

Vrije wil en verantwoordelijkheid in evolutionair perspectief*

Arno Wouters

Erasmus Universiteit Rotterdam

Van sommige menselijke gedragingen kunnen we ons nauwelijks voorstellen hoe ze te verenigen zijn met de evolutietheorie. Denk bijvoorbeeld aan een echtpaar dat een niet-verwant kind adopteert. De keus om van eigen kinderen af te zien teneinde een wees een gelukkiger leven te bieden is op het eerste gezicht moeilijk te verenigen met een theorie volgens welke neigingen die helpen om de eigen genen te verspreiden zich zullen uitbreiden ten koste van neigingen die vreemde genen bevoordelen. Zelfs in het geval van onvruchtbaarheid is de verspreiding van de genen gediend met ander gedrag: de onvruchtbare partner zou zijn of haar geld en energie beter kunnen investeren in de kinderen van familieleden (deze hebben immers ten dele dezelfde genen), terwijl de vruchtbare partner beter een ander zou kunnen zoeken.

Dit soort voorbeelden is gemakkelijk uit te breiden. Denk aan mensen die voor een celibatair leven kiezen, soldaten die zich opofferen voor het vaderland, zelfmoordterroristen en mensen die met gevaar voor eigen leven hulp bieden bij ongelukken en rampen. Dergelijke keuzen lijken erop te duiden dat mensen de vrijheid hebben zich te onttrekken aan wat hun genen voorschrijven. Die vrijheid lijkt een noodzakelijke voorwaarde voor verantwoordelijkheid, één van de drie begrippen van vrije wil die in de inleiding van dit boek onderscheiden zijn. Het lijkt immers volslagen bizar om mensen verantwoordelijk te houden voor hun gedrag als dat gedrag bepaald wordt door een in hun genen vastgelegde reactie op de omgeving.

De moeilijkheid om te begrijpen hoe deze vrijheid in het licht van de evolutietheorie kan bestaan, roept echter de vraag op of deze voorbeelden bij nadere beschouwing inderdaad op die vrijheid duiden. Misschien zijn dit soort keuzen niet meer dan een incidenteel nadelig effect van een neiging die in het algemeen de verspreiding van onze genen dient. Een echtpaar dat op vakantie impulsief een weeskind adopteert handelt misschien tegen het ‘belang’ van hun genen, maar niet noodzakelijk tegen hetgeen hun genen voorschrijven. In de *March of the Penguins* (2005), een documentaire die vijf jaar geleden volle bioscoopzalen trok, zien we een pinguïnmoeder die haar jong verloor het jong van haar buurvrouw kidnappen. De verspreiding van haar genen zou meer gediend zijn met een terugkeer naar zee zodat zij het volgende broedseizoen extra gevoed tegemoet kan treden. Kennelijk is in de barre omstandigheden waarin deze pinguïns leven de ontwikkeling van een zeer sterke neiging om een jong te verzorgen zo belangrijk dat de enkele keer dat deze neiging ongunstig uitwerkt er niet toe doet.

* Manuscript van “Vrije wil en verantwoordelijkheid in evolutionair perspectief”, verschenen in Maureen Sie (red.) *Hoezo vrije wil? Perspectieven op een heikele kwestie* (Uitgeverij Lemniscaat, 2011), hoofdstuk 10, pp. 190–209, ©Arno Wouters.

De gedachte dat onze praktijk van verantwoordelijkheid in het licht van de evolutietheorie niet te handhaven is wordt door biologen en neurowetenschappers met veel aplomb uitgedragen. Misschien herinnert u zich het stuk van de gedragsbioloog Richard Dawkins (de belangrijkste popularisator van het idee dat alle eigenschappen van organismen de verspreiding van de genen dienen) in het NRC Handelsblad van 14 januari 2006 nog? Onder de kop “straf is wetenschappelijk achterhaald” claimde deze hoogleraar ‘Maatschappelijk besef van natuurwetenschap’ (Oxford) dat begrippen als verantwoordelijkheid en schuld wetenschappelijk gezien onhoudbaar zijn, hoewel ze “door millennia darwinistische evolutie in onze hersenen zijn gegrift.” Eigenlijk zouden we volgens hem het beste maar eens hard kunnen lachen om ons gestuntel met deze begrippen in de juridische praktijk.

Begin dit jaar (2010) ging de plantenbioloog Anthony Cashmore in de lezing waarmee hij het lidmaatschap aanvaardde van de Nationale Academie van Natuurwetenschappen van de Verenigde Staten uitgebreid in op wat hij “het geloof in vrije wil” noemt. Dit geloof komt volgens hem neer op het halsstarrig vasthouden aan een vorm van vitalisme (de opvatting dat de levende wereld niet geheel door fysische en chemische krachten beheerst wordt) die, volgens hem, dankzij het werk van Darwin, in de biologie al meer dan honderd jaar verlaten is. Ook hij pleit voor een zodanige hervorming van het recht dat begrippen als toerekeningsvatbaarheid daar geen rol meer in spelen.

Al deze wetenschappers gaan er, zonder daarvoor enig argument te geven, van uit dat onze praktijk van verantwoordelijkheid gebaseerd is op het idee dat mensen het vermogen hebben om zich aan de natuurwetten te onttrekken. Veel filosofen denken daar anders over. De bekende Amerikaanse filosoof Daniel Dennett bijvoorbeeld, legt in zijn *Freedom Evolves* (2003) (één van de weinige filosofische publicaties over vrije wil en evolutie) uit hoe de vrijheid die redelijkerwijs nodig is om mensen verantwoordelijk te houden voor wat ze doen, zich zou kunnen ontwikkelen in de strijd om het bestaan. Daarmee doelt hij echter niet op de vrijheid om je aan de natuurwetten te onttrekken, maar op de vrijheid om op basis van redenen te handelen. Volgens hem is het dus niet nodig dat mensen het vermogen hebben zich aan de natuurwetten te onttrekken om hen verantwoordelijk te houden voor wat ze te doen. We zijn volgens hem verantwoordelijk voor wat we doen als we in staat zijn ons een oordeel te vormen over wat goede redenen zouden zijn voor ons handelen (en wat niet) en ons handelen op basis van die redenen aan te passen.

In dit hoofdstuk ga ik eerst in op de vraag of organismen die in een biologische strijd om het bestaan verwickeld zijn het vermogen kunnen ontwikkelen op basis van redenen te handelen. Vervolgens behandel ik de vraag of het vermogen op basis van redenen te handelen voldoende is om deze organismen verantwoordelijk te houden voor wat ze doen. Ik betoog dat het antwoord op beide vragen ‘ja’ is.

1 Het vermogen op basis van redenen te handelen¹

1.1 Zelfen

Om überhaupt over een *belang* te kunnen spreken, biologisch of anderszins, moet er iets bestaan als een zelf dat een belang kan hebben.² De elementen van zo'n zelf moeten op een speciale manier een eenheid vormen. Die eenheid mag niet puur subjectief zijn zoals bij het sterrenbeeld Grote Beer dat geen andere eenheid heeft dan het patroon dat wij er in zien.

Anders dan de sterren van de Grote Beer vormen de onderdelen van een atoom (elektronen, neutronen en protonen) een objectieve eenheid. Dat komt doordat zij in chemische reacties als een geheel functioneren: bij verbranding worden zuurstofatomen aan koolstofatomen gekoppeld, bij andere reacties worden ze weer losgemaakt of verplaatsen ze zich van het ene molecuul naar het andere. De moleculen van een sneeuwbal vormen eveneens een objectieve eenheid. Ditmaal doordat wij ze als dusdanig behandelen: we stapelen sneeuwballen op, gooien ze naar voorbijgangers en duiken weg voor een sneeuwbal die ons dreigt te raken.

Hoewel een atoom of een sneeuwbal een objectieve eenheid vormt is hier geen sprake van een zelf. Daar is meer voor nodig. Maar wat?

Anders dan atomen en sneeuwballen dreigen organismen voortdurend uiteen te vallen. Zij kunnen in de omstandigheden waarin zij bestaan alleen bestaan doordat zij (en zolang zij) zichzelf actief onderhouden (met actief wordt hier bedoeld dat er energie voor het onderhoud nodig is): de meeste lichaamcellen worden voortdurend vervangen, celmembranen, het DNA en andere organellen worden voortdurend verbouwd en gerepareerd, weefsels worden aangemaakt, omgevormd, weer afgebroken, afvalstoffen worden verwijderd, voedingsstoffen opgenomen, vreemde cellen onschadelijk gemaakt etc. Het is een activiteit van jewelste. De delen van een organisme vormen dus niet alleen een eenheid doordat er (net als bij atomen en sneeuwballen) allerlei uitwendige processen zijn waarin zij als geheel functioneren, maar ook doordat zij tezamen het geheel waarvan zij deel uit maken onderhouden, een geheel dat alleen dankzij dit onderhoud kan bestaan. Dit maakt het zinvol om over een zelf te spreken.

¹ De evolutiegeschiedenis die ik in deze paragraaf presenteer is in grote lijnen ontleent aan Dennetts *Freedom Evolves*.

² Volgens Dennett krijgen begrippen als 'eenheid', 'zelf' en 'schadevermijding' betekenis door vanuit het perspectief van een ontwerper naar de levende natuur te kijken. Een eenheid is een door de ontwerper in elkaar gezet object, alles wat dat object aantast is een bedreiging etc. Deze manier van kijken is volgens Dennett gerechtvaardigd omdat het evolutieproces volgens hem ook echt een proces is waarin door vallen en opstaan objecten ontworpen worden. Ik vind deze benadering niet erg overtuigend. Zoals ik laat zien kunnen we deze begrippen zonder een beroep te doen op een ontwerpersperspectief inhoud geven op basis van de noodzaak tot zelfonderhoud.

Omdat organismen alleen kunnen bestaan doordat zij zichzelf actief onderhouden, heeft het zin om onderscheid te maken tussen enerzijds iets wat een organisme zelf doet en anderzijds dat wat het ondergaat of passief veroorzaakt. Een kwal ondergaat de beweging van de zee, een vis niet: de beweging van de vis wordt voor een groot deel bepaald door z'n eigen activiteit, die van een kwal niet. Een plant werpt passief een schaduw (de schaduw is een fysisch gevolg van de vorm van de plant, daar gebruikt de plant geen energie voor), maar neemt actief wat erop (met behulp van energieverbruikende cellulaire pompen). Een atoom of een sneeuwbal kunnen niet iets doen of iets ondergaan, omdat het voor hun bestaan niet nodig is dat ze hun eigen organisatie onderhouden.

In relatie tot een zelf krijgen begrippen als 'bedreiging' en 'schadevermijding' betekenis: iedere omgevingsverandering die tot gevolg heeft dat een zelf zich minder goed kan handhaven (minder goed dan zonder die verandering) vormt een bedreiging. Van schadevermijding is sprake als dankzij een onderliggende actief mechanisme de activiteit van een zelf systematisch varieert in reactie op de omgeving op een zodanige manier dat het zelf zich met deze variatie beter kan handhaven dan zonder.

Organismen worden niet alleen voortdurend in hun bestaan bedreigd, ze produceren bovendien doorgaans meer nakomelingen dan er kunnen bestaan en zich voortplanten. Deze laatste situatie noemde Darwin 'de strijd om het bestaan.' Als in de strijd om het bestaan de verschillende nakomelingen van één ouderpaar enigszins van elkaar verschillen, als die verschillen erfelijk zijn en sommige typen nakomelingen meer kans hebben om aan de volgende generatie bij te dragen dan andere typen, dan hebben we te maken met het verschijnsel dat Darwin 'evolutie door natuurlijke selectie' noemde: de samenstelling van de populatie van dergelijke organismen zal langzamerhand opschuiven in de richting van die typen die het meest bijdragen aan de volgende generatie. In zo'n situatie zullen de organismen eigenschappen ontwikkelen die hen helpen bij te dragen aan de volgende generaties. Niet omdat ze zich dat ten doel stellen of daar naar streven, maar als automatisch gevolg van de situatie waarin ze zich bevinden.

De strijd om het bestaan in combinatie met een variatiemechanisme verklaart hoe schadevermijdende organismen kunnen bestaan: mechanismen die tot gevolg hebben dat een organisme zich beter kan handhaven, hebben immers vaak ook tot gevolg dat de voortplantingskans van dat organisme verhoogd wordt en dus kunnen dergelijke mechanismen zich uitbreiden.

Merk op dat begrippen als 'zelf', 'bedreiging' en 'schadevermijding' geen evolutionair geladen begrippen zijn. Een beroep op evolutie is nodig om te verklaren dat er zelden, bedreigingen en schadevermijding bestaan, niet om te begrijpen wat iets tot een zelf, een bedreiging of schadevermijding maakt.

Merk tevens op dat deze notie van zelf niet vereist dat een zelf een besef van zichzelf heeft of zichzelf op een of andere manier herkent. Organismen vormen een zelf doordat ze als eenheid fungeren in het proces van zelfonderhoud. Zelfherkenning en zelfbesef maken een zelf dus niet tot een zelf. Andersom kan het feit dat organismen zelden zijn

wel verklaren dat zich in de loop van de evolutie allerlei vormen van zelfbesef ontwikkeld hebben: bewegende organismen kunnen zich immers beter handhaven als ze zichzelf niet achterna zitten of opeten.

1.2 Keuzemechanismen³

In een eenvoudig organisme bestaat schadevermijding uit vaste reacties op veranderingen in de omgeving: bepaalde omgevingsprikkelers hebben een bepaalde reactie tot gevolg. Dergelijke mechanismen noemen we ‘situatie-reactiemechanismen’. Als er voldoende voedsel in de omgeving is blijf je eten, als dat er niet is ga je zwemmen. Als er plotseling een schaduw valt maak je een sprong.

Er zijn allerlei situaties waarin het gunstig is om niet altijd op dezelfde manier op een omgevingsprikkel te reageren. Als je net gegeten hebt, kun je je bij het minste geringste gevaar beter uit de voeten maken; als je heel lang niet gegeten hebt, is je overlevingskans groter als je wat minder voorzichtig bent (als vossen jongen hebben, in juli en augustus, zie je ze veel vaker dan in de rest van het jaar—in die tijd weegt het extra risico dat ze lopen door zich overdag te laten zien op tegen het extra eten dat ze zo binnenhalen).

Situatie-reactiemechanismen kunnen dit soort ‘afwegingen’ maken door het beslissingsmechanisme van informatie over de toestand van het organisme te voorzien (als de maagspier niet gerekt is en er komt een prooi voorbij dan hap je, als ie uitgerekt is blijf je liggen). Ze kunnen echter niet profiteren van ervaring. Om van ervaring te profiteren moet je de gevolgen van een reactie evalueren en in staat zijn om een volgende keer op hetzelfde inputpatroon anders te reageren. Daartoe is het nodig dat de ervaring op een of andere manier binnen het mechanisme opgeslagen (gerepresenteerd) wordt.

In variërende omgevingen die bovendien niet goed voorspelbaar zijn, is het vaak voordelig om te kunnen leren van ervaring. In een omgeving waarin de hapbare dingen ten dele giftig en ten dele voedzaam zijn bijvoorbeeld, gaat de evolutionaire ontwikkeling van een mechanisme dat in staat is om te *leren* wat giftig is en wat niet, sneller dan de evolutionaire ontwikkeling van een erfelijk mechanisme dat in staat is dit onderscheid te maken. Bovendien heeft een leermechanisme het voordeel dat je ook in nieuwe omgevingen het onderscheid tussen giftig en voedzaam kunt verwerven. Een goed voorbeeld van een dergelijk mechanisme vinden we bij ratten. Een rat die kort nadat hij iets gegeten of gedronken heeft ziek wordt, zal wat hij at of dronk geen tweede

³ De hier gepresenteerde gedachtegang is volledig ontleent aan Dennetts *Freedom Evolves*. In *Freedom Evolves* tracht Dennett zijn in eerder werk geïntroduceerde onderscheid tussen Skinner-mechanismen en Popper-mechanismen te verhelderen met behulp van een aan Drescher (1991) ontleent onderscheid tussen situatie-reactiemechanismen en keuzemechanismen. Het leek mij duidelijker daar een driedeling van te maken: situatie-reactiemechanismen, conditioneringsmechanismen (ofwel Skinner-mechanismen) en keuzemechanismen (ofwel Popper-mechanismen).

keer naar binnen werken. Ratten moeten dus wel beschikken over een aangeleerde representatie van wat eetbaar is en wat niet.

Het type leermechanisme dat ratten de mogelijkheid geeft te leren wat in de omgeving als voedsel kan dienen, noemen we een ‘conditionerings-mechanisme’. Hierbij wordt een reactie op de omgeving aangepast aan de hand van een specifieke ervaring.

Zogenaamde ‘keuzemechanismen’ hebben nóg meer flexibiliteit. Dit is het type mechanisme dat we vinden in bijvoorbeeld een schaakcomputer. Keuzemechanismen berekenen de verwachte gevolgen van een aantal mogelijke reacties op een aantal waarden. Op basis van deze waardescore wordt vervolgens een reactie berekend en uitgevoerd. In deze berekening heeft elke waarde een variabele weegfactor die aangeeft hoe belangrijk de betreffende waarde op dat moment is. Na de reactie wordt het resultaat op de verschillende waarden gescoord en vergeleken met de verwachte score. Op grond daarvan wordt de berekening van de te verwachten waardescore aangepast.

Keuzemechanismen verschillen van conditionerings-mechanismen doordat de aanpassing van toekomstige reacties niet op basis van één specifiek gevolg plaats vindt (maar op basis van de score op een *aantal* waarden), maar vooral ook doordat de score op een aantal mogelijke reacties bepaald wordt *vóór* de reactie plaats vindt en bepalend is voor welke reactie er uitgevoerd wordt. Dat maakt dat we van ‘keuzen’ kunnen spreken.

De reactie van keuzemechanismen hangt af van waarden en variabele weegfactoren die dus, op een of andere manier, in dat mechanisme opgeslagen moeten zijn. Keuzemechanismen hebben derhalve een representatie van wat op dat moment belangrijk is en die representatie bepaalt uiteindelijk hoe het mechanisme op de buitenwereld reageert. Keuzemechanismen reageren, met andere woorden, op basis van interne waarden.

Bij organismen komen die interne waarden en de manier waarop ze gewogen worden tot stand in het evolutieproces. Daardoor zijn ze zodanig van aard dat ze tot reacties leiden die het organisme helpen om zich te handhaven en zich voort te planten (in de omgeving waarin de mechanismen ontstaan zijn). De waarden worden in eerste instantie bepaald door materiële en fysische factoren (zoals genen en omgevings-temperatuur) die ten dele erfelijk zijn. Ze worden door individuele ervaring aangepast. Ze representeren het voortplantingssucces in de zin dat het criteria zijn voor welke reacties de voortplantingskans vergroten.

1.3 Culturele overdracht en redenen

Ik ben dit artikel begonnen met de vraag, hoe het in het licht van de evolutietheorie mogelijk is dat organismen keuzen maken die ingaan tegen wat we het belang van hun genen zouden kunnen noemen.⁴ Een eerste stap is de ontwikkeling van keuzemechanis-

⁴ Om uit te leggen hoe het kan dat mensen tegen het belang van hun genen ingaan, vat Dennett culturele constructies op als parasieten (hij noemt ze ‘memes’) die het evolutionair gevormde mechanisme uitbuiten voor hun eigen verspreiding. Dit aan Dawkins’ *The Selfish Gene* ontleende idee (dat te weinig

men. In de vorige paragraaf zagen we keuzemechanismen een innerlijke representatie van wat belangrijk is met zich meedragen. De vraag is nu hoe het in het licht van de evolutietheorie mogelijk is dat deze innerlijke representatie van wat belangrijk is af kan wijken van wat de verspreiding van de eigen genen zou dienen. Een deel van het antwoord is gelegen in het ontstaan van een mechanisme voor culturele overdracht.

Een organisme dat kan profiteren van de informatie die anderen vergaard hebben, kan in veel omstandigheden adequater reageren op wat er gebeurt dan wanneer de representatie van informatie over de omgeving, over de gevolgen van bepaalde typen reacties etc. alleen op basis van geërfde factoren en eigen ervaring tot stand komt. Als een rat van anderen zou kunnen leren wat eetbaar is, hoeft ie dat niet zelf uit te zoeken, een risicovol en tijdrovend proces. Daardoor kan zich in het evolutieproces een mechanisme voor culturele overdracht ontwikkelen, waarbij de interne representaties voor culturele invloeden vatbaar zijn. Bij keuzemechanismen hoeft het dan niet alleen om informatie over de effecten van bepaalde reacties te gaan, het kan, zoals ik zal bespreken, minstens zo nuttig zijn waarden cultureel over te dragen.

Een beperkte mate van culturele overdracht is bij enkele soorten vissen, vogels en zoogdieren vastgesteld. Chimpansees beschikken over uitgebreide systemen van culturele overdracht die met de genetische overdracht interfereren en behalve door de mens door geen andere diersoort worden geëvenaard. Dankzij de ontwikkeling van taal heeft culturele overdracht bij de mens een ongekende vlucht genomen waardoor genetische overdracht, culturele overdracht en het leren uit ervaring tot een samenhangend netwerk verweven zijn geraakt. Het vermogen te communiceren d.m.v. abstracte, contextonafhankelijke symbolen maakt het mogelijk van de ervaring van heel veel mensen te profiteren: mensen in onze omgeving, mensen die te ver weg wonen om ze te ontmoeten, mensen uit het verleden.

In eerste instantie bepalen geëvolueerde mechanismen wat in welke omstandigheden geleerd wordt. We hebben bijvoorbeeld de neiging om oudere mensen meer te vertrouwen dan jongere, bekenden meer dan vreemden, en om meer te vertrouwen op dat wat je zelf waarneemt dan op dat wat mensen zeggen. Deze ingebouwde 'criteria' volstaan niet meer als de culturele overdracht belangrijker, complexer en gedetailleerder wordt. In zo'n situatie kan het voordelig zijn als ook de leermechanismen (de mechanismen die bepalen wat je in welke omstandigheden leert) zelf niet genetisch vastgelegd zijn, maar zich (ten dele) op basis van eigen ervaring en culturele overdracht ontwikkelen.

Als ik je vertel dat ik kleurenblind ben, neem je dat waarschijnlijk zonder meer aan (een vorm van leren op basis van culturele overdracht). Vervolgens geloof je me niet als ik zeg dat de aardbeien in mijn tuin ondanks de vele zon maar niet rood worden. Wat wij eerder geleerd hebben (bijvoorbeeld dat iemand kleurenblind is) bepaalt dus mede wat

uitgewerkt is om het een theorie te noemen, al spreken sympathisanten vaak over de 'meme-theorie'), heeft weinig aanhang onder biologen. Zoals ik laat zien kan hetzelfde punt ook zonder een beroep op memen gemaakt worden.

wij nu leren (van iemand die kleurenblind is leren we niets over de kleur van aardbeien). Dankzij deze flexibiliteit kunnen wij op uiterst genuanceerde wijze van de ervaring van anderen profiteren. Als ik beweer dat de aardbeien in mijn tuin niet rijp zijn, zou je mij kunnen vragen waarom ik dat denk. Baseer ik dat op de kleur? Of heb ik ze geproefd? In het eerste geval geloof je me niet zonder meer, in het tweede wel.

Met andere woorden, hoe wij op iets reageren hangt voor een groot deel af van een netwerk van cultureel overgedragen opvattingen over goede redenen. Als iemand beweert dat hij kleurenblind is, is dat in het algemeen een goede reden om aan te nemen dat hij kleurenblind is. Iemands kleurenblindheid is een goede reden hem niet te geloven als wat hij beweert op kleurwaarneming gebaseerd is.

Dit netwerk van redenen wordt in de loop van onze ontwikkeling gevormd en voortdurend aangepast. Onze opvoeding is erop gericht het netwerk dat in onze omgeving gangbaar is over te dragen en ons te leren overeenkomstig dit netwerk te handelen. Het resultaat is dat wij doorgaans automatisch in overeenstemming met dit netwerk reageren, maar de match is natuurlijk niet perfect. Als we een reactie overwegen dan gaan we na hoe reacties in dit netwerk zouden passen. Als ons gevraagd wordt naar het waarom van een bepaalde reactie dan interpreteren we wat we deden in termen van dit netwerk. Als we merken dat bepaalde automatische reacties niet passen in het netwerk zullen we proberen om de reactie aan te passen of onze opvattingen te wijzigen. Voorzover we ons handelen in overeenstemming proberen te brengen met dit netwerk handelen we op basis van redenen.

De mate waarin ons handelen bepaald wordt door een cultureel overgedragen netwerk van opvattingen over wat goede redenen zijn voor wat, vormt het begin van een antwoord op de vraag hoe het kan dat in de strijd om het bestaan verwikkelde organismen niet uitsluitend gericht zijn op de verspreiding van hun genen. Maar meer dan een begin is het niet. Het blijft immers nog de vraag hoe het kan dat dit netwerk als geheel gericht kan zijn op waarden als rechtvaardigheid, de belangen van andere mensen, het algemeen belang of het goede (wat dat ook moge zijn). Zou je evolutionair gezien niet verwachten dat het netwerk gericht is op het produceren van een reactie die de verspreiding van onze genen zo goed mogelijk dient? Deze vraag staat in de volgende paragraaf centraal.

1.4 De bereidheid tot coöperatie

In een beroemde passage in *The Selfish Gene* (1976) merkt Dawkins (de in de inleiding van dit hoofdstuk geïntroduceerde popularisator van het idee dat organismen slechts dienen ter verspreiding van hun genen) op dat wij mensen in de unieke situatie verkeren dat wij, als enige organismen op aarde, in staat zijn in opstand te komen tegen onze zelfzuchtige genen. Dit vermogen danken wij volgens hem aan de belangrijke rol die culturele overdracht in ons leven speelt. Vaak trekt men hieruit de conclusie dat Dawkins van mening is dat de bereidheid het goede te doen een kwestie van culturele beschaving is die ingaat tegen onze natuurlijke neiging om datgene te doen wat de verspreiding van onze genen ten goede komt. Een dergelijke positie roept, zoals ik

hierboven suggereerde, onmiddellijk de vraag op hoe het in het licht van de evolutietheorie überhaupt kan dat culturele overdracht tegen onze natuurlijke neigingen ingaat.

Het is echter de vraag of de neiging het goede te doen echt wel zo in strijd is met onze genen als Dawkins suggereert en of wij niet een natuurlijke neiging hebben om in het algemeen belang te handelen, anderen te helpen en/of datgene te doen wat goed geacht wordt. Ik zal dergelijke neigingen samen nemen onder de noemer 'coöperatieve neigingen'.

Onze dierlijke voorouders waren, lang voor de ontwikkeling van taal en de grote vlucht die culturele overdracht nam, sociaal levende dieren. Het valt dan ook te verwachten dat onze natuur ten dele aangepast is aan een sociaal leven, ondermeer door de aanwezigheid van coöperatieve neigingen. De apenbioloog Frans de Waal heeft in zijn *Good Natured* (1996) en een aantal latere publicaties vele voorbeelden aangereikt van helpend gedrag in het dierenrijk dat zouden kunnen wijzen op een erfelijke neiging tot coöperatie. De vraag rijst daarom hoe coöperatieve neigingen in de loop van de evolutie zouden kunnen ontstaan.

Een begin van een antwoord op deze vraag is het inzicht dat coöperatieve neigingen een oplossing kunnen vormen voor wat biologen 'het probleem van de samenwerking' noemen. Dit probleem wordt doorgaans geanalyseerd aan de hand van het, aan de Amerikaanse rechtspraak ontleende, *prisoner's dilemma*. De politie heeft twee inbrekers op heterdaad betrapt en in hechtenis genomen. De twee hebben bovendien een aantal bankrekeningen geplunderd met behulp van gestolen pinpassen en afgekeken pincodes. Het bewijs voor dit laatste misdrijf is echter wankel. De officier van justitie bedenkt daarom een list die de arrestanten er toe moet overhalen zelf het bewijsmateriaal te leveren. Zij doet elk van de arrestanten afzonderlijk een aanbod tot strafvermindering in ruil voor controleerbare informatie waaruit de betrokkenheid van de ander bij de pinpasplundering blijkt. Als één van beide verdachten de informatie levert, wordt deze niet vervolgd, terwijl de ander een jaar gevangenisstraf krijgt. Als beide verdachten de informatie leveren, draaien ze beiden een half jaar de bak in. Als geen van tweeën op het aanbod ingaat, krijgen ze beide 5 weken gevangenisstraf vanwege de poging tot inbraak.

Als de arrestanten het voorstel van de officier beoordelen op basis van hun streven om zo kort mogelijk de gevangenis in te gaan, zullen ze allebei tot de conclusie komen dat ze het aanbod maar beter kunnen accepteren. Beiden zullen immers als volgt redeneren: als ik het aanbod accepteer en de ander doet dat niet, dan krijg ik geen straf (terwijl ik 5 weken zou krijgen als ik niet accepteer); mocht de ander het aanbod ook accepteren dan krijg ik een half jaar (terwijl ik een jaar zou krijgen als ik niet accepteer); wat de ander ook doet, ik ben dus beter af als ik het aanbod accepteer. Op basis van deze redenering zullen beide arrestanten op het aanbod ingaan, met als uitkomst dat ze allebei een half jaar gevangenisstraf krijgen. Ze zouden echter beiden beter af zijn, als ze allebei hun mond houden, dan krijgen ze elk vijf maanden gevangenisstraf.

In de evolutionaire versie van dit dilemma gaat het om situaties waarin individuen die zich coöperatief gedragen minder bijdragen aan de volgende generatie dan individuen die zich niet coöperatief gedragen, terwijl de voortplantingskans van alle betrokkenen groter is als iedereen het coöperatieve gedrag vertoont, dan wanneer geen van de betrokkenen coöpereert.⁵ Een mus die een geschikte voedselbron vindt, verkondigt dit luidkeels. Daardoor kunnen ook andere mussen van het gevonden voedsel profiteren. Doordat elke mus zo enthousiast reageert op de ontdekking van een voedselbron, vindt de groep als geheel meer voedsel. Daardoor heeft elke mus meer te eten dan wanneer iedere mus het goede nieuws voor zich zou houden. Toch is het voor elke mus voordeliger om over de vondst te zwijgen. Dan kun je er immers een groter deel van de voedselbron zelf naar binnenwerken. In barre tijden hebben de mussen die hun bek houden daardoor een grotere overlevingskans dan de schreeuwers.

Voor keuzemechanismen heeft het *prisoner's dilemma* een paradoxaal karakter: een evaluatie van mogelijke reacties op basis van coöperatieve waarden dient het uiteindelijk belang (zo weinig mogelijk straf, resp. de beste voortplantingskans) beter dan een evaluatie met betrekking tot dat uiteindelijk belang (mits alle betrokken mechanismen de afweging op basis van coöperatieve waarden maken). Als geen van de arrestanten zich druk maakt om de straf die hen te wachten staat, maar de keus uitsluitend baseert op waarden als hun reputatie als bendelid, het houden van een belofte, het niet verlinken van je maat etc. komen ze er beiden met minder straf vanaf dan wanneer het vermijden van gevangenisstraf voor ieder een hogere prioriteit heeft.

Een betrouwbare coöperatieve instelling kan dus gezien worden als een oplossing voor het *prisoner's dilemma*. Het blijft echter de vraag hoe een dergelijke instelling in de evolutie kan ontstaan en blijven bestaan. In een gemeenschap van niet-coöpererende individuen zijn coöpererende individuen immers in het nadeel. Een genetische neiging om te coöpereren zal zich hier dus niet kunnen verbreiden. In een gemeenschap van coöperatieve individuen zijn individuen die niet coöpereren in het voordeel en de neiging tot non-coöperatie kan zich daar dus wel verbreiden.

Er zijn verschillende wiskundige modellen ontwikkeld die kunnen verklaren hoe en onder welke omstandigheden zich toch een genetische neiging tot coöperatie kan ontwikkelen. Deze modellen tonen dat een genetische neiging tot coöperatie zich kan verbreiden indien individuen met de neiging tot coöperatie geclusterd zijn. Hierdoor hebben individuen met coöperatieve neigingen vooral te maken hebben met individuen die eveneens coöperatieve neigingen hebben, terwijl niet-coöperatieve individuen vooral te maken hebben met individuen die evenmin coöperatieve neigingen hebben. Omdat de meeste organismen voornamelijk omgaan met degenen met wie ze ook de

⁵ Er is weinig bekend over de feitelijke rol van het *prisoner's dilemma* in de evolutie. Mijn mussenvoorbeeld dient slechts om het dilemma duidelijk te maken. Het stemt niet overeen met de realiteit. Mussen roepen er inderdaad anderen bij als ze een voedselbron ontdekken, maar ze bevinden zich waarschijnlijk niet in de situatie van het *prisoner's dilemma* omdat de aanwezigheid van anderen voor de vinder voordelig is—zo worden gevaarlijke katten sneller ontdekt.

genen uitwisselen—hun naaste burenen—is deze conditie niet helemaal onrealistisch. In modellen waarin coöperatieve organismen weigeren samen te werken met individuen waarvan ze op een of andere manier opgemerkt hebben dat deze niet coöpereren is de kans op de ontwikkeling van een genetische neiging tot coöperatie nog groter. Nog beter is het als coöperatieve individuen voor anderen als dusdanig herkenbaar zijn en ze elkaar actief opzoeken. Als niet-coöpererende individuen bovendien gestraft worden (door hun levenskansen actief te verlagen) staat niets de ontwikkeling van een neiging tot coöperatie in de weg, zoals Boyd en Richerson in 1992 aantoonde in een artikel met de veelzeggende titel “Punishment Allows the Evolution of Cooperation (or Anything Else) in Sizable Groups”.

1.5 De bereidheid datgene te doen wat goed geacht wordt⁶

In de vorige paragraaf zagen we dat in een strijd om het bestaan verwickelde keuzemechanismen, in situaties terecht kunnen komen waarin coöperatieve waarden het belang van de genen beter dienen dan een afweging op basis van de directe opbrengst voor de genen. In zo’n situatie kunnen zich coöperatieve waarden ontwikkelen als coöperatieve individuen elkaar als dusdanig herkennen en opzoeken.

Een coöperatieve waarde die het probleem van de samenwerking kan oplossen is de neiging datgene te doen wat als goed beschouwd wordt. Het mechanisme van herkenning dat de ontwikkeling van deze waarde mogelijk maakt, zou gebaseerd kunnen zijn op culturele overdracht: iemands reputatie geeft een indicatie van haar staat van dienst m.b.t. de bereidheid datgene te doen wat als goed beschouwd wordt (als die reputatie tenminste gebaseerd is op bewezen coöperatie).

In eerste instantie zal coöperatie op basis van reputatie eerder leiden tot de ontwikkeling van de bereidheid goed te *lijken* dan tot de bereidheid datgene te doen wat als goed beschouwd wordt. Als er veel situaties zijn waarin je het zonder reputatieschade kunt nalaten om het goede te doen (niemand merkt het, degenen die het merken kunnen het niet doorvertellen of worden niet geloofd, je hebt in de toekomst niet te maken met degenen die ten gevolge van het gebrek aan coöperatie een negatief beeld van je hebben etc.) is de vaardigheid om de indruk te wekken dat je goed bent (terwijl je in feite zoveel mogelijk profiteert) immers voordeliger dan de neiging om ongeacht de omstandigheden datgene te willen doen wat goed geacht wordt. Daardoor ontstaat er een situatie waarin zowel een verbeterde capaciteit om met zo min mogelijk coöperatie een reputatie te vestigen als een beter vermogen om dit toneelspel te ontmaskeren in het voordeel zijn. In zo’n situatie kan er een “wapenwedloop” ontstaan waarin beide capaciteiten verbeteren. Als gevolg daarvan wordt profiteren langzamerhand steeds minder lonend: je moet steeds meer investeren in het vermogen om te onderscheiden tussen situaties waarin profiteren reputatieschade oplevert en situaties waarin dit niet

⁶ De gedachtegang in deze paragraaf is geheel ontleend aan *Freedom Evolves*. Er is echter een belangrijk verschil: ‘goed zijn’ betekent in mijn opvatting ‘datgene doen wat goed geacht wordt (in de gemeenschap waarin je leeft)’ terwijl het bij Dennett betekent ‘datgene te doen wat (objectief gezien) goed is’.

het geval is, je moet de zelfbeheersing ontwikkelen om de neiging tot profiteren te bedwingen in situaties waarin dat reputatieschade oplevert en je moet steeds vaker zichtbaar en feitelijk coöpereren om je reputatie te vestigen en te behouden. Bovendien doen zich steeds minder situaties voor waarin het mogelijk is te profiteren zonder je reputatie te verliezen. Deze wapenwedloop kan daardoor uitmonden in een situatie waarin de neiging om te doen wat goed geacht wordt, voordeliger is dan de neiging tot profiteren. Dit is een situatie waarin, zoals Dennett het uitdrukt, goed zijn, de beste manier is om goed te lijken: je loopt dan niet voortdurend het gevaar ontmaskerd te worden en hoeft minder zelfbeheersing op te brengen. In zo'n situatie kan zich dus een genetische neiging ontwikkelen om kandidaat-handelingen te beoordelen op hun bijdrage aan dat wat als goed beschouwd wordt.

In paragraaf 1.3 besprak ik hoe onze reacties op onze omgeving ingebed zijn in een cultureel overgedragen netwerk van opvattingen over wat goede reacties zijn. In de daaropvolgende paragraaf suggereerde ik dat het niet noodzakelijk in strijd is met de evolutietheorie als dit netwerk gericht is op het realiseren van coöperatieve waarden. Ten slotte besprak ik (in deze paragraaf) hoe de neiging om kandidaat-reacties te beoordelen op datgene goed geacht wordt zich in de strijd om het bestaan zou kunnen ontwikkelen. Daarmee is aangetoond dat het bestaan van organismen die in staat en bereid zijn te handelen op basis van een netwerk van redenen dat ondermeer gericht is op het realiseren van coöperatieve waarden niet in strijd hoeft te zijn met de evolutie theorie. Rest de vraag is of dit vermogen (dat ik in het vervolg kortweg 'het vermogen op basis van redenen te handelen' zal noemen) voldoende voorwaarde is om mensen verantwoordelijk te houden voor wat ze doen.

2 Morele verantwoordelijkheid,

Vrijwel iedereen zal tijdens het lezen van het bovenstaande af en toe scepsis hebben gevoeld met betrekking tot de invulling van noties als schadevermijding, afweging en keuze. Automatische reacties die maken dat de levenskans groter is dan wanneer je die reactie niet vertoont, dat is toch geen schadevermijding? Automatische afweging van mogelijke handelingen, dat kun je toch geen keuze noemen? Als je de pech hebt dat jouw ingeboren waarden wat minder coöperatief zijn dan gemiddeld, dan kun je daar toch niet verantwoordelijk voor gehouden worden? De vraag die hier rijst is de vraag of de invulling van het begrip vrije wil als het vermogen op basis van redenen te handelen een acceptabele invulling is, of vrije wil in deze zin een voldoende voorwaarde is om mensen verantwoordelijk te houden voor wat ze doen. In de volgende paragrafen zal ik verdedigen dat dit het geval is.

2.1 De vrije wil en de natuurwetten

In de inleiding van dit boek wordt de vraag gesteld of het redelijk is om mensen wel en de maan en dieren niet verantwoordelijk te houden voor wat ze doen, zonder aan te nemen dat het menselijk handelen zich aan de natuurwetten kan onttrekken. Mijn antwoord is: ja dat is redelijk. Het vermogen op basis van redenen te handelen

onderscheidt de mens van de maan en de dieren, dit vermogen kan bestaan zonder het vermogen ons aan de natuurwetten te onttrekken en is bovendien voldoende grond om mensen verantwoordelijk te houden voor wat ze doen.

Het gedrag van een mens hangt, anders dan de beweging van de maan om de aarde, niet alleen af van de eigenschappen van de materie maar ook van de manier waarop die materie georganiseerd is. Het gedrag van mensen wordt anders dan dat van de maan (voor een deel) bepaald door een door de mens zelf onderhouden innerlijke organisatie. Dat heeft de mens met andere organismen gemeen. Die organisatie komt tot stand door evolutie, door individuele leerprocessen en door culturele overdracht.

De innerlijke organisatie van de mens is anders dan die van de meeste organismen (maar net als die van een aantal andere mensapen) zodanig dat ons gedrag in hoge mate beïnvloed wordt door de cultuur. Dankzij de taal heeft de mens als enige soort de culturele overdracht zo hoog ontwikkeld dat wij in staat zijn om opvattingen over wat het geval is, wat wenselijk is en hoe te handelen op basis van redenen te beoordelen en daar naar te handelen.

In paragraaf 1 hierboven heb ik betoogd dat de ontwikkeling van dit vermogen niet in strijd is met de evolutietheorie. De vraag is nu of dit vermogen een goede basis is om mensen verantwoordelijk te houden voor wat ze doen.

2.2 Morele verantwoordelijkheid

Laten we om te beginnen het begrip morele verantwoordelijkheid eens onder de loep nemen. Vaak wordt aangenomen dat dit begrip vooral slaat op het aanrekenen van fouten en verdiensten (met nadruk op de eerste). Je houdt iemand verantwoordelijk voor wat ze doet als je vindt dat ze verwijten of straf *verdient* als ze het fout doet. In een poging uit te leggen waarom het redelijk is om mensen verantwoordelijk te houden voor wat ze doen, wordt daarom vaak verwezen naar een uniek vermogen dat vrijwel alle mensen hebben en dieren niet. ‘Mensen zijn verantwoordelijk voor wat ze doen omdat ze anders dan dieren de consequenties van hun handelen kunnen overzien en er voor kunnen kiezen om anders te handelen’ is een voorbeeld van zo’n poging.

‘Verantwoordelijk zijn’ kan ook een heel andere betekenis hebben. Een secretaresse die zegt dat ze verantwoordelijk is voor de fouten van haar baas, bedoelt natuurlijk niet dat ze straf verdient voor die fouten. Zij bedoelt dat het haar *taak* of *functie* is om de fouten die haar baas maakt op te lossen. Wie wil beargumenteren dat het redelijk is om secretaresses verantwoordelijkheid te houden voor de fouten van hun baas zoekt niet naar een unieke secretaresse-eigenschap maar wijst op het nut van deze taakverdeling binnen de organisatie waar beide werken. Deze notie van verantwoordelijkheid noemen we ‘rolverantwoordelijkheid’.

Mijns inziens is de legitimiteit van onze praktijk om mensen moreel verantwoordelijk te houden voor wat ze doen niet gebaseerd op individuele verdienste maar op de functie van deze praktijk in onze samenleving. Morele verantwoordelijkheid heeft daardoor veel meer weg van rolverantwoordelijkheid dan vaak wordt aangenomen.

Zoals bij de bespreking van het *prisoner's dilemma* aan de orde kwam, berust samenwerking er vaak op dat er ten eerste een zekere overeenstemming bestaat over wat in een bepaalde situatie verwacht wordt en ten tweede dat mensen in die situatie ook inderdaad doen wat van ze verwacht wordt. Er zijn in onze samenleving verschillende praktijken om dit te bereiken. Autoriteiten, traditie en stemprocedures bepalen wat er in bepaalde situaties verwacht wordt; conditionering, dwang en omgevingsmanipulatie beogen te bereiken dat mensen doen wat er van ze verwacht wordt. Onze praktijk van morele verantwoordelijkheid fungeert in aanvulling op genoemde praktijken en ten dele als alternatief daarvoor als praktijk waarin wij onze opvattingen over wat er verwacht mag worden en ons handelen op elkaar afstemmen.

Een tijdje geleden liet ik mijn fietspomp achter in een hoekje van een café omdat ik er niet mee wilde sjouwen tijdens een wandeling. De caféhouder had mij beloofd het ding in de gaten te houden. Toen ik aan het eind van de dag terugkwam was mijn pomp weg. Ik sprak de caféhouder hierop aan. Deze bleek de pomp uitgeleend te hebben aan een vrouw met een lekke band en deze had de pomp niet teruggebracht. Hij beloofde de 'dief' de volgende keer dat zij het café zou bezoeken ter verantwoording te roepen. Een week later kwam een vriendelijke vrouw mij de pomp terugbrengen. Zij had aangenomen dat de pomp door de eigenaar vergeten was en hem meegenomen om onderweg naar huis haar lekke band nog een keer te kunnen oppompen.

In dit voorbeeld zien we hoe we onze verwachtingen en ons handelen op elkaar afstemmen doordat we ons verantwoordelijk voelen voor wat we doen en mensen die niet doen wat we van ze verwachten ter verantwoording roepen. Ik verwacht dat de caféhouder erop let dat mijn pomp niet meegenomen wordt. Omdat ik denk dat aan die verwachting niet voldaan is, verwijt ik de caféhouder achteloosheid. Daardoor ontdekt hij dat niet voldaan is aan zijn verwachting dat iemand aan wie hij een pomp uitleent deze na gebruik zal terugbrengen, terwijl ik ontdek dat hij wel op de pomp gelet heeft en tot de conclusie komt dat hem niets te verwijten valt. Omdat de caféhouder zich verantwoordelijk voelt voor het verdwijnen van mijn pomp, maakt hij op zijn beurt de vrouw die de pomp meenam verwijten. Deze leert daardoor dat niet alles wat in een hoekje van een café staat vergeten is. Ze ziet in dat ze verkeerd gehandeld heeft en omdat ze zich daar verantwoordelijk voor voelt brengt ze de pomp terug. Bovendien neemt ze zich voor om, als ze weer eens denkt dat er geen bezwaar tegen is om iets wat ze mocht gebruiken ook mee te nemen, dit na te vragen. Door haar uitleg en handelwijze wordt het mij en de caféhouder duidelijk dat de vrouw geen diefstal te verwijten valt, maar slechts een verkeerde opvatting over wat je mag doen met artikelen waarvoor je toestemming hebt om ze te gebruiken. Omdat zij deze fout erkent en we alle drie de indruk hebben dat ze de volgende keer anders zal handelen is de kwestie nu de wereld uit.

De afstemming van verwachtingen en handelen gaat niet altijd zo harmonieus als hierboven. Iemand anders had wellicht gevonden dat de caféhouder de hem toevertrouwde pomp niet had mogen uitlenen en hem daarom om een vergoeding gevraagd. Als dan zou blijken dat de caféhouder een andere mening toegedaan is hebben we te maken met een normatief conflict. We kunnen dan proberen het conflict

op te lossen door redenen te geven voor de respectievelijke verwachtingen. Als het niet lukt om tot overeenstemming te komen kunnen we het erbij laten zitten, of gebruik maken van beschikbare andere praktijken (emotionele beïnvloeding, vechten, de rechter er bij halen etc.).

Een tweede functie van onze praktijk van morele verantwoordelijkheid is gelegen in de psychologische ondersteuning die deze praktijk aan individuen biedt: het deel uitmaken van een praktijk van verantwoordelijkheid helpt je om in te zien wat in een bepaalde situatie verwacht mag worden en om daar vervolgens ook naar te handelen. Doordat je je verantwoordelijk voelt voor je gedrag houdt je beter in het oog wat redelijkerwijs van je verwacht mag worden en doordat je weet waarom het gedrag dat verwacht wordt redelijk is, is het gemakkelijker te doen wat verwacht wordt.

De gerechtvaardigheid van de praktijk van morele verantwoordelijkheid is gelegen in deze maatschappelijke functies en niet in een of andere individuele, uniek-menselijke capaciteit, zoals de vrijheid je aan de natuurwetten te onttrekken of het vermogen op basis van redenen te handelen. Dit betekent natuurlijk niet dat dergelijke persoonlijke capaciteiten niet van belang zijn. Zoals gezegd, het bestaan van secretaresses wordt niet gelegitimeerd door een of andere unieke secretaressecapaciteit, maar door het nut van een bepaalde taakverdeling binnen een organisatie. Om de functie van secretaresse naar behoren te vervullen zijn echter wel bepaalde capaciteiten vereist en de organisatie zal proberen er voor te zorgen dat hun secretaresses die capaciteiten hebben.

Hetzelfde geldt voor de praktijk van verantwoordelijkheid. Die praktijk wordt niet gelegitimeerd door het vermogen op basis van redenen te handelen, maar om in deze praktijk te functioneren, is het wel nodig dat je in staat bent om in een groot aantal situaties in te zien wat redelijkerwijs van je verwacht mag worden en daar naar te handelen. Aangezien regulatie door verantwoordelijkheid de basis is voor heel veel andere praktijken in onze samenleving is het terecht om mensen die niet in voldoende mate in staat zijn op basis van redenen te handelen geheel of ten dele van onze maatschappelijke praktijken uit te sluiten. Aan de andere kant proberen we onze praktijken natuurlijk zo in te richten dat zoveel mogelijk mensen er aan kunnen deelnemen.

2.3 Verantwoordelijkheid en straf

De verhouding tussen de strafpraktijk en de verantwoordelijkheidspraktijk is complex. De functies van beide praktijken overlappen waar het er om gaat te bereiken dat mensen doen wat van hen verwacht wordt. Regulatie door verantwoordelijk houden is wat dit betreft een mensvriendelijk alternatief voor regulatie door straf.

Een jaar of twintig geleden ging ik voor het eerst met de auto naar Zweden. Telkens als we om een auto te passeren harder reden dan toegestaan, werd er om ons heen driftig met de lichten geknipperd. ‘Wat een heetgebakerde lui,’ dachten wij, ‘zouden alle Zweden zich continu ergeren aan mensen die de verkeersregels overtreden?’ Toen we dit aan een paar Zweden vroegen, reageerden ze hoogst verbaasd. Ze zagen het knippen niet als uiting van ergernis, maar als een vriendelijke manier om de

overtreder op z'n overtreding te attenderen. Het was niet in hen opgekomen dat iemand de snelheidslimiet met opzet zou overtreden. Nu was het onze beurt om verbaasd te zijn. 'Wat gebeurt er als de politie je betrapt? Wordt je dan niet aangehouden? Krijg je dan geen boete?' 'Een boete? Nee hoor. Aangehouden? Waarschijnlijk wel. Ze zullen je wel attenderen op de snelheidslimiet.' 'En als we ons daar vervolgens niets van aantrekken, wat gebeurt er dan?' 'Tja, geen idee eigenlijk. Een enkele buitenlander die zich niet aan de regels houdt, is geen gevaar op de weg. Waarschijnlijk halen ze minachtend hun schouders op. Misschien ontnemen ze je je rijbewijs, of je auto, of misschien zetten ze je het land uit. Wat jammer dat je niet beseft hoe goed wij het hier geregeld hebben.' Wat heerlijk om in zo'n land te leven.

Het probleem is natuurlijk dat de praktijk van verantwoordelijkheid vaak niet volstaat ter regulatie van ons gedrag. In Nederland zou het zonder boetes voor overtreding van de maximumsnelheid snel een puinzooi worden op de weg. Vaak werkt een verantwoordelijkheidspraktijk alleen maar als deze aangevuld wordt met een strafpraktijk.

Het zou te ver voeren hier in te gaan op de maatschappelijke functie en legitimiteit van stafpraktijken. Mijn punt is slechts dat de notie van verantwoordelijkheid niet, zoals vaak aangenomen wordt, in eerste instantie dient om straf te legitimeren (zo van: je bent in staat om op basis van redenen te handelen dus ben je verantwoordelijk voor wat je doen en dus verdien je straf als je het fout doet). Integendeel, de legitimiteit van de praktijk van verantwoordelijkheid berust ten dele op haar mensvriendelijke karakter als alternatief voor de strafpraktijk. In onze maatschappij is straf iets wat toegepast wordt als mensen niet in staat blijken op basis van redenen te handelen.

Voor beide praktijken geldt dat het onredelijk is ze toe te passen in situaties waarin de betreffende praktijk zijn maatschappelijke functies niet kan vervullen. Vaak heeft een conditie die tot gevolg heeft dat de verantwoordelijkheidspraktijk zijn functies niet kan vervullen, tevens tot gevolg dat de strafpraktijk zijn functies niet kan vervullen. Hieruit mag echter niet geconcludeerd worden dat de legitimiteit van het straffen gelegen is in de mate waarin mensen verantwoordelijk gehouden kunnen worden voor wat ze doen. Omdat de verantwoordelijkheidspraktijk een dove niet helpt om op noodkreten te reageren, is het onredelijk om het een dove kwalijk nemen dat zij niet op dergelijke kreten reageert. Kreten zijn simpelweg niet het soort zaken waar een dove op kan reageren. Als de enige maatschappelijke functie van straf gelegen zou zijn in het bevorderen dat mensen zich aan de regels houden, heeft het onvermogen van doven om op geluid te reageren tevens tot gevolg dat de strafpraktijk hen niet kan helpen om op noodkreten te reageren. Daarom is het niet terecht om een dove die niet op een noodkreet reageert te straffen voor het overtreden van de wettelijke plicht tot hulpverlening aan mensen in nood die in een aantal Europese landen van toepassing is. De onredelijkheid om doven te straffen voor het niet reageren op een noodkreet is echter niet gelegen in de onredelijkheid om het doven kwalijk te nemen dat zij noodkreten niet horen, maar in het feit dat de strafpraktijk wat dit betreft haar functies niet kan vervullen.

2.4 Vrije wil als het vermogen op basis van redenen te handelen

Ik heb hierboven betoogd dat de legitimiteit van onze praktijk van verantwoordelijkheid gelegen is in haar maatschappelijke functie in het op elkaar afstemmen van verwachtingen en handelingen. Het zal duidelijk zijn dat onze verantwoordelijkheidspraktijk deze functie ook kan vervullen als mensen niet beschikken over het vermogen zich te onttrekken aan de natuurwetten en ook als de uitkomsten van het afstemmingsproces volledig gedetermineerd zijn. Indeterminisme en het vermogen de natuurwetten te doorbreken zijn dus geen voorwaarden om mensen verantwoordelijk te houden voor wat ze doen. Het vermogen in een groot aantal situaties op basis van redenen te handelen is daarentegen zoals we gezien hebben een noodzakelijke en voldoende voorwaarde voor het goed functioneren van de praktijk van verantwoordelijkheid. De vrije wil als voorwaarde voor verantwoordelijkheid bestaat dus in het in de loop van de biologische en culturele evolutie ontwikkelde vermogen op basis van redenen te handelen.

3 Conclusie

Biologen zoals Dawkins en Cashmore die de praktijk van verantwoordelijkheid verwerpen omdat deze zou berusten op de illusie dat wij ons kunnen onttrekken aan de natuurwetten hebben geen gelijk. Die praktijk veronderstelt slechts het vermogen op basis van redenen te handelen en dit vermogen kan, zoals we gezien hebben, bestaan in een populatie van organismen die verwickeld zijn in een biologische strijd om het bestaan. In deze strijd kunnen zich keuzemechanismen ontwikkelen die hun handelen afwegen op basis van waarden die mede bepaald worden door het cultureel discours en die tot op zekere hoogte bereid en in staat zijn in een grote verscheidenheid aan omstandigheden op basis van redenen te beoordelen wat van hen verwacht mag worden en daar naar te handelen. Dit vermogen maakt de ontwikkeling van een praktijk van verantwoordelijkheid mogelijk, maar legitimeert deze niet. De legitimatie van de praktijk is gelegen in haar maatschappelijke functie in de afstemming van verwachtingen en handelen.

Aanbevolen ter verdere studie

Dennett, Daniel C. (2003). *Freedom Evolves*. New York: Viking. Nederlandse vertaling door Hans Bosman: *De evolutie van de vrije wil*, Amsterdam: Contact, 2004.

Dawkins, Richard (1976). *The Selfish Gene*. Oxford: Oxford University Press. Nederlandse vertaling door Henny Scheepmaker en Peter van Huizen: *De zelfzuchtige genen: over evolutie, eigenbelang en altruïsme*, Amsterdam: Olympus, 2006. Een ongenueanceerde, onderhoudende uiteenzetting van het idee dat organismen begrepen moeten worden als instrumenten waarmee de genen hun verspreiding bevorderen.

Waal, Frans de (1996). *Good Natured: The Origins of Right and Wrong in Humans and other Animals*. Cambridge: Harvard University Press. Nederlandse vertaling door

Fieke Lakmaker: *Van nature goed*, Amsterdam, Contact, 1996. Door middel van talloze voorbeelden maakt De Waal duidelijk dat dieren lang niet zo zelfzuchtig zijn als vaak wordt aangenomen.

Waal, Frans de et al. (2006). *Primates and Philosophers: How Morality Evolved*. Princeton University Press. Nederlandse vertaling door Nico Groen: *De aap en de filosoof: hoe de moraal is ontstaan*, Amsterdam: Contact, 2007. Na een inleidend opstel waarin De Waal zijn opvattingen over de menselijke natuur samenvat, volgt een bespreking van die opvattingen door een aantal vooraanstaande filosofen en een antwoord van De Waal. Mijns inziens de beste inleiding tot de filosofische opvattingen van De Waal.

De beste uitwerking van het idee dat de legitimatie van onze praktijk van verantwoordelijkheid gelegen is in onze gevoeligheid voor redenen is te vinden in het werk van Susan Wolf en Jay Wallace:

Wolf, Susan (1990). *Freedom Within Reason*. Oxford: Oxford University Press.

Wallace, R. Jay (1994). *Responsibility and the Moral Sentiments*. Cambridge: Harvard University Press.

De uitwerking die John Fischer en Mark Ravizza gaven aan het idee van gevoeligheid voor redenen heeft meer aandacht gekregen dat die van Wolf en Wallace, maar gaat een heel andere richting uit:

Fischer, John Martin & Mark Ravizza (1999). *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge: Cambridge University Press.

Overige bronnen

Boyd, Robert & Peter J. Richerson (1992). "Punishment allows the evolution of cooperation (or anything else) in sizable groups" *Ethology and Sociobiology* 13(3): 171-195.

Cashmore, Anthony R. (2010). "The Lucretian swerve: The biological basis of human behavior and the criminal justice system" *Proceedings of the National Academy of Sciences* 107(10): 4499-4504.

Drescher, Garry (1991). *Made-Up Minds: A Constructivist Approach to Artificial Intelligence*. Cambridge, MA: MIT Press.