren door de filmploeg. Day-Lewis vertelt in interviews dat hij de grootst mogelijke moeite heeft om na de filmopnames los te komen van zijn rol. Hij blijft hangen in zijn personage. Het doet denken aan Janki. En jawel, het past ook in Trivers' idee dat zelfbedrog op de korte termijn een voordeel oplevert, want Day-Lewis wordt geroemd vanwege de natuurgetrouwe manier waarop hij personages speelt. Hij is zijn personages en dat heeft hem al menige Oscar opgeleverd.

Anti-Janki's

Minstens zo interessant zijn mensen *geen* zelfbedrog aan de dag leggen ofschoon dat wel beter voor hen zou zijn geweest. Neem de Canadees die zich ergens in 2012 meldde op het politiebureau in Den Haag. Hij kwam op verzoek van zijn onlangs overleden vader een verkeersboete betalen. De overleden vader had de bekeuring van een paar gulden in 1958 gekregen, omdat hij toen met zijn brommer een overtreding had begaan. Kort daarop was hij naar Canada geëmigreerd. Maar al die tijd was de openstaande boete aan zijn geweten blijven knagen. Op zijn sterfbed had de man zijn zoon geïnstrueerd om de boete alsnog te gaan betalen. Waarom hief de man zijn gewetenswroeging niet op met zelfbedrog? Dat zou voor zijn gemoedsrust het allerbeste zijn geweest. Als je er naar zoekt, blijkt het te wemelen van zulke anti-Janki's en ze schreeuwen om een verklaring.

Door zijn geringe appreciatie van de psychologie ziet Trivers de Janki's en anti-Janki's over het hoofd. Jammer. Een analyse ervan had Trivers meer inzicht opgeleverd in het verschijnsel van zelfbedrog dan de studie van hoe de koekoek en de gele kwikstaart elkaar over en weer bedotten. Vooral als hij de Janki's en hun tegenpolen naar het psychologisch laboratorium zou hebben gebracht. Daarom: iemand moet nog eens tegen Robert Trivers zeggen: *get thee to a laboratory* (Dunning, 2011).

Literatuur

- Robert Trivers (2011). *The folly of fools: The logic of deceit and self-deception in human life*. New York, Basic Books. 2011
- Zoe Chance, Michael L. Norton, Francesca Gino & Dan Ariely (2011). Temporal view of the costs and benefits of self-deception. *Proceedings of the National Academy of Sciences*, 108, 15655-15659.
- David Dunning (2011). Get thee to a laboratory. *Behavioral and Brain Sciences*, 34, 18-19.
- David Livingstone Smith. Aiming at self-deception: Deflationism, intentionalism, and biological purpose (2011). *Behavioral and Brain Sciences*, 34, 37-38.

Harald Merckelbach is als hoogleraar psychologie verbonden aan de Universiteit Maastricht.