

Omslag: wordt door de drukker geleverd

Sympathie over en weer

Waaruit bestaat de biologische basis van moraal?

Cor van der Weele

CIP-GEGEVENS KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Weele, Cor van der

Sympathie over en weer. Waaruit bestaat de biologische basis van moraal?

Preadvies uitgebracht t.b.v. de jaarvergadering van

de Nederlandse Vereniging voor Bioethiek

op 20 januari door C. van der Weele – Utrecht: NVBE

ISBN 978-90-367-4222-1

trefw: biologie, moraal, sociale psychologie, sympathie, empathie,
gehoorzaamheid, Darwin, De Waal, Adam Smith

Copyright 2010, C. van der Weele - Rotterdam

Alle rechten voorbehouden. Niets in deze uitgave mag worden
verveelvoudigd, opgeslagen in een geautomatiseerd gegevens-
bestand of openbaar gemaakt in enige vorm of op enige wijze,
hetzij elektronisch, mechanisch, door fotokopieën, opnamen of
enig andere manier, zonder voorafgaande toestemming van
de auteur.

SYMPATHIE OVER EN WEER

Waaruit bestaat de biologische basis van moraal?

Cor van der Weele

Nederlandse Vereniging voor Bio-ethiek



Preadvies uitgebracht t.b.v. de jaarvergadering van
de **Nederlandse Vereniging voor Bioethiek**
op 20 januari 2010 - Utrecht: NVBE

INHOUDSOPGAVE

1. Inleiding: Biologie, Psychologie en Moraal	1
2. De Waal: Van Nature Goed	5
3. Darwin: “The Workings of Sympathy”	11
4. Adam Smith en “het Spook van Mandeville”	17
5. Tijdgeesten	23
6. Gehoorzaamheid (etc)	29
7. Empathie	33
8. Situaties en Innerlijke Krachten	37
Dankwoord	43
Referenties.	45
De gehoorzaamheidsexperimenten van Milgram	49

“For to what purpose is all the toil and bustle of this world? (...) To be observed, to be attended to, to be taken notice of with sympathy, complacency and approbation, are all the advantages which we can propose to derive from it.”

Adam Smith (1759): *The Theory of Moral Sentiments, Part I, Section III, Chapter II*

1 INLEIDING

BIOLOGIE, PSYCHOLOGIE EN MORAAAL

Toen E.O. Wilson in 1975 vond dat het tijd werd om de bestudering van moraal tijdelijk uit handen te nemen van filosofen en ter beter begrip voor te leggen aan biologen, was hij te vroeg en te eenzijdig. Biologie was in die tijd onder filosofen en in de sociale wetenschappen vooral verdacht, en wat Wilson voorstelde kwam arrogant en gevaarlijk over. De laatste jaren lijkt de tijd veel rijper; het gaat nu niet om uit handen nemen, maar om toenadering en synthese over en weer. Het idee dat moraal een biologische basis heeft ontmoet al dan niet gematigde sympathie bij zowel biologen, psychologen als filosofen.

Denken, zo is inmiddels alom geaccepteerd, is een door en door lichamelijke activiteit, die sterk emotioneel en intuïtief gekleurd is en die bovendien grotendeels onbewust verloopt. Biologie is dan meteen relevant om denken te begrijpen. Theorieën van “dual processing” verdelen onze cognitieve activiteiten in een snel, intuïtief, onbewust en evolutionair oud deel en een traag, bewust, talig en evolutionair jonger deel. Er bestaat in grote lijnen consensus over het idee dat filosofen de moraal teveel hebben gesitueerd in dat laatste, rationele, talige en bewuste deel, en de oude intuïtieve kanten hebben verwaarloosd. De tegenstelling tussen cognitie en emotie heeft gaandeweg plaats gemaakt voor onderscheidingen tussen cognitieve processen die meer dan wel minder emotioneel geladen zijn.

Jonathan Haidt (2001, 2007, Haidt and Kesebir 2010) loopt voorop

in het verwoorden van een “nieuwe synthese” op het gebied van morele psychologie. Een van de centrale onderdelen daarvan is dat intuïties bij het vormen van morele oordelen leidend zijn (zonder een absoluut dictaat op te leggen). Haidt ziet Frans de Waal’s *Good Natured* (1996), waarin duidelijk werd dat de “bouwstenen” van menselijke moraal ook bij mensapen voorkomen, als een van de twee centrale boeken in wat hij noemt de “affectieve revolutie” in de morele psychologie. Het andere is *Descartes’ Error* van Antonio Damasio (1994), dat liet zien dat emoties van cruciale betekenis zijn voor morele oordeelsvorming.

Het werk van Frans deWaal heeft inderdaad grote invloed en draagt sterk bij aan de terreinwinst van het idee dat moraal diep in onze evolutionaire geschiedenis wortelt. Dat idee is inmiddels nauwelijks nog omstreden. Ook in dit essay is het een uitgangspunt; ik juich de toenadering tussen filosofie, biologie en psychologie van harte toe. De vraag is welke wegen daarbij te bewandelen.

2

Ik zal beginnen met het werk van De Waal, die sociale instincten, met als kern empathie, aanwijst als de evolutionaire basis van moraal. De Waal bouwt expliciet voort op Darwin’s idee, gepresenteerd in *The Descent of Man* (2004/1871), dat elk dier met duidelijke sociale instincten onvermijdelijk een gevoel voor moraal ontwikkelt zodra de intellectuele vermogens die van de mens benaderen. Hoewel Darwin’s ideeën op het eerste gezicht inderdaad een directe voorloper vormen van die van De Waal, is dat bij nader inzien toch maar gedeeltelijk het geval, zo zal ik vervolgens beargumenteren. Een nadere beschouwing van het werk van Adam Smith, op wie Darwin zich baseerde, maakt dat nog duidelijker. Empathie, of liever de voorloper ervan, sympathie, vervult bij Smith en Darwin een rol die sterk verschilt van de rol die De Waal eraan toekent.

Deze verschillen zijn aanleiding om in de rest van het essay uitgebreider te kijken naar de vraag welke sociale instincten, en op welke manier, in de huidige sociale psychologie worden geacht de basis voor moraal te vormen,

wat daar op aan te merken valt en wat het uitmaakt. De huidige sfeer valt kort samen te vatten als ‘gehoorzaamheid is uit, empathie is in’. Ik sluit mij aan bij Jonathan Haidt, die de huidige basis te smal vindt om moraal te begrijpen. Ik zal ervoor pleiten om duidelijk onderscheid te maken tussen sociale instincten en moraal, en om alle voor moraal relevante instincten of intuïties te betrekken bij het onderzoek naar de biologische basis van moraal, dus ook gehoorzaamheid, conformisme etc. Empathie alleen is niet zaligmakend. Bovendien heeft ook empathie morele keerzijden en schaduwkanten, die meer aandacht verdienen dan ze nu krijgen. Darwin en Adam Smith bieden daarbij belangrijke aanknopingspunten; zij hadden meer oog voor de complexiteiten van empathie (nu ja, sympathie) dan De Waal.

Allerlei andere kwesties rond de biologische basis van moraal zal ik onbesproken laten, en dat zijn er talloze. Dit essay gaat bij voorbeeld niet in op de nature-nurture kwestie, niet op groepsselectie versus verwanteselectie,¹ en niet op de omstrede verdiensten van de evolutionaire psychologie. Ik zal ook maar heel beperkt stilstaan bij conceptuele en terminologische kwesties. Zo spreek ik onbekommerd over een biologische “basis” voor moraal, hoewel die metafoor wel bedenkingen oproept.² Evenmin ga ik in op vragen en theoretische posities rond de begrippen “intuïtie” en “instinct”, bij voorbeeld in verband met de mechanismen van moreel oordelen.³ Ik laat die vragen onaan-geroerd vanuit de inschatting dat ze geen klein omweggetje zouden opleveren maar een heel eigen reis nodig hebben. In de hoop dat er tussen deze onderwerpen zinvolle scheidingen vallen te maken laat ik kortom veel liggen.

Dit essay heeft zijn oorsprong in een Studium Generale lezing in Wageningen, binnen een serie in het kader van het Darwinjaar waarin onderzoekers en docenten uit allerlei disciplines vertelden hoe Darwin hun werk beïnvloedt.

1 Groepsselectie wil zeggen dat niet alleen individuen, maar ook groepen eenheid van selectie kunnen zijn. Als een eigenschap goed is voor een groep zou dat dus in directe zin evolutionair van betekenis zijn. Darwin gaf ruimte aan groepsselectie. Sinds de krachtige tegenargumenten van George Williams (1966) houden de meeste evolutiebiologen verwanteselectie en reciprociteit altruïsme aan als selectiemechanismen. Dit geldt ook voor De Waal. De laatste jaren beleeft groepsselectie een come-back (zie bv Wilson en Wilson 2008), ook via argumenten over culturele groepsselectie (bv Boyd & Richerson 2005). Haidt en Kesebir (2010) staan vrij uitgebreid stil bij ontwikkelingen rond groepsselectie.

2 Bijvoorbeeld: hij suggereert causaal eenrichtingsverkeer in plaats van interactie.

3 Hebben we bij voorbeeld een moreel orgaan of een moreel zintuig (Hauser 2006; Joyce, 2006)?

Binnen de ethiek komen evolutionaire overwegingen vooral binnen door het werk van De Waal. De herdenking van Darwin bood een aanleiding, zelfs een noodzaak, om Darwin zelf (*The Descent of Man*) eindelijk eens in te zien. Ik heb deze volgorde, en sporen van de verrassing bij het lezen van Darwin, laten bestaan in de structuur van dit essay.

Het resultaat is het product van een generalistische manier van kijken. Ik ben geen expert in evolutie of ethologie, noch in de geschiedenis van biologie en filosofie, noch in sociale psychologie. Wel probeer ik verbanden te leggen tussen deze disciplines. Dat gaat gepaard met waarneming vanuit vogelvlucht, waarbij eigenaardigheden opvallen die onzichtbaar kunnen blijven als je het onderwerp vanuit één discipline bestudeert. Zo kan de eenvoudige constatering dat een bepaald onderwerp ergens niet aan bod komt veel zeggen; het opsporen van blinde vlekken is een belangrijk product van deze manier van kijken.

4

Ook al mag er veel als je filosoof bent, geen echte deskundige te zijn geeft vaak een ongemakkelijk gevoel. Specialisering is immers een meer prominente academische trend dan integratie, en als generalist mis je veel van wat de specialisten bezighoudt. Bovendien, en dat is pas echt ongemakkelijk, je weet nooit wat je allemaal mist. Maar hoeveel gaten en frustrerende losse einden dit essay ook onvermijdelijk bevat, ik ben ervan overtuigd dat het belangrijk en nodig is verbanden tussen verschillende disciplines te leggen. De problemen van de wereld hebben geen fragmentarische maar geïntegreerde kennis (in meervoud) nodig.⁴

⁴ Vanuit die overtuiging betreur ik ook nog steeds de opheffing van de Centrale Interfaculteiten, waarin de Nederlandse filosofie van 1960 tot 1987 was ondergebracht. Verbanden leggen tussen disciplines was precies het beoogde doel van deze interfaculteiten.

2 DE WAAL: VAN NATURE GOED

Frans de Waal werd opgeleid in een tijd waarin onder invloed van Konrad Lorenz het thema agressie de ethologische debatten beheerste. Maar De Waal zag niet alleen agressie als hij naar chimpansees keek; hij zag ook verzoeningsgedrag. Vanuit zijn onderzoek naar chimpansees, en later ook bonobo's en ander apen, bouwt hij nu al vele jaren aan een visie waarin ook het goede een plaats heeft in de natuur, en waarin menselijke moraal is geworteld in sociale instincten die we delen met andere dieren. Mijn bespreking van De Waal's ideeën is met name gebaseerd op *Primates and Philosophers* (2006), waarin hij zijn visie beknopt uiteenzet in dialoog met filosofen, *Good Natured* (1996) en *The Age of Empathy* (2009).

5

Onze *state of nature* is niet die van door eigenbelang gedreven individuen, die door pure noodzaak gedwongen met pijn en moeite samenlevingen op gang krijgen. De Waal sluit zich van harte aan bij Dewey's visie dat we tot diep in onze botten sociale dieren zijn; een pre-sociale mens zullen we vergeefs zoeken. Dit betekent tegelijk dat moraal niet (alleen maar) een gevecht tegen de natuur is maar ook bouwt op en met de natuur. We zijn moreel tot in onze (evolutionaire) kern, schrijft hij in *Good Natured*, en een theorie van menselijk gedrag die moraal niet 100 % serieus neemt valt zelf niet serieus te nemen (De Waal 1996; 2/3). Geleidelijk is De Waal die kern steeds meer gaan identificeren met empathie. De titel van zijn laatste boek, *The Age of*

Empathy, maakt dat duidelijk, en de eerste zin van dat boek (“Greed is out, empathy is in”) kondigt De Waal’s missie kort en bondig aan. We hebben een nieuw biologisch mensbeeld nodig. Tot nu toe werd biologie te eenzijdig geassocieerd met zelfzuchtige menselijke neigingen, maar biologie heeft óók de lijm voortgebracht die samenlevingen bij elkaar houdt. We delen deze lijm met andere dieren, en hij verdient meer erkenning.

De Waal sluit zich daarin aan bij Darwin. Hij citeert met instemming en in diverse boeken Darwin’s visie dat ieder dier met sociale instincten moraal zal ontwikkelen zodra het intelligent genoeg wordt. Darwin (2004/1871; 120):

Any animal whatever, endowed with well-marked social instincts, the parental and filial affections being here included, would inevitably acquire a moral sense or conscience, as soon as its intellectual powers had become as well, or nearly as well developed, as in man.

6

In *The Age of Empathy* stelt De Waal na dit citaat de vraag wat die sociale instincten dan wel zijn, en hij formuleert de vraag meteen opnieuw: “What is it that makes us care about the behavior of others, or about others, period?” (De Waal 2009; 8) De sociale instincten gaan dus in zijn ogen over zorg, en ze leiden ons direct, via belangstelling voor anderen, naar empathie. Ook in *Primates and Philosophers* stelt hij onomwonden: “Human morality is firmly anchored in social emotions, with empathy at its core.” Naast empathie zijn er nog andere bouwstenen voor moraal, door De Waal ook “morele instincten” genoemd: wederkerigheid, vergelding, conflictoplossing en een gevoel van rechtvaardigheid (De Waal 2006; 167).

Dat er aan moraal een biologische basis te ontdekken valt, en dat die basis, of liever kern, bestaat uit empathie, maakt De Waal in *Primates and Philosophers* heel aanschouwelijk. Zijn betoog richt zich expliciet tegen diegenen—T.H. Huxley, Richard Dawkins, e.a.—die moraal niet zien als behorend tot onze kern. In hun benadering, die hij “vernistheorie”(vener theory) noemt, is de mens een grote zelfzuchtige bol, omgeven door een dun laagje

morele vernis. Veelzeggend is de formulering “Scratch an altruist, and watch a hypocrite bleed”, afkomstig van M. Ghiselin. Daartegenover stelt De Waal voor moraal te zien als een Russisch poppetje, waaruit bij openmaken steeds nieuwe poppetjes tevoorschijn komen. Helemaal binnenin, als allerkleinste poppetje, vinden we “emotionele besmetting” en dit is onze morele kern, de rechtstreekse voorloper van empathie. Emotionele besmetting houdt in dat we, op een volkomen automatische manier, gevoelig zijn voor de emotionele toestand van anderen. Spiegelneuronen vormen de biologische verklaring van het mechanisme. Het zijn neuronen die zowel actief zijn wanneer een dier zelf handelt als wanneer het dier diezelfde handeling bij anderen observeert. Deze neuronen zijn te vinden bij alle sociale dieren en ze worden sinds een aantal jaar intensief onderzocht door neuro-onderzoekers. In volgende lagen van evolutie worden aan emotionele besmetting vermogens toegevoegd die deze kern omvatten en erop voortbouwen: filters die verschil opleveren tussen het zelf en de ander, het vermogen tot cognitieve evaluatie, en tenslotte het volledige vermogen om het perspectief van iemand anders aan te nemen.

7

Het ontstaan van moraal wordt in een evolutionaire benadering functioneel verklaard: moraal is voor sociale dieren nodig om het leven in goede banen te leiden. De wortels van moraal liggen bij empathie voor familieleden. De grootste stap in de evolutie van moraal was de verschuiving van focus van louter interpersoonlijke relaties naar een meer omvattende vorm van het goede, ofwel een sprong naar “community concern”, aldus De Waal (2006; 54).⁵ Een van de krachtigste stimulansen van zulk groepsgoed is vijandigheid tegenover buitenstaanders, wat tot de gevolgtrekking leidt dat het meest “nobele” dat we hebben bereikt (moraal) evolutionair verbonden is met ons “laagste” gedrag: oorlog. Oorlog verschafte het gemeenschapsgevoel dat het leven in grote groepen nodig heeft.

De samenwerkingsfunctie van moraal verklaart tegelijk de beperking van onze morele aard: we zijn niet onpartijdig. Een onderscheid tussen groepsleden en buitenstaanders dragen we in onze botten mee. De morele

⁵ De Waal is niet overtuigd dat groepsselectie (zie noot 1) nodig is om het ontstaan van moraal te verklaren; we komen “een heel eind” met verwantselectie en wederkerig altruïsme. “Community Concern” weerspiegelt het belang dat ieder lid van een groep heeft bij een coöperatieve sfeer; zonder goed functionerende groep is ieder individu veel slechter af. Sociale druk speelt daarbij een rol. Volgens De Waal (2006) vormen dus op zich zelfzuchtige overwegingen de onderbouwing van een onpartijdig moreel perspectief.

kring kan wel groter worden, maar alleen als en voor zover ons leven in de kleinere kring voorspoedig verloopt. Dat we een bias hebben ten gunste van de nauwere kringen (onze familie, onze gemeenschap, onze soort), zegt De Waal, is niet zomaar een empirisch gegeven, het is een morele plicht (2006;165).

Zoals De Waal het hier stelt mist menselijke moraal onvermijdelijk een eigenschap die filosofen er juist cruciaal voor achten: onpartijdigheid. Het is een belangrijk spanningsveld in de communicatie met filosofen. In hun ogen laat deze bias in onze morele instincten namelijk zien dat er een belangrijke discontinuïteit bestaat tussen de empathische voorlopers van moraal en de normatieve kracht van morele regels. Het verschil tussen is en ought, de naturalistic fallacy, en het vermogen te worden gemotiveerd door normatieve overwegingen zijn thema's die regelmatig opduiken in de discussie tussen De Waal en filosofen (De Waal 2006). Ook in dit essay zal de partijdigheid van empathie nog terugkomen.

8

Zowel in *Primates and Philosophers* als in *The Age of Empathy* citeert De Waal niet alleen Darwin over de sociale instincten als basis voor moraal, maar ook de eerste zin van Adam Smith's *Theory of Moral Sentiments* (1759). Smith:

How selfish soever man may be supposed, there are evidently some principles in his nature, which interest him in the fortune of others, and render their happiness necessary to him, though he derives nothing from it except the pleasure of seeing it.

Het zegt veel over de onderscheidingen die we moeten maken, zegt de Waal, dat Smith, die toch vooral bekend is om zijn theorie van eigenbelang als de basis van de economie, ook over het universele vermogen voor sympathie schrijft (2006;15).

Ik zal verderop betogen dat De Waal's lezing van Smith eenzijdig is, maar hier worden wel de onderscheidingen zichtbaar die De Waal zelf maakt. In de aard van apen en mensen zijn twee tendensen werkzaam: aan de ene kant eigenbelang, aan de andere kant empathie. De Waal is niet heel pre-

cies in zijn weergave van de polen van deze dichotomie. Tegenover empathie staat soms hebzucht, soms zelfzuchtigheid, soms slechtheid, soms agressie, soms competitie. Maar hij suggereert hoe dan ook een grote, overkoepelende tweedeling in de aard van de mens: het goede (samenwerking, altruïsme, met als basis empathie) tegenover het slechte (agressie, competitie, met als basis zelfzuchtigheid).⁶ Deze tweedeling levert soms karikaturale karakteriseringen van mensapen op. De bonobo's zijn dan goed, de chimpansees slecht, en mensen vormen daarvan een complex mengsel; mensen zijn "bipolar apes", zegt hij bij voorbeeld in *The Age of Empathy*, met enerzijds trekken van de "gentle, sexy" bonobo en anderzijds van de "brutal, domineering" chimpansee (2009; 203).

In zijn nadruk op sociale instincten als basis voor moraal zegt De Waal dus op Darwin en Smith voor te bouwen. Weliswaar stellen de laatsten sympathie centraal als basis voor moraal, terwijl De Waal inzet op empathie, maar aanzien het woord empathie pas in het begin van de 20e eeuw in het Engels werd geïntroduceerd, lijkt hier eerder sprake van nuance-verschuivingen in de taal dan van een groot verschil. Ook bij De Waal zelf is sprake van zo'n verschuiving; in *Good Natured* koos hij nog voor sympathie omdat dat meer actiegericht lijkt dan empathie (1996; 41). Empathie ziet hij als een meer basaal vermogen, dat evolutionair verder teruggaat.

Over het ontstaan van het begrip empathie en de precieze afbakening van empathie en sympathie ten opzichte van elkaar is veel geschreven (voor een geschiedenis van empathie zie Eisenberg & Strayer, 1987). Er zijn vele definities in omloop en beide begrippen worden op veel verschillende manieren ten opzichte van elkaar afgebakend, waarbij empathie geleidelijk aan prominentie lijkt te winnen. Voor de doelen van dit essay zijn die nuances niet heel belangrijk. Wel van belang is dat het hier gaat om de Engelse betekenis van het woord sympathie, die breder is dan de Nederlandse. Naast

6 In een recent interview in NRC Handelsblad (Koelewijn en Spiering 2009) zegt De Waal, in reactie op de opmerking dat *The Age of Empathy* ook veel negatieve opmerkingen over de menselijke natuur bevat: "Als je een heel boek schrijft over zoiets moois als empathie, kan je maar beter de kritiek voor zijn dat er ook nog wel andere minder mooie dingen zijn in het leven. En dat is natuurlijk ook zo!" Even verderop stelt hij dat empathie op zichzelf neutraal is: het is óók negatief te gebruiken, bijvoorbeeld ten behoeve van martelen. Deze laatste lijn van denken lijkt op gespannen voet te staan met de visie dat empathie inherent goed is. Maar in *The Age of Empathy* associeert De Waal een immoreel gebruik van het vermogen het perspectief van een ander in te nemen vooral met psychopathologie, met name het ontbreken van de diepere lagen van empathie. Zie ook noot 14.

genegenheid en waardering (in het Nederlands de voornaamste associaties) verwijst het Engelse woord bv ook naar medelijden en deelneming.

Laat ik ingaan op rol van sympathie bij Darwin. Het woord is dus het punt niet.

3 DARWIN: “THE WORKINGS OF SYMPATHY”

The Descent of Man heet voluit *The Descent of Man and Selection in Relation to Sex*. In twee van de 21 hoofdstukken is moraal aan de orde: in hoofdstuk 4 bespreekt Darwin de evolutionaire oorsprong van “the moral sense”, in hoofdstuk 5 volgt de vraag hoe die moral sense zich in de loop van de geschiedenis bij mensen verder heeft ontwikkeld. Meteen aan het begin van hoofdstuk 4 komt Darwin met de basisgedachte dat moraal niet kan uitblijven als sociale dieren intelligent genoeg worden. Hij ziet daarbij vier ingrediënten. Ten eerste houden dieren met sociale instincten van elkaars gezelschap, en voelen ze in enige mate sympathie met elkaar⁷. Ten tweede spelen de geestelijke vermogens een rol: zodra die genoeg zijn ontwikkeld, zullen individuen hun acties, motieven en verlangens voortdurend in zich laten omgaan. In de derde plaats gaat taal krachtig meespelen zodra hij beschikbaar is: de wensen van de gemeenschap kunnen dan tot uitdrukking worden gebracht en een leidraad worden voor gedrag. En tenslotte is daar gewoontevorming.

11

Darwin laat eerst met allerlei voorbeelden zien dat sociaal gedrag en de kern ervan, sympathie, ook bij dieren voorkomen. Een roerende kleine anecdote uit zijn eigen ervaring gaat bij voorbeeld over een hond die warm bevriend was met een kat. Toen die kat ziek in haar mandje lag gaf de hond haar elke keer als hij langs kwam een

⁷ Sympathie “voor elkaar” zou de Nederlandse betekenis van sympathie suggereren.

paar likjes; een duidelijk geval van vriendelijk gevoel, zegt Darwin.

Niet dat het gemakkelijk is om de verschillende instincten en hun onderlinge verhouding goed te bepalen, of te besluiten hoe ze ontstaan zijn. Moed bijvoorbeeld zal waarschijnlijk onafhankelijk van sympathie zijn geëvo-lueerd, zegt Darwin, en ook eerder, maar de gewoonte van mannetjesdieren om de gemeenschap te beschermen is misschien wel ontstaan uit wederzijdse sympathie. In ieder geval is duidelijk dat instincten met elkaar in botsing kunnen komen. Het voorbeeld dat hij geeft betreft trekvogels. Het instinct om te trekken wint het in de herfst van het toch ook heel sterke ouderinstinct; als zwaluwen nog afhankelijke jongen hebben als het tijd wordt om naar het zuiden te trekken laten ze de jongen achter in het nest, zodat die onherroepelijk een miserabele dood tegemoet gaan (p.131).

12

Iedereen zal erkennen, schrijft Darwin, dat de mens een sociaal wezen is. We kunnen dat zien aan onze afkeer van eenzaamheid en onze behoefte aan gezelschap. Eenzame opsluiting is voor mensen een van de ergste straffen. Veel instincten hebben bij mensen aan kracht verloren, en zo ook de sociale, maar we hebben toch nog wel een zekere mate van bijvoorbeeld trouw, gehoorzaamheid, liefde en sympathie behouden. Zij worden aangevuld door rede en ervaring.

Voor de basale emotie van sympathie verwijst Darwin naar Adam Smith en naar Alexander Bain. De basis van sympathie, zegt hij met verwijzing naar hen, ligt in ons sterke vermogen om onze ervaringen van plezier en pijn vast te houden en ons deze later weer voor te stellen. Als we bij anderen honger, koude, pijn of vermoedheid zien worden die herinneringen geactiveerd, en dat leidt tot de impuls het lijden van anderen te verlichten, en ook met hen mee te genieten.

Sympathie is op zichzelf nog geen moraal, want een wezen dat puur instinctief handelt kan geen moreel wezen worden genoemd. Er is ook een normatief oordeel nodig: een moreel wezen is in staat om acties en motieven uit verleden en toekomst met elkaar te vergelijken, en om die goed- of af te

keuren (p. 135). Voor een uitwerking van de vraag hoe sympathie tot moraal leidt, begint Darwin bij de constatering dat mensen, anders dan bij voorbeeld zwaluwen, spijt kunnen voelen. Wat een kwelling zou het voor zwaluwen zijn als ze steeds moesten denken aan de kleintjes die daar in het vreselijke noorden verkommerden van hongersnood en kou. Zulk terugdenken is precies wat mensen niet kunnen voorkomen. Wat ons vooral bezig houdt als we nadenken over wat we hebben gedaan is wat anderen van ons denken, zegt Darwin, en dit is het gevolg van sympathie, dat centrale aspect van onze sociale instincten (p. 136).

Dit lijkt op het eerste gezicht wonderlijk. Sympathie gaat toch over medegevoel met anderen, niet over de vraag wat anderen van ons denken? Toch is het de waarde die we hechten aan het oordeel van anderen die Darwin keer op keer benadrukt als het over sympathie gaat. “As Mr Bain has clearly shewn,” schrijft hij, “the love of praise and the strong feeling of glory, and the still stronger horror of scorn and infamy, ‘are due to the workings of sympathy.’” (Darwin 2004/1871; 133)⁸ Zijn sympathie heeft hierdoor een heel ander karakter dan De Waal’s empathie. Bij Darwin ligt de nadruk niet alleen op de sympathie die we geven, maar ook en vooral op de behoefte die we hebben aan de sympathie van anderen

13

Het is de moeite waard hier preciezer naar te kijken. Als we handelen volgen we vaak sterke impulsen en dat zullen vaak impulsen zijn die op onze eigen verlangens zijn gericht, schrijft Darwin. Maar als die behoeften bevredigd zijn ebt hun belang weg, terwijl andere overwegingen sterker worden; we gaan onze daden dan bekijken in het licht van de altijd-aanwezige sociale instincten en van ons grote ontzag voor het gunstige oordeel van anderen. Dan kunnen we spijt gaan voelen, en wroeging en schaamte, waarbij vooral dat laatste berust op het oordeel van anderen, en dan gaan we ons voornemen in de toekomst anders te handelen. Hier zien we het geweten in actie: het geweten kijkt terug en dient als gids voor de toekomst. (p. 138) Zolang anderen er geen last van hebben kun je met een gerust geweten je eigen verlangens bevredigen, maar om echt vrij te zijn van zelfverwijt, of op zijn minst van angst, is het nodig om

8 De sectie in Bain’s boek waarnaar Darwin verwijst, “Love of Approbation”, maakt geen deel uit van het hoofdstuk over sympathie, maar van het hoofdstuk *Emotions of Self*. “The feeling of being approved, admired, praised by others, is a heightened form of self-gratulation, due to the workings of sympathy”, schrijft Bain daar. De sympathie van anderen, zo legt hij uit, versterkt ons in onze emoties en opinies en in onze zelfwaardering. (Bain 1868; 254).

de afkeuring van anderen te vermijden (p. 140).

Keer op keer blijft Darwin herhalen dat sympathie ons ertoe aanzet veel waarde te hechten aan de goed- en afkeuring van anderen, en dat het vervolgens onze cognitieve vermogens zijn die het ons mogelijk maken daar conclusies voor de toekomst aan te verbinden, die in sociaal verband leiden tot morele regels en gewoonten. Hij doet dat opnieuw in hoofdstuk 5, als hij probeert een historische reconstructie te geven van het ontstaan van moraal. Als primitieve stammen met elkaar in botsing komen zijn sociale instincten van levensbelang voor de groep: eigenschappen als trouw, moed en gehoorzaamheid bevorderen de sociale cohesie en het succes van de groep.⁹ Maar hoe kunnen deze sociale en morele kwaliteiten ontstaan zijn binnen een groep, vraagt Darwin zich af. Moedige mensen die zich opofferen voor de groep zullen immers minder nakomelingen hebben; hoe kan zo'n eigenschap zich dan toch verbreiden? Darwin geeft een paar verklarende mogelijkheden. Ten eerste zullen behulpzame mensen leren dat ze op tegenprestaties kunnen rekenen, en zo kan hulp zich ontwikkelen tot een sociale gewoonte. Maar een ander en veel krachtiger stimulans is de waardering en veroordeling door onze medemensen. Het sympathie-instinct leidt ertoe dat we anderen prijzen en kritiseren, en dat we het fantastisch vinden om te worden geprezen en bang zijn voor kritiek. Anderen te behandelen zoals je zelf behandeld zou willen worden is de basis van moraal. "It is, therefore, hardly possible to exaggerate the importance during rude times of the love of praise and the dread of blame." (p.157)¹⁰

In feite schetst Darwin hier zowel met de eerste als met de tweede verklaring van het ontstaan van moraal een mechanisme van reciprociteit. Beide mechanismen vallen uit te drukken met gulden-regel spreekwoorden. "Wie goed doet, goed ontmoet" schetst het beloningsmechanisme dat ons tegenprestaties in het vooruitzicht stelt als we goed doen. "Wat gij niet wilt dat u geschiedt, doe dat ook een ander niet" verwijst naar het angstmechanisme van de tweede verklaring.

9 Een mechanisme van groepsselectie (zie noot 1).

10 Ook mensen die eigenlijk niet de instinctieve neiging hebben zich op te offeren voor anderen, zullen dat toch doen door de waardering (glory) die ze daarmee winnen. Gewoonte, ervaring en denkracht zullen gezamenlijk uiteindelijk tot een hoogst complex moreel instinct leiden: "originating in the social instincts, largely guided by the approbation of our fellow-men, ruled by reason, self-interest, and in later times by deep religious feelings, and confirmed by instruction and habit." (p. 157)

Zoals gezegd verwijst Darwin voor een nader begrip van sympathie naar Alexander Bain en Adam Smith. Bain was een tijdgenoot van Darwin, Smith schreef meer dan honderd jaar eerder. Het spoor leidt dus terug naar hem.

4 ADAM SMITH EN “HET SPOOK VAN MANDEVILLE”

Adam Smith's *The Theory of Moral Sentiments* verscheen in 1759, precies een eeuw eerder dan *The Origin of Species* en 112 jaar voor *The Descent of Man*. Niemand heeft sympathie zo centraal gesteld bij het ontstaan van moraal als Smith. Het boek begint met een hoofdstuk over sympathie en valt met de deur in huis met de vaak en ook door De Waal geciteerde zin waarin hij stelt dat we ons mensen niet alleen zelfzuchtig moeten voorstellen; er zijn ook kanten aan onze natuur die zich voor anderen interesseren (zie boven).¹¹

17

De Waal houdt het bij die eerste zin, waarin onze sympathie met anderen ten tonele verschijnt. Maar deze verwijzing geeft een incompleet beeld, want na hoofdstuk 1 volgt hoofdstuk 2, en hier luidt de eerste zin:

But whatever may be the cause of sympathy, or however it may be excited, nothing pleases us more than to observe in other men a fellow-feeling with all the emotions of our own breast; nor are we ever so much shocked as by the appearance of the contrary.

Wie bij voorbeeld een gezelschap probeert te vermaken, en merkt dat er niemand lacht behalve hijzelf, is “mortified”, schrijft Smith, terwijl het voelt als een geweldig applaus als mensen meelachen.

Veel van de rest van het boek besteedt Smith aan het zoeken naar een evenwicht in het sociale spel van het geven en nemen van sympathie. In de reconstructie daarvan zal ik gebruik maken van het verhelderende proef-

¹¹ Smith omschrijft sympathie als “fellow feeling”. Oorspronkelijk was het begrip beperkt tot medelijden en compassie, schrijft hij, maar het kan heel goed worden gebruikt om medegevoel aan te geven met welke emotie dan ook. Ik ga daar niet nader op in, noch volg ik het sympathie-spoor verder terug naar de voorlopers van Smith: Hume en Hutcheson.

schrift van A.J. Kerkhof uit 1992.

Een probleem van Smith's theorie is vanaf het begin geweest dat hij zowel verklarend als normatief wil zijn en dat dat bijzonder lastig met elkaar te rijmen valt, schrijft Kerkhof. Volgens hem slaagt Smith in zijn verklarende ambities veel beter dan in de normatieve, en hij (Kerkhof) stelt zich tot taak die verklarende kracht te verhelderen. De kern van zijn analyse is te vinden in de titel die hij zijn boek heeft meegegeven: *De Mens is een Angstig Dier*. Deze uitspraak van Smith verwijst ernaar dat deugd berust op het inacht nemen van de mening van anderen en dat dit inacht nemen iets angstigs is omdat het voortkomt uit een voortdurende behoefte aan goedkeuring en erkenning.

Om deze kern te verhelderen laat Kerkhof zien dat Smith veel dichter bij Mandeville staat dan hij zelf wil toegeven. De in Rotterdam geboren en naar Engeland geëmigreerde arts/filosoof Bernard Mandeville (1670-1733) is de auteur van het satirische *The Fable of the Bees: or Private Vices, Publick Benefits* uit 1705 en later daaraan toegevoegde dialogen en essays. De boodschap van deze geschriften is dat ondeugd de bron is van maatschappelijk welzijn. Trots en ijdelheid in het bijzonder liggen ten grondslag aan morele deugd; de trots, oftewel de vrees voor schande, is zo sterk dat ze zowel eigenbelang als de vrees voor de dood kan overwinnen. Dat laatste verklaart volgens Mandeville waarom soldaten in oorlogen niet weggrennen. Mensen zijn volgens dit mensbeeld niet van nature agressief, of op een rationele en berekenende manier uit op eigenbelang; ze hebben er in de eerste plaats behoefte aan boven anderen te staan, zich superieur te voelen, gezien en erkend te worden. Hoewel dit volgens Mandeville betekent dat mensen geen sociale wezens zijn, stelt Kerkhof terecht dat het hier misschien niet om positief gerichte sociale neigingen gaat, maar dat ze wel degelijk sociaal gericht zijn in de zin dat ze mensen afhankelijk maken van andere mensen.

Smith probeert het "spook van Mandeville" buiten de deur te houden maar slaagt daar volgens Kerkhof niet werkelijk in. Smith mag dan be-

zwaar hebben tegen de term ijdelheid, maar de “kern van waarheid” die hij bij Mandeville ontwaart en die consistent is met zijn eigen theorie “steekt hierin dat de moraal uiteindelijk berust op een inachtneming van de mening van anderen” (Kerkhof 1992; 156). Smith’s bezwaren tegen ijdelheid komen er in feite op neer dat ijdelheid dom is en averechts werkt: het levert eerder verachting op dan achting. Zijn eigen alternatieven zijn gewoon betere strategieën om je te verzekeren van de goedkeurende blik van anderen (p. 158). Deze kern vormt zowel de verklarende kracht van Smith’s *Theory* als haar zwakte als moreel systeem, aldus Kerkhof.

Hoe valt dit te rijmen met een moreel mechanisme waarin sympathie centraal staat? Sympathie, in Smith’s visie niet zomaar een passieve gevoelsinfectie maar een actief vermogen om ons dingen voor te stellen die anderen zijn overkomen, zet een moreel mechanisme in werking dat naar evenwicht leidt tussen geven en nemen. Kerkhof karakteriseert dit mechanisme met behulp van de “wetten van de sympathie”¹² die inhouden a. dat wij ervan genieten als mensen met ons sympathiseren en het onaangenaam vinden als mensen dat niet doen en b. dat we het plezierig vinden met anderen te sympathiseren en het onaangenaam vinden als we dat niet kunnen doen.

19

Het gaat hier dus om tweerichtingverkeer, met de eigen behoefte voorop. In de praktijk van het sociale leven zijn mensen voortdurend bezig hun gevoelens op elkaar af te stemmen.¹³ Daarbij vallen verschillende deugden te onderscheiden. “Beminnelijke deugden” bestaan uit een meer dan gemiddeld geven van sympathie. “Achtenswaardige deugden” bestaan uit een minder dan gemiddeld nemen ervan, en berusten op zelfbeheersing. De zoektocht kan dus tot verschillende evenwichtspunten leiden, geassocieerd met verschillende culturele tradities. Een stoïcijnse moraal legt de nadruk op zelfbeheersing en dus de achtenswaardige deugden, terwijl een Christelijke moraal de beminnelijke deugden vooropstelt.

¹² In navolging van T.D. Campbell’s *Adam Smith’s Science of Morals* uit 1971.

¹³ **Ambitie ontstaat bijvoorbeeld doordat we gemakkelijker sympathie verwerven met onze successen en vreugden dan met onze tegenslagen en bedroefdheden:** “It is because mankind are disposed to sympathize more entirely with our joy than with our sorrow, that we make parade of our riches, and conceal our poverty. (...) For to what purpose is all the toil and bustle of this world? What is the end of avarice and ambition, of the pursuit of wealth, of power and preheminance? (...) To be observed, to be attended to, to be taken notice of with sympathy, complacency and approbation, are all the advantages which we can propose to derive from it.” (Smith, 1976/1759; 50).

Het punt waarop vraag en aanbod elkaar in evenwicht houden is het “punt van gepastheid”. De bepaling van dit punt is een morele prestatie waarvoor Smith de figuur van de “onpartijdige toeschouwer” te hulp roept, een hypothetisch persoon die met een blik van buitenaf vaststelt wat een gepaste vorm en mate is om bij voorbeeld lastige emoties als hevige woede of verdriet al dan niet te uiten. Ongeremd woedend zijn is in de meeste situaties niet gepast, maar als ons onrecht aangedaan is wordt enige vorm van woede toch verwacht, omdat de zaak daartoe dwingt, of omdat het anders lijkt of we over ons heen laten lopen. Nagaan wat de onpartijdige toeschouwer zou vinden helpt ons om het punt van gepastheid te vinden.

20

Smith volgt Mandeville in de visie dat het grootste deel van menselijk geluk bestaat uit de erkenning en goedkeuring van ons gedrag door anderen, dat wil zeggen dat anderen de spiegel zijn waaraan we onze eigenwaarde afmeten. De onpartijdige toeschouwer heeft voor Smith de functie ons daarvan los te maken; deze figuur dient niet om vast te stellen hoe anderen ons zien maar hoe ze ons zouden moeten zien, en dit levert ons een eigen moreel oordeel op, waarbij het niet gaat om de “love of praise” maar om de “love of praise-worthiness”. Ons eigen oordeel over prijzenswaardigheid neemt daarbij de plaats over van daadwerkelijk prijzen en geprezen worden. De deugd lijkt zo gered, stelt Kerkhof, en de afstand tot Mandeville gewaarborgd: het is niet alles ijdelheid (p. 103). Maar is dit wel zo? Aangezien de onpartijdige toeschouwer niet echt bestaat geeft deze figuur onze eigen inschatting weer van het evenwichtspunt op de markt van gepastheid. Hij is daarmee een afgeleide van de spiegel die anderen ons voorhouden. Smith moet toegeven, schrijft Kerkhof, dat zowel liefde voor roem als liefde voor de deugd een element van ijdelheid bevatten, en ook dat een gevoel verdienstelijk te zijn afhankelijk is van de oordelen van anderen. In de woorden van Smith: “De deugd wordt niet beminnelijk en verdienstelijk genoemd omdat hij het object van zijn eigen liefde of zijn eigen dankbaarheid is, maar omdat hij die gevoelens bij anderen oproept.” (Kerkhof 1992; 104).

Smith heeft dus eerder een sociologische verklaring van moraal gegeven dan een normatieve theorie; behoefte aan sympathie is de motor van het morele aanpassingsmechanisme. Aan het eind van zijn boek stelt Kerkhof, met verwijzing naar Rousseau, dat de samenleving wel vaart bij het mechanisme dat ons de goedkeurende blik van anderen doet zoeken, maar dat het individu met de brokken blijft zitten: een voortdurende angst tekort te schieten in de ogen van anderen en daaruit voortkomend een permanente onrustige gedrevenheid. “Man is an anxious animal”.

In *Wealth of Nations* en *The Theory of Moral Sentiments* lijkt Adam Smith op het eerste gezicht twee heel verschillende fundamentele menselijke neigingen centraal te stellen. In het ene boek de neiging handel te drijven en te ruilen (“the propensity to truck, barter, and exchange one thing for another”). In het ander de neiging mee te voelen met anderen en anderen gelukkig te willen zien. Smith heeft deze neigingen niet expliciet geïntegreerd in een totaalbeeld, en ze worden vaak als fundamenteel verschillend of zelfs inconsistent gezien. Maar er valt wel degelijk eenheid te ontdekken, en wederkerigheid lijkt dan de kern. Volgens de econoom Vernon Smith (Smith 1998) ging Adam Smith in feite uit van één overkoepelend gedragsaxioma, namelijk “the propensity to truck, barter and exchange one thing for another.” Dit axioma, breed geïnterpreteerd, omvat de sociale uitwisseling van goederen zowel als gunsten; het genot dat we ontlenen aan het geluk van anderen is gebaseerd op de verwachting van wederkerige gunsten, zo stelt Vernon Smith. “Kindness is the father of kindness” schrijft Adam Smith inderdaad, “and if to be beloved by our brethren be the great object of our ambition, the surest way of obtaining it is by our conduct to show that we really love them” (Smith 1976/1959; 225).

Ook Kerkhof, zo bleek hierboven, benadrukt de wederkerigheid in Smith's analyse van sympathie, en ook hij is ervan overtuigd dat aan de twee boeken eenzelfde mensbeeld ten grondslag ligt. Hij ziet daarbij zelfs nog een diepere laag dan wederkerigheid, namelijk de fundamentele behoefte van de

mens te overtuigen, waarop voor Adam Smith niet alleen de neiging tot handel is gebaseerd maar ook het zo ongemakkelijk Mandeville-achtige mechanisme van de moraal (Kerkhof 1992; 63-72).

5 TIJDGEESTEN

Terug naar Darwin en De Waal. In evolutionaire benaderingen staan “sociale instincten” aan de basis van de ontwikkeling van moraal. De sociale instincten en de daaruit voortgekomen moraal zijn functioneel: ze helpen mensen overleven door samenlevingen te laten functioneren. Zowel Darwin als De Waal spreken daarbij in meervoud: Darwin acht sympathie, loyaliteit, gehoorzaamheid en moed onmisbaar, De Waal noemt empathie, wederkerigheid, vergelding, conflictoplossing en een gevoel voor rechtvaardigheid. Beiden leggen de nadruk op de centrale betekenis van één van deze bouwstenen: sympathie respectievelijk empathie, globaal te omschrijven als een gevoeligheid voor en betrokkenheid bij de emotionele toestand van anderen. Opvallend is ook dat de rijtjes sociale instincten van Darwin en De Waal elkaar verder niet overlappen.

23

Dit centrale instinct leidt bij Darwin en De Waal via heel verschillende mechanismen tot moraal. Darwin benadrukt, in het voetspoor van Adam Smith en in overeenstemming met tijdgenoten als Blain, onze behoefte aan sympathie (erkenning, lof, bijval) en de angst voor de afwezigheid ervan in de vorm van krietiek en uitstoting. Het is vooral deze angst die, met behulp van ons vermogen tot terugzien, vooruitzien, het vormen van gewoonten en het volgen van regels, tot moraal leidt. Sociale instincten, inclusief sympathie, zijn bouwstenen voor

moraal, maar ze zijn op zichzelf niet moreel goed of slecht.

Vanuit een Smithiaans / Darwiniaans perspectief is De Waal's begrip van het functioneren van empathie eenzijdig: het bestrijkt alleen de roze kant ervan. Empathie is voor De Waal tot in zijn diepste aanzetten iets goeds, oftewel wij zijn moreel tot in onze diepste kern. Hoewel hij zich er van bewust is dat sociale vermogens op zich niet impliceren dat dieren "aardig" voor elkaar zijn (verginging hoort bij voorbeeld ook bij de sociale bouwstenen) beschouwt hij empathie wel als inherent goed.¹⁴ In overeenstemming daarmee ziet hij, anders dan bij voorbeeld Kerkhof en Vernon Smith, bij Adam Smith wel twee heel verschillende kanten: in *Wealth of Nations* komt onze zelfzuchtige helft aan bod, in *The Theory of Moral Sentiments* de helpende (De Waal 2009; 222).

De Waal maakt geen duidelijk onderscheid tussen sociale instincten en morele sentimenten; hij gebruikt deze begrippen als inwisselbaar; in *Primates and Philosophers* hanteert hij nu eens de ene, dan weer de andere aanduiding voor de bouwstenen van de moraal (bv p.167). Om recht te doen aan de minder goede kanten in onze natuur is er daarom naast deze inherent goede ook een inherent slechte neiging nodig: wij zijn zowel fundamenteel nobel als fundamenteel brutaal. De Waal karakteriseert deze brute neiging wisselend, met name als zelfzuchtigheid, hebzucht en agressie.

Hoe verschillend sympathie/empathie hier worden begrepen blijkt ook uit de plaats van reciprociteit. Voor Darwin is reciprociteit gegeven met sympathie, in de vorm van het tweerichtingsverkeer van geven en nemen; bij De Waal, voor wie empathie een kwestie is van louter geven, krijgt reciprociteit een aparte plaats onder de sociale instincten.

Door empathie te zien als louter moreel goed is ook het denken van De Waal in overeenstemming met de geest van zijn tijd, bijvoorbeeld de geest in de sociale psychologie. Deze discipline heeft uiteraard een voorname rol bij het begrijpen van sociaal gedrag, en ze speelt een belangrijke rol bij de huidige integratie van moraal, biologie en psychologie. De verbanden zoals Darwin

¹⁴ Hij zegt weliswaar dat het vermogen het perspectief van de ander in te nemen ook kan worden ingezet voor martelen (De Waal 2009; 211). Maar vervolgens ziet hij hier toch vooral pathologie, die onder meer kan bestaan uit "a permanent disconnect between perspective-taking and the deeper regions of empathy." Bij psychopaten lijkt de oude zoogdierkern van het Russische poppetje te ontbreken; ze bezitten alleen de meer cognitieve buitenste lagen, schrijft hij. Meer in het algemeen: "A lot of trouble in the world can be traced to people whose Russian doll is an empty shell."

en tijdgenoten die legden tussen evolutie en moraal zijn in de huidige sociale psychologie niet meer in al hun complexiteit te vinden. De illustratie daarvan kan kort zijn; de inhoudsopgaven van sociaal-psychologische standaardwerken volstaan in feite om de patronen te zien.

Wie een van deze boeken openslaat (ik heb er drie gekozen) op zoek naar evolutie en moraal, ziet onmiddellijk dat ze samen voorkomen in met name in één specifiek hoofdstuk, namelijk het hoofdstuk over “Prosociaal Gedrag” (Hewstone et al, 2008, Aronson et al, 2007), ook wel genoemd “Helpen en Samenwerken” (Smith and Mackie, 2007). De vraag waarom mensen andere mensen helpen leidt allereerst tot een evolutionair antwoord: verwantselectie en wederzijds altruïsme zijn de mechanismen die tot altruïstische hulpvaardigheid hebben geleid. Het psychologisch centrale mechanisme dat individuen aanzet tot zulke hulpvaardigheid is empathie (Hewstone et al, p. 187). Het verband tussen de twee wordt gelegd door de “empathy-altruïsme hypothesis” van Batson, die wordt geopperd als verklaring voor niet-zelfzuchtige vormen van hulpvaardig gedrag. Experimenten van Batson en zijn medewerkers hebben laten zien dat “echt altruïsme”, dat dus niet voortkomt uit egocentrische overwegingen, bestaat.

Hedendaagse evolutionaire en psychologische verklaringen vullen elkaar dus aan: empathie is de psychologische kern van altruïsme. Als psychologische eigenschap is het onzelfzuchtig en moreel goed.

Net zo opvallend is dat evolutie in andere hoofdstukken vrijwel afwezig is. Met Darwin in het achterhoofd is dit met name opmerkelijk in de hoofdstukken die handelen over “Sociale Invloed” (Hewstone et al), “Conformisme” (Aronson et al) of “Normen en Gedrag” (Smith en Mackie). In die hoofdstukken gaat het ondermeer over conformisme, loyaliteit en gehoorzaamheid, sociale instincten die Darwin ook rekende tot de biologische basis voor moraal. Verklaringen voor zulke gedragingen worden in de sociale psychologie gezocht via de menselijke gevoeligheid voor normen. Ter verklaring van de invloed van normen wordt uitvoerig gesproken over de motieven die eraan ten grondslag liggen, zoals behoefte aan eigenwaarde, de behoefte erbij

te horen en het belang van reputatie, maar evolutie komt daarbij niet ter sprake. Moraal is met name aanwezig in de vorm van morele zorgen, want het gedrag dat in deze hoofdstukken aan de orde is dreigt voortdurend in immoreel vaarwater te komen. Gehoorzaamheid, conformisme en loyaliteit verschijnen als moreel gevaarlijke fenomenen.

Deze verdenking hangt rechtstreeks samen met een vaakgenoemde hoofdreden van de naoorlogse opbloei van de sociale psychologie: de fascinatie met de vraag hoe zoveel mensen hebben kunnen meegaan met Hitler. Door Hannah Arendt's *Eichmann in Jerusalem* (1963) werd Eichmann de verpersoonlijking van een antwoord: hij was niet slecht, alleen maar extreem in trekken die op zich heel gewoon zijn; behoefte aan erkenning en de behoefte erbij te horen, en een daarmee samenhangend onvermogen tot zelfstandige oordeelsvorming. Gefascineerd door dezelfde vragen deed Stanley Milgram zijn beroemde experimenten, die ik in de bijlage beknopt heb weergegeven. Uit gehoorzaamheid diende de grote meerderheid van heel gewone mensen in deze experimenten steeds sterkere stroomschokken toe aan een "leerling" die fouten maakte.

26

Gegeven de grote invloed van Arendt's analyse en Milgram's onderzoek hoeft het misschien geen verbazing te wekken dat gehoorzaamheid een slechte naam heeft gekregen en zijn toch al bescheiden plaats op het toneel van de evolutionaire basis voor moraal vrijwel helemaal verloor. Empathie lijkt vandaag de dag te zijn overgebleven als het enige of in ieder geval het voornaamste moreel onverdachte (namelijk pro-sociale) sociale instinct dat de biologie ons heeft gebracht.

Maar deze beperking van de biologische basis is onbevredigend. Om een goed beeld te krijgen van bio- en psychologische instincten of gevoeligheden die voor moraal van belang zijn doen we er goed aan ze niet van meet af aan langs een morele oordeelslat te leggen.

In het vervolg van dit essay wil ik ingaan op twee implicaties voor het denken over de biologische basis voor moraal. In beide gevallen gaat het

om een verbreding. Ten eerste verdienen ook instincten die in diskrediet zijn geraakt, zoals gehoorzaamheid en conformisme, aandacht. Ten tweede verdienen ook de moreel neutrale en de negatieve aspecten van empathie aandacht.

6 GEHOORZAAMHEID (ETC)

Met een pleidooi voor grotere erkenning van gehoorzaamheid, loyaliteit, trouw, conformisme e.d. als krachten die niet alleen relevant zijn voor moraal door haar in de weg te staan, sluit ik mij aan bij het recente werk rond “morele psychologie” van de sociaal psycholoog Jonathan Haidt, die ik in de inleiding al noemde. Haidt speelt een belangrijke rol in de synthese tussen biologie, psychologie en moraal, waarbij hij morele intuïties ziet hij als de bouwstenen voor moraal. Net als evolutiebiologen benadert hij moraal functionalistisch: moraal dient volgens hem zowel handelen in het algemeen als samenwerken in het bijzonder.

29

Ideeën over wat moraal inhoudt zijn in zijn ogen ernstig versmald geraakt. Bij het zoeken van een verklaring gaat hij terug naar de filosofie van de Verlichting, die zowel utilisme als deontologie voortbracht. Beide stromingen probeerden ethiek af te leiden vanuit een enkele regel, die moest leiden tot het beantwoorden van de vraag “wat te doen” in allerlei dilemma-situaties. Ook de psychologie raakte aangetast door deze versmalling tot dilemma’s. In de jaren ‘70 van de afgelopen eeuw draaide het om de vraag of Heinz het medicijn moest stelen om het leven van zijn vrouw te redden en of abortus wel of niet de goede keuze was. Deze versmalling heeft er volgens hem ook toe geleid dat morele psychologie nog maar twee typen morele overwegingen erkent: ten eerste rond rechtvaardigheid en wederkerigheid, en ten tweede rond zorg en schade. De

laatste hebben door het werk van Carol Gilligan terrein gewonnen gewonnen op de eerste, geassocieerd met het werk van Kohlberg (Haidt en Kesebir 2010).

Maar morele intuïties zijn breder en veelsoortiger, betoogt Haidt. Met behulp van psychologische experimenten en door bestudering van antropologische literatuur heeft hij met een aantal co-auteurs geprobeerd een overzicht te creëren van de bouwstenen voor moraal. Dat resulteerde in vijf velden van morele intuïties: 1. schade en zorg, 2. rechtvaardigheid en wederkerigheid, 3. groepsloyaliteit, 4. autoriteit en respect en 5. zuiverheid en heiligheid (o.m. Haidt & Graham, 2007; Haidt & Kesebir 2010).

30

De implicatie is dat de huidige academische benaderingen van moraal, die zich vrijwel uitsluitend op de eerste twee gebieden richten, een aantal belangrijke bouwstenen voor moraal verwaarlozen, namelijk intuïties die te maken hebben met loyaliteit, autoriteit en zuiverheid. Anders dan de meer individualistisch georiënteerde intuïties die wél aandacht krijgen, zijn deze intuïties groepsgericht, en ze zijn door het werk van Milgram en anderen in een kwaad licht komen te staan. Ondertussen zijn ze volop in de samenleving aanwezig, niet alleen in verre of primitieve samenlevingen maar ook bijvoorbeeld in de moraal van Amerikaanse conservatieven. Conservatieven blijken in hun morele overwegingen alle vijf de gebieden te beslaan, terwijl liberalen vooral naar zorg en rechtvaardigheid verwijzen. De academische versmalling van moraal is dan ook gekoppeld aan het liberale deel van de (Amerikaanse) bevolking. Die liberalen bepalen wel voor het overgrote deel het onderzoek naar moraal; zo bleek uit een onderzoek in 2004 dat 95 % van een groep ondervraagde onderzoekers in de geestes- en sociale wetenschappen voor de democraat Kerry had gestemd, en niemand voor Bush. Haidt & Kesebir (2010) stellen dat hierdoor een collectieve blinde vlek is ontstaan voor de groepsgerichte morele intuïties.¹⁵

Ook de Belgische filosoof Jan Verplaetse pleit, in zijn boek *Het Morele Instinct* (2008), voor een verbreding van ons begrip van moraal. Zonder

¹⁵ Verguisde morele intuïties kunnen overigens op onverwachte manieren opduiken. K.A. Appiah wijst er bij voorbeeld op dat onder liberalen veel roep om zuiverheid rond voedsel voorkomt (Appiah 2008; 245).

de pretentie uitputtend te zijn onderscheidt hij een hechtingsmoraal, een geweldmoraal, een reinigingsmoraal, een samenwerkingsmoraal en een beginselenmoraal. De eerste vier daarvan zijn morele systemen die zich hebben gegroepeerd rond natuurlijke instincten. De hechtingsmoraal is in zijn optiek gegrond op empathie, de geweldmoraal op angst en agressie, de reinigingsmoraal op walging en de samenwerkingsmoraal op welwillendheid. Omdat natuurlijke vormen van moraal ons wel motiveren maar uiteindelijk normatief tekort schieten zullen hun dictaten onderdrukt moeten worden: het is de vijfde moraal, de beginselenmoraal, die volgens Verplaetse uiteindelijk uitkomst moet bieden. Deze moraal leidt tot een afweging van verschillende neigingen en argumentaties; “we leven met onze emoties en we beheersen die met onze rede” (Verplaetse 2008; 285).

Milgram's experimenten hebben eraan bijgedragen dat gehoorzaamheid een slechte reputatie kreeg en ze hebben daarmee ook bijgedragen aan de ver-smalling van moraal. Toch beschouwde Milgram zelf gehoorzaamheid niet als iets dat alleen maar remmend werkt op moraal. In feite dacht hij, net als Haidt en Verplaetse, dat gehoorzaamheid een *eigen* gebied van moraal is, of met zich meebrengt, dat in conflict kan komen met andere gebieden. In zijn boek *Obedience to Authority* (2005/1974) bespreekt en interpreteert hij zijn gehoorzaamheidsexperimenten, en de psychologische mechanismen die hij daarin aan het werk ziet. Mensen die zich voegen in een systeem van autoriteit komen volgens zijn analyse terecht in een staat van “agentisme” waarin ze de wensen van anderen uitvoeren en een deel van hun autonomie kwijt raken. De meest vergaande consequentie is dat de betekenis van het handelen, en de verantwoordelijkheid ervoor, bij de autoriteiten komt te liggen. Moraal is dan niet verdwenen, schrijft Milgram, maar heeft een andere focus; de vraag is dan namelijk hoe adequaat iemand heeft uitgevoerd wat de autoriteit verlangt, en dit leidt tot gevoelens van schaamte of trots (p. 147). De morele termen die hier bij horen zijn loyaliteit, plicht, en discipline.

De macht van gehoorzaamheid in Milgram's experimenten overtrof

destijds alle verwachtingen, ook die van Milgram zelf (Milgram 2005/1974, Goethals 2007), omdat gehoorzaamheid hier zo flagrant botste met indringende morele overwegingen. Hoewel de meeste proefpersonen zich ervan bewust waren dat de gevolgen van hun handelingen op den duur verschikkelijk waren (althans, dat dachten ze), en zij rationeel wel tot het besluit kwamen dat ze moesten ophouden, lukte het de meesten niet deze overtuiging in handelen om te zetten. Meer dan 60 % van de deelnemers gehoorzaamde tot het eind, vaak ten prooi aan enorme spanningen. Krachtige maar slechtgedefinieerde banden hielden hen gevangen, schrijft Milgram. De meeste mensen hebben volgens hem een gebrek aan “inner resources” om deze krachten te bestrijden.

32

Deze enorme kracht van gehoorzaamheid bevestigt dat het om een neiging gaat die onder invloed van immorele autoriteiten tot verschrikkingen leidt. Maar dat wil niet zeggen dat gehoorzaamheid op zichzelf immoreel is of dat er geen positieve moraal bij kan horen. Gehoorzaamheid is duidelijk een van de sociale instincten die aan de basis staan van het functioneren in groepen. Vanuit een functionele benadering van ethiek is het daarom niet vreemd dat er morele recepten ontstaan die het functioneren van hiërarchische verbanden mogelijk maken. Gehoorzaamheid, van kinderen aan ouders en leerkrachten, van uitvoerders aan beslissingsbevoegden, van werknemers aan chefs, van burgers aan de oordelen van rechters, van soldaten aan hun superieuren: dit alles is onmisbaar en omgeven met moraal. Binnen welke categorisering van typen moraal dat alles een plaats krijgt is een volgende vraag. De categorisering van Verplaetse heeft geen pretentie van volledigheid en doet ook enigszins arbitrair en heterogeen aan. Die van Haidt heeft meer pretentie, omdat hij is gebaseerd op een systematische analyse van antropologische literatuur. Maar deze vragen naar categorisering laat ik verder open. Het gaat mij er vooral om dat ook sociale neigingen van de mens die onaantrekkelijke of gevaarlijke kanten hebben op positieve manieren met moraal verbonden zijn. Gehoorzaamheid is daarvan een voorbeeld.

7 EMPATHIE

Ook empathie is een voorbeeld van een instinct met morele beperkingen en keerzijden. De erkenning daarvan is de tweede verbreding die het denken over de biologische basis van moraal nodig heeft. Empathie geven en nodig hebben zijn met moraal verbonden langs verschillende wegen. Laat ik kort naar elk van deze wegen kijken, om vervolgens te bepleiten dat de diverse mechanismen in samenhang bestudeerd worden, bijvoorbeeld onder de noemer “de wetten van de empathie”, of wellicht eenvoudig reciprociteit.

33

Over het centrale morele belang van het “*geven*” van sympathie en/of empathie is geen twijfel mogelijk. Het is de basis van emotionele betrokkenheid, van verbondenheid, zorg, compassie, verantwoordelijkheid en hulp. Het is moeilijk voor te stellen wat sociaal leven zonder empathie zou moeten inhouden. De kanttekeningen die volgen hebben niet de functie om het morele belang van empathie te ontkennen, integendeel; ze ontstaan juist door de grote kracht ervan.

Dat de morele implicaties van empathie niet alleen maar positief zijn komt vooral doordat empathie op uiterst gespannen voet staat met onpartijdigheid. Empathie veronderstelt nabijheid, in enige vorm. Empathie met mensen ver weg, die we niet kennen en niet kunnen zien, vereist veel meer inspanning dan empathie met mensen die dichtbij, zichtbaar en vertrouwd zijn.

Daniel Batson en co-auteurs (1995) trekken op basis van twee experimenten de conclusie dat “empathie-geïnduceerd altruïsme” tot immorele consequenties kan leiden, bij voorbeeld tot het bieden van schaarse hulp aan mensen die die hulp niet het meest nodig hebben. Proefpersonen kregen onder meer interviews te zien met zieke kinderen en lieten zich vervolgens in het vaststellen van prioriteiten in de hulp meeslepen door empathie ten aanzien van deze kinderen. Het verschijnsel is, zoals de auteurs in de discussie opmerken, alom aanwezig. Beelden van lijdende kinderen op tv leiden tot uitbarsingen van hulpgedrag, maar zijn niet gebaseerd op enige analyse van de vraag waar hulp het hardst nodig is. Alleen wie in beeld komt maakt kans op empathie-geïnduceerd hulpgedrag. Deze eenvoudige basisvoorwaarde voor empathie is een rechtstreekse achtergrond van media-oorlogen op het gebied van internationale hulpverlening. Zonder beelden geen hulp.

34

De relatieve ongevoeligheid voor mensen die ver weg zijn of buiten beeld is een psychologisch gegeven met moreel rampzalige gevolgen. De Milgram-experimenten lieten de invloed van nabijheid duidelijk zien: als de proefpersoon en de “leerling” zich in dezelfde kamer bevonden nam de bereidheid tot het toedienen van schokken sterk af, en als zij elkaar konden aanraken was dit effect nog sterker. Milgram merkt naar aanleiding daarvan op dat niets zo gevaarlijk is voor het menselijk voortbestaan als immorele autoriteiten in combinatie met het dehumaniserende effect van psychologische buffers, in de vorm van tijd, afstand of fysieke barrières. Een mens stenigen is psychologisch veel moeilijker dan het droppen van bommen vanuit een vliegtuig; tegen dat laatste bestaan haast geen psychologische remmingen (Milgram 2005/1974; 158-59).

De keerzijde van de emotionele banden die empathie smeedt zetelt dan ook in de onvermijdelijke partijdigheid daarvan. In antwoord op de vraag wat hij aan de menselijke natuur zou willen veranderen als hij God was schrijft De Waal “If I were God, I’d work on the reach of empathy.” (De Waal 2009; 204).

Dat we sympathie *nodig hebben* is moreel gezien op zijn best neutraal; het

kan ook leiden tot een Mandeville-achtige cynische visie op de oorsprong van moraal. De zorg voor reputatie die door deze behoefte centraal komt te staan krijgt ook in de huidige psychologie veel aandacht. Dit gebeurt niet in de context van empathie, maar van de sociale functionaliteit van moraal. Jonathan Haidt brengt die sociale functionaliteit naar voren in zijn uitgangspunten dat moreel denken in dienst staat van sociaal handelen, en dat het een bindende functie heeft.

Als moraal in dienst staat van ons sociale functioneren dan gaat het ons in de praktijk eerder om het handhaven van een positieve identiteit dan om het vinden van morele waarheid. Daarom stellen Haidt en Kesebir (2010) voor, onder verwijzing naar het werk van Philip Tetlock¹⁶, mensen met betrekking tot hun sociale motivatie niet te karakteriseren als intuïtieve wetenschappers maar als intuïtieve politici, die voortdurend bezig zijn rekenschap af te leggen. Ze verwijzen naar het grapje dat geweten de innerlijke stem is die ons waarschuwt dat er misschien wel iemand kijkt. Maar binnen de metafoor van de intuïtieve politicus die rekenschap af te leggen heeft is dat meer dan een grapje. Interne dialoog, waarin mensen anticiperen op tegenwerpingen en alvast antwoorden proberen te formuleren, vervult voor een intuïtieve politicus een sleutelfunctie in het denken, schrijft Tetlock (2002; 452). Haidt en Kesebir nemen de politicusmetafoor over en presenteren een review van onderzoeksresultaten waaruit blijkt dat mensen bereid zijn de waarheid te verdraaien als dat hun reputatie en/of zelfbeeld ten goede kan komen.¹⁷

Reputatiemanagement valt ook te ontdekken in de manieren waarop mensen zich onttrekken aan morele verantwoordelijkheid door het reframen van situaties in hun eigen voordeel, die uitvoerig zijn bestudeerd door Albert Bandura; het gaat dan bijvoorbeeld om het dehumaniseren van slachtoffers, slachtoffers de schuld geven, eufemismen gebruiken, de gevol-

¹⁶ Philip Tetlock (2002) bespreekt functionele aannames waarmee psychologen keuzen en oordelen begrijpen. Traditionele zienswijzen waarin mensen in de eerste plaats als intuïtieve wetenschappers werden gezien, voor wie het bij normatieve oordelen om de waarheid gaat, hebben aan invloed verloren. In de plaats daarvan stelt Tetlock een functioneel pluralisme voor: hij beschrijft drie functionele frames, waarin menselijk respectievelijk worden voorgesteld als metaforische politici, theologen en aanklagers.

¹⁷ In een van die experimenten, ook weer van Batson en zijn medewerkers, mochten proefpersonen bij voorbeeld taken verdelen tussen zichzelf en een ander, waarbij de ene taak veel aantrekkelijker was dan de andere. De proefpersonen mochten de taken simpelweg toekennen, of, als ze dat liever wilden, erom loten door een munt op te gooien die in een verzegeld zakje zat. Degenen die het zakje niet openden kenden zichzelf in 80-90 % van de gevallen de aantrekkelijke taak toe. Maar dat gold ook voor degenen die het zakje wel hadden geopend, en dus (?) het lot hadden laten beslissen.

gen minimaliseren en diffusie of ontkenning van verantwoordelijkheid (zie o.a. Bandura 1999). Daarnaast is reputatie een belangrijk onderwerp in de gedragseconomie, een groot en groeiend gebied van fascinerende inzichten en experimenten dat hier ook een omweg waard zou zijn.¹⁸

Mijn pleidooi zal niet meer verrassen: reputatiemanagement valt te formuleren in termen van behoefte aan empathie, en de beide richtingen van empathie, het geven en het nodig hebben, verdienen bestudering in samenhang. Het werk van Darwin en Adam Smith, waarin wederkerigheid rondom sympathie centraal stond, biedt daarbij nog steeds belangrijke aanknopingspunten.

¹⁸ In *The Economy of Esteem* vragen Brennan en Pettit (2004) aandacht voor de herontdekking van achting als een fundamentele menselijke drijfveer. Zie b.v. ook Ellingsen and Johannesson (2007), een programmatisch artikel in deze zelfde stroom.

8 SITUATIES EN INNERLIJKE KRACHTEN

Het idee dat denken en moraal overheersend bewuste en rationele processen zijn heeft terecht veel terrein verloren. Denken is lichamelijk en functioneel en dit leidt tot erkenning van de grote cognitieve rol van imperfecte instrumenten als intuïties, beelden, metaforen, heuristiek en allerlei vormen van bias. In welke mate kunnen we voorkomen dat we de speelbal zijn van die processen?

37

De bekendste bias uit de sociale psychologie heet “correspondentie bias”, vroeger aangeduid als “fundamentele attributiefout”. Deze bias houdt in dat mensen geneigd zijn het gedrag van andere mensen veel meer dan cognitief verantwoord is toe te schrijven aan interne factoren, zoals karaktereigenschappen. Ze onderschatten in het beoordelen van het gedrag van anderen consequent het belang van de situatie. De sociale psychologie neemt dan ook stelling tegen een psychologie en moraal die de verschillen tussen mensen vooral toeschrijft aan verschillen in persoonlijkheidskenmerken, karaktereigenschappen of deugden. In *Lack of Character* (2002) bespreekt John Doris implicaties van situationistische inzichten voor moraalfilosofie, met name de deugdenethiek. Hij bespreekt enkele van de grote sociaal-psychologische experimenten, waaronder die van Milgram.¹⁹ Verklaringen op basis van persoonlijkheid of disposities kunnen geen goede verklaringen geven voor het gedrag van Milgram’s proefpersonen, concludeert hij, en de experimenten geven

¹⁹ Daarnaast gaat hij o.a. in op het Stanford Gevangenisexperiment van Zimbardo, en het Goede Samaritaan-experiment van Darley en Batson.

daarom aanleiding te twijfelen aan de robuustheid van disposities. Het immorele gedrag in de experimenten moet in de eerste plaats worden verklaard door kenmerken van de omgeving. In zijn interpretatie van de experimenten is die omgeving niet heel dwingend; hij legt er nadruk op dat de proefleider beleefd was en dat de proefpersonen konden ophouden als ze dat wilden. Met deze interpretatie past het experiment bij wat situationisme in zijn ogen inhoudt, namelijk dat omgevingsfactoren die op zichzelf niet bepaald indrukwekkend zijn disproportioneel grote effecten kunnen hebben op wat mensen doen (Doris 2002; 28, 49). Mensen zijn zich van veel situationele invloeden nauwelijks bewust, schrijft Doris, en door die alarmerend fragiele greep op ons motivationele universum (p. 139) leven we in een wereld vol onverwachte morele gevaren. Zo bevinden wij ons misschien wel veel vaker dan we ons realiseren in Milgram-achtige situaties, tijdens een seminar bijvoorbeeld, of een vergadering. Zijn aanbeveling is daarom onze morele aandacht te verleggen van karaktertrekken naar die aspecten van de omgeving die ons gedrag beïnvloeden (p. 146).

38

Aandacht voor de morele verleidingen van situaties lijkt me een uitstekende aanbeveling. Toch is een verwijzing naar “de situatie” op zichzelf onbevredigend. Doris’ boek lezend zou je haast denken dat de morele gevaren werkelijk in alle kleine hoekjes kunnen loeren, en dat we er geen flauw idee van hebben. Maar valt er niet meer te zeggen over de specifieke hoeken waaruit we de gevaren kunnen verwachten? Is dit niet precies waar onze sociale instincten in zicht komen? Gehoorzaamheid? De behoefte gezien, erkend, geliefd en geprezen te worden? Als we ervan uitgaan dat autoriteitsrelaties een belangrijk veld van sociale krachten zijn krijgen de situationele stimuli in de Milgram-experimenten een andere lading: nog steeds niet dramatisch in hun fysieke verschijning, maar wel in hun motivationele betekenis.

Doris’ nadruk op situaties vormt een complement van Milgram’s eigen interpretatie. Volgens Milgram hebben mensen te maken met een tekort aan “in-

ner resources” om de kracht van autoriteiten te weerstaan. De meeste mensen weten op zich wel wat ze zouden moeten doen, zoals duidelijk bleek toen hij aan psychiaters, studenten en een heterogene groep volwassenen vroeg hoe ze verwachtten dat ze zich in zijn experiment zouden gedragen. De meesten verwachtten dat ze niet verder zouden gaan dan 150 volt, en niemand dacht verder te zullen gaan dan 300 volt. Algemeen verwachtte men zich door empathie en rechtvaardigheid te zullen laten leiden (Milgram 2005/1974; 29). Ook als het ging om de vraag wat anderen zouden doen verwachtten de meeste respondenten dat alleen een heel kleine en pathologisch gestoorde minderheid, niet meer dan 1 of 2 %, tot het eind zou doorgaan. Onder dergelijke verwachtingen ligt volgens Milgram niet alleen de aanname dat mensen in grote lijnen fatsoenlijk zijn, maar ook de aanname dat gedrag, behalve in geval van fysiek geweld, uit eigen overtuigingen voortkomt.

In een later hoofdstuk beschrijft Milgram de obstakels die mensen tegenkomen bij het mobiliseren van die eigen overtuigingen. De meeste mensen hebben te maken met een toenemende spanning over de discrepantie tussen wat hen wordt opgedragen en de impact van hun daden. Omdat ongehoorzaamheid psychologisch zo moeilijk is, is ontkenning de eenvoudigste manier om deze spanning op te lossen. Ook wordt de spanning vaak omgezet in fysieke ontladingen als zweten, trillen of nerveus lachen. Ongehoorzaamheid is uiteindelijk een veel effectievere manier om de spanning op te lossen, maar er zijn veel barrières te nemen; de stappen die Milgram beschrijft zijn het uiten van twijfel, het uiten van afkeuring, het dreigen met stopzetten van de medewerking, en uiteindelijk echte weigering. Slecht een minderheid van de mensen is hiertoe in staat. Voor de meesten blijkt het verbreken van de belofte aan de autoriteit te moeilijk.

Mensen toonden dus weinig inzicht in het web van krachten dat in werkelijke situaties werkzaam is. In het voorwoord bij zijn boek geeft Milgram aan hoe akelig de door hem bestudeerde situaties zijn: “Nothing is bleaker than the sight of a person striving yet not fully able to control his own behaviour in a situation of consequence to him.”(p. xix).²⁰

²⁰ Een soortgelijke “bleakness”, maar dan in een veel onontkoombare situatie, beschrijft Sebastian Haffner in *Het Verhaal van een Duitser* (2002), in verband met de manieren waarop gewone Duitsers zich probeerden te verhouden tegenover de autoriteit van de nazi’s. “De situatie van de Duitsers die geen nazi waren, was in de zomer van 1933 stellig een van de moeilijkste waarin mensen zich kunnen bevinden, namelijk een toestand van volledig en reddeloos overmeesterd zijn, in combinatie met de nawerkingen van de shock van de ergste overrompeling”

Dat een dergelijk gebrek aan inzicht, in situaties en in onze innerlijke reacties erop, van meer dan academisch belang is behoeft geen betoog, en dit leidt tot het idee dat het nuttig zou zijn inzicht en weerbaarheid ten aanzien van zulke situaties te trainen. Philip Zimbardo reflecteert in *The Lucifer Effect* (2007) op het Stanford Prison Experiment, waarin het onder meer conformisme was dat tot immorele daden leidde en waarvan hij de actualiteit en bredere betekenis illustreert met de martelingen in de gevangenis van Abu Ghraib. In het laatste hoofdstuk bepleit hij het versterken van onze weerstand tegen dergelijke situationistische verleidingen via een training tot “alledaags heldendom”. Sleutels ziet hij in een bewuste houding, een gevoeligheid voor verleidingen die uit de situatie voortkomen, en in wijsheid van de straat. De tien stappen die hij als oefenprogramma aanbeveelt zijn aanmoedigen voor individuen om zich te oefenen in innerlijke onafhankelijkheid en integriteit. Zo bestaat de eerste stap uit de bereidheid om fouten toe te geven, volgende stappen roepen bij voorbeeld op tot een “mindful” en verantwoordelijke opstelling. Een van de stappen draagt het motto “I want group acceptance, but I value my independence.”

Hans Achterhuis heeft het in een recent interview in *Filosofie Magazine* (De Wit, 2009) ook over het belang van zulk alledaags heldendom. Hij beschrijft zijn verbazing over zijn eigen passiviteit in een aantal situaties; als iemand luidruchtig aan het bellen is in een stiltecoupé zal hij bij voorbeeld maar stiltejes ergens anders gaan zitten. Zoekend naar verklaringen voor zijn gedrag verwijst hij naar conformisme en ander sociaal-psychologische mechanismen zoals het “bystander effect”, dat inhoudt dat de kans dat omstanders ingrijpen kleiner is naarmate ze met meer zijn. Om wél in actie te komen is ten eerste bewustwording nodig en ten tweede actie-vaardigheid: weten hoe je effectief en niet-escalierend kunt handelen in een lastige situatie. Zoiets vereist oefening en training, zoals bij voorbeeld politie-agenten die in hun opleiding krijgen. Zij leren, als het goed is, kalm te blijven, mensen tot de orde te roepen zonder meteen agressief over te komen, tegelijk vriendelijk en standvastig

(p. 190). Haffner beschrijft een aantal verleidingen om deze spanning te reduceren op niet-heldhaftige manieren, waarvoor de meeste mensen bezweken.

te zijn. Ook de minder heldhaftigen onder ons kunnen door oefeningen en met vallen en opstaan toch leren zulke obstakels als het bystander effect te overwinnen, zegt Achterhuis: “je moet het vooral een paar keer meegemaakt hebben.”

Gehoorzaamheid, conformisme en het afschuiven en verdunnen van verantwoordelijkheid staan bekend als gevaarlijke neigingen. Het idee dat we inzicht en training nodig hebben om deze gevaren te weerstaan ligt dus voor de hand. Empathie lijkt een heel ander geval: het heeft zo’n goede reputatie dat argwaan nogal vreemd lijkt, misschien zelfs wel moreel verdacht. Hoogstens lijken we het, met Frans de Waal, te kunnen betreuren dat empathie niet een grotere reikwijdte heeft.

Maar ik denk dat ook de dagelijkse mechanismen van empathie meer inzicht kunnen gebruiken, volgens de lijnen die ik hierboven heb geschetst. In een geglobaliseerde wereld waarin de media ons alle uithoeken van de aarde kunnen tonen heeft het spel van het geven en nemen van empathie heel nieuwe vormen gekregen. Het leidt tot een media-strijd die bepaald geen garantie is voor de meest evenwichtige verdeling van aandacht en van de inspanningen die daarop volgen.²¹ Beelden zijn heel effectief, ze bespelen onze empathie en creëren daarmee hypes en blinde vlekken. Om te beginnen zijn de dingen van thuis nu eenmaal dichterbij, en wat betreft de gebeurtenissen van verder weg gaat de aandacht voor plotselinge rampen zoals de tsunami ten koste van aandacht voor meer sluipende, minder mediagenieke maar in hun gevolgen veel grotere rampen. Aandacht van de media is door de effecten op materiële hulp op zichzelf inzet van strijd geworden, waarbij het er om gaat de empathie van de toekijkende wereld te winnen, of (voor anderen) de wereld daar juist van weg te houden.

Het effect van de decennialange oproepen tot mondiale hulp van Peter Singer is teleurstellend geweest. Maar in zijn laatste boek, *The Life You Can Save* (2009), sluit hij aan bij de toenadering tussen moraal, psychologie en biolo-

²¹ In *Het Zijn Net Mensen* (2006) beschrijft Joris Luyendijk voorbeelden van de media-strijd uit Israël en de Palestijnse gebieden. Aart Zeeman verzucht in *Tot Zover Darfur* (2007) “Geen beelden, geen verhaal, geen aandacht. Dan bestaat de oorlog in Darfur niet” (p.29).

gie en probeert hij zijn oproep psychologisch realistischer te maken. In twee hoofdstukken over “Human Nature” stelt hij bij voorbeeld voor om giften publiek bekend te maken, zodat er prikkels ontstaan vanuit onze behoefte aan een goede reputatie. Ook pleit hij ervoor onze empathie te helpen richten door de armen een gezicht te geven. Een kind in nood prikkelt ons immers meer dan de statistieken. Zijn voorstellen komen neer op een uitbreiding van de nu zo populaire strategie van *Nudge* (Thaler & Sunstein 2008): gebruik de groeiende inzichten in de menselijke natuur ten behoeve van een “zacht paternalisme”, dat ons helpt goed te doen en gelukkig te worden, zonder ons van onze vrijheid te beroven. Het is een veelbelovende strategie, die inderdaad aansluit bij de verbindingen tussen biologie, psychologie en moraal waarover dit essay gaat. Maar genoeg is het niet. De verdeling van aandacht en middelen in de wereld is van groot moreel belang²² en kan het niet stellen zonder inzicht in de wereld en in onszelf. Het vrije spel van empathie-bespelende krachten trekt zich daar niets van aan, maar ook de goede bedoelingen van bestuurders en fundraisers schieten tekort als het om inzicht gaat.

42

Empathie heeft dankzij de inspanningen van Frans de Waal een ereplaats in de biologische basis van moraal. Maar ook empathie heeft zijn morele voetangels en klemmen, en onze weerbaarheid daartegen valt wellicht te vergroten via training, op basis van onderzoek in het voetspoor van Darwin en Adam Smith.

²² Over een “ethiek van aandacht” heb ik ook geschreven in de context van nature-nurture vragen in de embryologie (bv Van der Weele 1999) en van de ethiek rond genomics (Van der Weele 2008 en <http://tenninetygap.wordpress.com/>). Een journalistiek initiatief ter bevordering van een meer evenwichtige verdeling van aandacht is *Het Vergeten Verhaal*, een project van de Dick Scherpenzeelstichting over blinde vlekken in de journalistiek, met als doel een betere berichtgeving over ontwikkelingsvraagstukken.

DANKWOORD

De NVBe concentreert zich volgens haar website op morele vragen die ontstaan in de levenswetenschappen. In dit essay waren de levenswetenschappen op een heel andere manier aan de orde, namelijk ter mogelijke verheldering van moraal. Ik dank het bestuur van de Nederlandse Vereniging voor Bio-ethiek, en in het bijzonder Dick Willems, van harte voor de uitnodiging dit pre-advies te schrijven en voor de ruimte die ik daarbij kreeg. Uiteraard hoop ik dat het resultaat sympathie zal ontmoeten. Dat zou mij in dit verband liever zijn dan empathie.

43

Verder wil ik vooral Henk van den Belt heel hartelijk bedanken. Zonder hem zou dit essay er niet zijn gekomen: hij wees mij op de Studium Generale lezingenserie, dacht aan mijn voordracht in verband met een mogelijk pre-advies over de biologische basis van moraal, attendeerde mij op het proefschrift van Kerkhof en las de conceptversies. Tenslotte dank ik Clemens Driessen, Joël van der Weele en Bruno van der Pol van harte voor hun diverse en waardevolle suggesties.

REFERENTIES

- Appiah, K.A. (2008): *Experiments in Ethics*. Cambridge (MA), Harvard U.P.
- Arendt, H. (1963): *Eichmann in Jerusalem*. Harmondsworth, Penguin.
- Aronson, E., T.D. Wilson and R.M. Akert (2007): *Social psychology*. 6th edition. Upper Saddle River (NJ), Pearson Education.
- Bain, A. (1868): *Mental and Moral Science. A Compendium of Psychology and Ethics*. London, Longmans, Green & Co.
- Bandura, A. (1999): "Moral Disengagement in the Perpetration of Inhumanities." *Personality and Social Psychology Review* 3, 193-209.
- Batson, C.D. et al (1995): "Immorality from Empathy-Induced Altruism: When Compassion and Justice Conflict." *Journal of Personality and Social Psychology* 68 (6); 1042-1054.
- Batson, C.D. et al (2007): "Altruism and Helping Behavior." In: M.A. Hogg and J. Cooper (eds): *The SAGE handbook of Social Psychology*. London, Sage.
- Boyd, R. and P.J. Richerson (2005): *The Origin and Evolution of Cultures*, Oxford, Oxford U.P.
- Brennan, G. and P. Pettit (2004): *The Economy of Esteem*. Oxford, Oxford U.P.
- Damasio, A. (1994): *Descartes' Error. Emotion, Reason and the Human Brain*. New York, Putnam.
- Darwin, C. (2004/1871): *The Descent of Man, and Selection in Relation to*

Sex. With an introduction by J. Moore and A. Desmond. London, Penguin.

- Doris, J.M. (2002): *Lack of Character. Personality and Moral Behavior*. New York, Cambridge U.P.
- Eisenberg, N. and J. Strayer (1987): *Empathy and its Development*. Cambridge (Ma), Cambridge U.P.
- Ellingsen, T. and M. Johannesson (2007): "Paying Respect." *Journal of Economic Perspectives* 21 (4); 135-149.
- Goethals, G.R. (2007): "A Century of Social Psychology: Individuals, Ideas, and Investigations." In: M.A. Hogg & J. Cooper (eds): *The SAGE handbook of Social Psychology*. London, Sage.
- Haffner, S. (2002): *Het Verhaal van een Duitser 1914-1933*. Amsterdam, Mets en Schilt.
- Haidt, J. (2001): "The Emotional Dog and its Rational Tail: A Social Intuitionist Approach to Moral Judgment." *Psychological Review* 108, 814-834.
- Haidt, J. (2007): "The New Synthesis in Moral Psychology." *Science* 316, 998-1002.
- Haidt, J. and J. Graham (2007): "When Morality Opposes Justice: Conservatives Have Moral Intuitions that Liberals May not Recognize." *Social Justice Research* 20 (1): 98-116.
- Haidt, J. and S. Kesebir (2010, in press). "Morality." In: S.T. Fiske, D.Gilbert & G. Lindzey (eds): *Handbook of Social Psychology*, 5th edition, Vol 2. Hoboken (NJ), Wiley.
- Hauser, M.D. (2006): *Moral Minds. How Nature Designed Our Universal Sense of Right and Wrong*. New York, Harper Collins.
- Hewstone, M, W. Stroebe and K. Jones (2008): *Introduction to Social Psychology. A European Perspective*. 4th edition. Malden, Blackwell Publishing.
- Joyce, R. (2006): *The Evolution of Morality*. Cambridge (Ma), MIT Press.
- Kerkhof, A.J. (1992): *De Mens Is een Angstig Dier. Adam Smith' Theorie van de Morele Gevoelens*. Meppel, Boom.
- Koelewijn, J. en H. Spiering (2009): "Iets Moois als Empathie." *NRC Handelsblad* Zaterdag 5 december.

- Luyendijk, J. (2006): *Het Zijn Net Mensen. Beelden uit het Midden-Oosten*. Amsterdam, Podium.
- Milgram, S. (2005/1974): *Obedience to Authority*. London, Pinter & Martin.
- Singer, P. (2009): *The Life You Can Save. Acting Now to End World Poverty*. New York, Random House.
- Smith, A. (1976/1759): *The Theory of Moral Sentiments*. With an Introduction of D.D. Raphael and A.L. Macfie. Oxford, Clarendon Press.
- Smith, E.R. and D.M. Mackie (2007): *Social Psychology*, 3rd edition. New York, Psychology Press.
- Smith, V.J. (1998): "The Two Faces of Adam Smith." *Southern Economic Journal* 65 (1), 1-19.
- Tetlock, P.E. (2002): "Social Functionalist Frameworks for Judgment and Choice: Intuitive Politicians, Theologians and Prosecutors." *Psychological Review* 109 (3); 451-471.
- Thaler, R.H & C.R. Sunstein (2008): *Nudge. Improving Decisions about Health, Wealth and Happiness*. New Haven, Yale U.P.
- Verplaetse, J. (2008): *Het Morele Instinct. Over de Natuurlijke Oorsprong van onze Moraal*. Amsterdam, Uitgeverij Nieuwezijds.
- Vorstenbosch, J. (1990): *Toestemming voor Onderzoek*. Nijmegen, Proefschrift K.U. Nijmegen.
- Waal, F. de (1996): *Good Natured. The Origins of Right and Wrong in Humans and Other Animals*. Cambridge (Ma), Harvard U.P.
- Waal, F. De (2006): *Primates and Philosophers. How Morality Evolved*. Princeton, Princeton U.P.
- Waal, F. de (2009): *The Age of Empathy. Nature's Lessons for a Kinder Society*. New York, Harmony Books.
- Weele, C. van der (1999): *Images of Development. Environmental Causes in Ontogeny*. Albany (NY), Suny Press.
- Weele, C. van der (2008): "Moral Agendas for Genomics: How to Find the Blind Spots?" *Configurations* 16 (1); 117-135.
- Weele, C. van der (2009): *The 10-90 Gap: What Can We Do?*

<http://tenninetygap.wordpress.com/>

- Williams, G. (1966): *Adaptation and Natural Selection: A Critique of some Current Evolutionary Thought*. Princeton, Princeton U.P.
- Wilson, D.S. and Wilson, E.O. (2008): "Evolution for the Good of the Group." *American Scientist* 96 (5); 380-389.
- Wilson, E.O. (1975): *Sociobiology*. Cambridge (Ma), Harvard U.P.
- Wit, A. de (2009): "Ik had natuurlijk moeten ingrijpen". *Filosofie Magazine* 18 (8); 13-15.
- Zeeman, A. (2007): *Tot Zover Darfur*. Amsterdam, Querido.
- Zimbardo, P. (2007): *The Lucifer Effect. How Good People Turn Evil*. New York, Random House.

DE GEHOORZAAMHEIDSEXPERIMENTEN VAN MILGRAM

Stanley Milgram voerde aan het begin van de jaren zestig aan de universiteit van Yale een aantal experimenten uit rond gehoorzaamheid. Via advertenties in kranten wierf hij proefpersonen, zogenaamd voor een experimenteel onderzoek naar geheugen en leren. In een gemanipuleerde procedure werd bepaald wie leerling werd en wie leraar; een medewerker werd de leerling, de proefpersoon werd de leraar. De proefpersonen kregen te horen dat het in dit experiment ging om de invloed van straffen op het leerproces van mensen. De leraar/proefpersoon moest paren woordjes voorlezen, die de leerling moest proberen te onthouden. Als de leraar vervolgens een woord uit het rijtje gaf moest de leerling het bijbehorende woord geven. De straf voor een fout antwoord bestond uit een stroomschok, toe te dienen door de leraar, die daartoe een bord met switches voor zich had; voor ieder fout antwoord een switch. Het voltage van de schok werd iedere keer 15 volt hoger: het begon bij het eerste foute antwoord met 15 volt, het eindigde met 450 volt. Bij elke switch stond het voltage genoemd en per groepje van vier waren de switches ook voorzien van een omschrijving, die opliep van "slight shock" via aanduidingen als "strong shock" en "danger: severe shock" tot XXX bij de allerhoogste voltages.

De leraar krijgt een proefschok van 45 volt, om een indruk te krijgen, en krijgt te horen dat de schokken pijnlijk kunnen zijn maar geen blijven-

de schade veroorzaken. De leraar en leerling bevinden zich in aparte ruimten; de leraar kan de leerling horen maar niet zien. De proefleider bevindt zich bij de proefpersoon.

Foute antwoorden zijn in de opzet van de proef ingebouwd, en de stroomsterkte neemt dus geleidelijk toe. Na een tijdje geeft de leerling te kennen dat de stoten pijnlijk zijn (van 90 tot 135 volt). Vanaf 150 volt reageert hij met steeds sterkere uitingen van pijn, en laat weten niet meer te willen meewerken. Bij 300 volt gilt hij dat hij geen antwoord meer zal geven. Vanaf 330 volt laat de leerling helemaal niets meer horen.

Voor als de proefpersoon, de leraar dus, aarzelt om door te gaan met het toedienen van schokken, heeft de proefleider een serie zinnen paraat, steeds rustig uit te spreken: 1. Gaat u alstublieft door 2. Het experiment vereist dat u doorgaat. 3. Het is absoluut noodzakelijk dat u doorgaat en 4. U hebt geen andere keus, u *moet* doorgaan. Als de leraar vraagt of de stroomstoten geen kwaad kunnen zegt de proefleider “Hoewel de schokken pijnlijk kunnen zijn is er geen permanente weefselschade, dus gaat u alstublieft verder”. Als de leraar opmerkt dat de leerling niet meer wil meewerken zegt de proefleider “Of de leerling het leuk vindt of niet, u moet doorgaan tot hij alle woordparen correct heeft geleerd, dus gaat u alstublieft door.” “Geen antwoord is een fout antwoord,” krijgt de leraar te horen als de leerling niets meer zegt, “gaat u door.” Zodra de proefpersoon weigert door te gaan, eindigt het experiment.

In werkelijkheid werden er geen stroomstoten uitgedeeld. Zowel de foute antwoorden als het gekreun, geschreeuw e.d. waren voorgeprogrammeerd. Maar de misleiding was zeer overtuigend en bracht veel van de proefpersonen/leraren in een staat van grote spanning. Alle proefpersonen ontmoetten na afloop de ongedeerde “leerling” weer. Bovendien werd er uitgebreid met hen gesproken over de ware aard van het experiment.

De mate van gehoorzaamheid overtrof alle voorspellingen van zowel deskundigen als leken: bijna tweederde van de proefpersonen ging door tot 450 volt. In een serie vervolgonderzoeken onderzocht Milgram de invloed van allerlei variabelen. Sommige variabelen (bv het geslacht van de proefper-

soon) bleken weinig of geen verschil te maken. Maar als de proefpersoon en de leerling zich in dezelfde kamer bevonden daalde de gehoorzaamheid naar ongeveer 40 % . Als de proefleider afwezig was en de instructies per telefoon gaf, daalde de gehoorzaamheid nog sterker, naar ongeveer 20% .

Een uitgebreide beschrijving van en reflectie op het onderzoek is te vinden in Milgram's *Obedience to Authority* uit 1974, dat in dit essay herhaaldelijk ter sprake komt. Het experiment gaf ook aanleiding tot hevige discussie over de ethiek van onderzoek met proefpersonen, een discussie die ik laat liggen. Jan Vorstenbosch geeft in het laatste hoofdstuk van zijn proefschrift *Toestemming voor Onderzoek* (1990) een helder overzicht.