

Hoofdstuk 13 Einsteins gemoedsrust

Drie perspectieven op het probleem van de vrije wil

Maureen Sie

Voorlaatste ongecorrigeerde versie, citeren alleen uit uiteindelijke versie.

Sie@fwb.eur.nl

Verschenen in: *Hoezo Vrije Wil, Perspectieven op een heikele kwestie*,
Lemniscaat 2011, pp pp 252-274

[k1]13

Einsteins gemoedsrust

Drie perspectieven op het probleem van de vrije wilⁱ

Maureen Sie

[t1] 'Het brein van pedoseksuelen wijkt af,' 'Stress vermindert zelfcontrole dramatisch,' 'Mannelijke omhelzing produceert hechting.' Elke dag hebben de Nederlandse media (de betere krant inclusief) wel een kleine traktatie voor diegene die zichzelf een kleine of flinke zonde moet vergeven. Val je op de verkeerde man? De oorzaak: zijn stevige omhelzingen. Lustgevoelens ten aanzien van kinderen? Het is je brein! Toch weer een peukie opgestoken? Te veel stress op je bordje! De vrije wil op een of andere wijze tot illusie verklaren, dat is marketingtechnisch een beproefd recept geworden. Kijk maar naar de titel van Daniel Wegners boek *The Illusion of Conscious Will* (Wegner 2002, zie de bijdrage van Slors in deze bundel) en het recent in Nederland verschenen *De vrije wil bestaat niet* van Victor Lamme (Lamme 2010, zie de bijdrage van Keizer in deze bundel).

Er zijn echter door de tijd heen altijd filosofen geweest die waarschuwen tegen zo'n boude claim. De Amerikaanse filosoof Daniel Dennett bijvoorbeeld, vindt Wegners boek briljant, hoewel hij opmerkt dat de titel misplaatst is (Dennett 2003). De vrije wil als zodanig een illusie noemen suggereert dat we iets missen, wat volgens hem niet het geval is. Een vergelijkbaar één-tweetje vinden we ook een paar eeuwen terug. Zo stelt Schopenhauer dat Spinoza helemaal gelijk heeft wanneer hij constateert dat '[...] als een in de lucht geworpen steen bewustzijn zou hebben, deze zou denken uit eigen wil te vliegen', maar hij bijt hem toe: 'Ik zou daaraan willen toevoegen dat de steen gelijk heeft!'

(Schopenhauer 1859, boek 2, sec. 24; voor Schopenhauers opvatting over de vrije wil, zie de bijdrage van Groot in deze bundel.)

Wegner sluit zijn boek af met een citaat van Einstein dat sterk aan Spinoza doet denken:

[t2]Als de maan in haar ronde om de aarde zelfbewustzijn zou bezitten, zou zij er volledig van overtuigd zijn dat zij haar weg uit eigen wil aflegt. [...] Op eenzelfde wijze zou een intelligenter wezen dat de mens en wat hij doet met een dieper inzicht zou aanschouwen, glimlachen om zijn illusie dat hij uit vrije wil handelt. (Wegner 2002, p. 342; vertaald uit het Engels, MSⁱⁱ)

[t1]Deze gedachte, aldus Einstein in zijn *Credo*, begeleidt hem in alle situaties van zijn leven en verzoent hem met de handelingen van anderen, zelfs als deze pijnlijk voor hem zijn. Het stelt hem in staat om zichzelf en anderen wat minder serieus te nemen als handelende en beslissende actoren én om wat minder zijn geduld te verliezen. Interessant is dat hij meent deze les aan Schopenhauer te ontleen, die hem ervan overtuigde dat de mens kan doen wat hij wil, maar niet kan willen wat hij wil.

Het metafysisch 'heen-en-weren' over de vrije wil kent noch tijd, noch duidelijke afbakening binnen één positie. Schopenhauer kan geciteerd worden als verdediger van de positie dat het onzinnig is om de vrije wil een illusie te noemen, maar hij wordt hier door Einstein aangehaald als iemand die hem heeft duidelijk gemaakt dat de vrije wil een illusie is.ⁱⁱⁱ Dat is natuurlijk verwarrend, zeker als je op zoek bent naar wat houvast omdat je je realiseert dat een stellingname over een

belangrijk onderwerp als de vrije wil implicaties heeft voor onze alledaagse morele praktijk. Immers, de vraag rijst of we als de vrije wil niet bestaat, mensen nog wel moreel verantwoordelijk kunnen houden voor hun daden.

Dit hoofdstuk behandelt drie verschillende perspectieven op het idee dat de vrije wil een illusie is. Ten eerste het perspectief dat het onzinnig is te denken dat de vrije wil een illusie is, omdat het teruggrijpt op een onderscheid dat wij dagelijks zonder problemen maken. Ten tweede het perspectief dat de vrije wil een illusie is, omdat niets in dit ondermaanse ontsnapt aan de gesloten keten van oorzaken die elkaar wetmatig opvolgen. Ten derde het perspectief dat de vrije wil bestaat en er wel degelijk gebeurtenissen zijn die zich onttrekken aan de wetmatigheid in dit ondermaanse, namelijk 'zelfvormende handelingen.' Ik zal laten zien dat er voor elk van deze perspectieven wat te zeggen valt, zonder daarbij gedetailleerd in te gaan op de verenigbaarheid of onverenigbaarheid van vrijheid en determinisme (zie daarvoor de bijdrage van Cuypers in deze bundel). Het is mij in dit hoofdstuk vooral te doen om de vraag waarom men zich wel of niet iets gelegen zou laten liggen aan dergelijke filosofische uitwerkingen, waarom men filosofisch geïnteresseerd zou kunnen raken in de haalbaarheid en uitwerking van één van deze perspectieven.

[k2]1 Het is onzinnig te denken dat de vrije wil een illusie is

[t1]Einstein ontleent troost en gemoedrust aan de gedachte dat de vrije wil niet bestaat, omdat hij in het dagelijkse leven gebukt gaat onder ergernis, wrok en verontwaardiging over het handelen van anderen (en

zichzelf). Zulke reacties geven uitdrukking aan onze verhouding tot de ander als een moreel verantwoordelijke actor. Bijvoorbeeld: als iemand ons in het verkeer op lompe wijze afsnijdt, reageren we zonder daar eerst over te hoeven nadenken met kwaadheid, ergernis en verontwaardiging: we ballen onze vuisten, vloeken binnensmonds of lopen rood aan. Als we vervolgens zien dan de persoon die ons zo lomp afsneed een epileptische aanval heeft, dan zullen die gevoelens als vanzelf afnemen. We realiseren ons dat hier geen sprake is van verwijtbaar gedrag door een daarvoor moreel verantwoordelijke actor, maar van een ongeluk veroorzaakt door een nare ziekte. We reageren met zogeheten morele emoties primair op het gedrag van andere volwassenen, en niet (althans niet op dezelfde wijze) op iemand die te lijden heeft onder een epileptische aanval, of op onze computers, huisdieren of heel jonge kinderen (zie ook de bijdrage van Anderson in deze bundel).

Peter F. Strawson wees er in de tweede helft van de vorige eeuw op dat deze natuurlijke en onmiddellijke emotionele reacties op elkaar een grote rol zouden moeten spelen in discussies over morele verantwoordelijkheid. Hiermee veranderde hij de tot dan toe gevoerde discussie over dat onderwerp drastisch (Strawson 1962). Deze werd voorheen gedomineerd door de volgende vragen. Als alles in dit ondermaanse uitputtend bepaald is door voorafgaande gebeurtenissen (de these van het determinisme), kunnen we dan nog wel spreken over vrije wil? En als dat niet het geval is, kunnen we elkaar dan nog wel moreel verantwoordelijk houden?

Strawson legt deze, en daaraan gerelateerde, vragen terzijde. Hij beweert zelfs geen idee te hebben wat de these van het determinisme zou kunnen betekenen. Volgens hem verliezen we het zicht op morele verantwoordelijkheid

zodra we afstand doen van de onmiddellijke en natuurlijke houdingen die wij tegenover elkaar innemen en de morele sentimenten die hiermee gepaard gaan. Met deze morele sentimenten reageren we op de door ons ervaren goede of slechte wil van de ander, op onze indruk of iemand ons goed of kwaad gezind is. Alleen als we zwaarwegende redenen hebben om een ander vanuit een zogeheten objectieve houding tegemoet te treden -- bijvoorbeeld omdat deze geestelijk uit balans of ziek is -- kunnen we onze interpersoonlijke verhouding tijdelijk opschorten en vervangen voor een houding waarin bijvoorbeeld de zorg voor, behandeling of manipulatie van de ander centraal staat. En dat is geen eenvoudige opgave, zoals iedereen weet die kinderen heeft opgevoed of veel te maken heeft met mensen die geestelijk niet geheel in balans zijn. Het komt niet vanzelf en spontaan om met geduld en begrip te reageren op een klein kind (laten we hem Niklas noemen) dat krijsend op de keukenvloer ligt, omdat je hem een rode in plaats van een blauwe beker gegeven hebt. Integendeel, dit kost inspanning, concentratie én energie. Ik kom hier verderop nog op terug.

Een groot aantal filosofen heeft deze door Strawson geïnitieerde beweging weg van de these van het determinisme omarmd en verder uitgewerkt. Zoals bijvoorbeeld de Amerikaanse filosofen Susan Wolf en R. Jay Wallace (Wolf 1990; Wallace 1994; zie ook de bijdrage van Meynen in deze bundel). Zij argumenteren geheel in de lijn van Schopenhauer en Dennett dat het niet zo zinnig is om aan het bestaan van de vrije wil te twijfelen en het af te doen als een illusie. Volgens hen is het overduidelijk dat er onderlinge verschillen bestaan in de mate waarin wij in staat zijn om ons moreel adequaat op te stellen en te gedragen. Die verschillen schuilen echter niet in een of ander mysterieus vermogen dat ons

in staat stelt om door de wetten van het universum te breken. Nee, die onderlinge verschillen tonen zich in onze redelijkheid, dat wil zeggen in ons vermogen te reflecteren op wat we behoren te doen en wat niet, en ons vermogen in overeenstemming te handelen met de uitkomsten van die reflectie.^{iv}

Het is op grond van die onderlinge verschillen dat we elkaar in meer of mindere mate moreel verantwoordelijk houden. En terecht. Hoe kunnen we immers iemand verwijten maken die niet (of in mindere mate) het vermogen heeft het goede te doen om de juiste redenen? Ik kom hier nog op terug. Laten we eerst even kijken naar de andere denker die de centrale rol heeft ondermijnd die de these van het determinisme lange tijd gespeeld heeft in de Angelsaksische debatten over morele verantwoordelijkheid: de eigenzinnige filosoof Harry G. Frankfurt.

[k3]1.1 De hiërarchische visie op de wil

[t1]Frankfurt wijst ons, eveneens in de tweede helft van de vorige eeuw, op het feit dat het niet zozeer de afwezigheid van dwang, drang of obstakels is die maakt dat onze zogenaamde vrije wil van belang is voor onze morele verantwoordelijkheid (zie de bijdrage van Cuypers in deze bundel). Naar zijn oordeel is bepalend dat in sommige gevallen deze dwang, drang of obstakels de *exclusieve* verklaring vormt voor ons gedrag (Frankfurt 1969). Wij verontschuldigen mensen voor gedrag dat afgedwongen is, omdat we ervan uitgaan dat ze onder andere omstandigheden *anders* zouden hebben gehandeld: 'Zonder bedreiging had ik mijn geld nooit afgegeven (dan had ik dus *anders* gehandeld).' Frankfurt benadrukt echter dat ook als iemand niets anders had kunnen doen, we die persoon nog steeds verantwoordelijk kunnen houden,

namelijk als iemand het deed omdat hij of zij het zelf zo wilde. Wanneer iemand 's ochtends in bed blijft liggen omdat hij gewoon liever doorslaapt, dan maakt het voor wat wij hem verwijten niets uit dat zijn wekker gesaboteerd is of een duivelse wetenschapper klaar staat om zijn brein te manipuleren indien hij op het laatste moment toch zou besluiten op te staan. Wat wij hem verwijten, aldus Frankfurt, is dat hij liever blijft liggen en het feit dat het déze wens is die maakt dat hij zich, lekker snurkend, nog eens omdraait.

In tegenstelling tot Spinoza en Einstein argumenteert Frankfurt dat wat ons onderscheidt van andere wezens niet zozeer ons bewustzijn is, maar de hiërarchische structuur van onze wil (Frankfurt 1972). Wij hebben niet alleen verlangens die gericht zijn op de wereld -- het verlangen om in bed te blijven liggen, te roken, rijk te worden of filosoof -- maar eveneens verlangens met betrekking tot deze verlangens. Hij noemt deze tweede categorie verlangens 'tweede-orde-verlangens'. Voorbeelden van zulke tweede-orde-verlangens zijn het verlangen om niet langer te verlangen naar sigaretten, het verlangen dat het verlangen om rijk te worden ons niet langer gijzelt en het verlangen dat het verlangen om filosoof te zijn ons ertoe zal aanzetten op tijd uit bed te komen. Wij hebben dit soort van tweede-orde-verlangens, aldus Frankfurt, omdat wij geven om wat voor een soort persoon wij zijn, welke verlangens ons handelen sturen en ons leven vormgeven. Het spreken in termen van vrije wil, aldus Frankfurt, moeten we begrijpen tegen deze achtergrond. Iemand die handelt op grond van eigen verlangens én vanuit een toestand van instemming en tevredenheid met deze verlangens, handelt uit vrije wil. Iemand die niet handelt op basis van eigen verlangens -- of dat wel doet, maar niet vanuit een

toestand van instemming en tevredenheid -- is onvrij. Iemand die 's morgens in bed blijft liggen omdat hij dat wil en het prima vindt dat hij dat wil, handelt uit vrije wil (ongeacht de vraag of hij ook anders had kunnen handelen). Op grond daarvan kunnen wij hem verantwoordelijk houden als hij niet op zijn afspraak verschijnt.

Volgens deze opvatting is iemand die bijvoorbeeld dwangmatig steelt, terwijl hij dat niet zou willen doen, onvrij. Iemand die een boek steelt omdat hij geen zin heeft ervoor te betalen voordat hij weet of het goed genoeg is, handelt uit vrije wil. Let wel, beide handelingen zijn uitputtend bepaald door een combinatie van de biologische constitutie, de opvoeding en sociale setting van de actor. Het enige verschil is dat de éne persoon zijn verlangen om te stelen omarmt en daarmee 'vrij' genoemd kan worden, terwijl de ander het verlangen om te stelen afwijst, maar er wel door gedreven wordt en daarmee onvrij genoemd kan worden. Voor Frankfurt staat iets als verslaving dan ook niet noodzakelijk op gespannen voet met de vrije wil. Als men zich identificeert met het verlangen om bijvoorbeeld te roken, dan is daarmee het opsteken van een sigaret niet langer te beschouwen als een onvrije handeling.

Deze visie op de vrije wil noemt men een 'hiërarchische' of 'a-historische' visie, omdat de ontstaansgeschiedenis van onze verlangens er in deze opvatting niet toe doet, maar slechts onze verhouding tot het verlangen dat ons motiveert op het moment van handelen. Met andere woorden, het gaat om de hiërarchische verhouding tot ons verlangen, niet om de oorzaken die aan dat verlangen ten grondslag liggen. Vanwege dit hiërarchische aspect van Frankfurts visie is de these van het determinisme volstrekt irrelevant voor

ons onderscheid tussen vrije en onvrije handelingen. Ook al is de herkomst van al onze handelingen en al onze verlangens volstrekt bepaald door voorgaande gebeurtenissen, wat ertoe doet is onze verhouding tot die verlangens en handelingen.

Aan Frankfurts werk kleven tal van bezwaren en problemen die tot op de dag van vandaag bediscussieerd worden. Om er één te noemen: Wat te denken van psychopaten die hun lugubere verlangens omarmen? Zijn deze mensen niet geestelijk gestoord en daarmee onvrij in plaats van, wat Frankfurts visie impliceert, vrij? En wat als iemand volledig geïsoleerd opgroeit en door zijn opvoeding een totaal vertekend beeld van de wereld krijgt en van wat belangrijk en moreel (on)juist is? Zo iemand kan het toch niet helpen dat hij instemt met verlangens die wij als moreel onjuist kwalificeren? Deze twee laatste vragen zijn opgepakt door de eerdergenoemde Wolf in haar boek *Freedom within Reason* (Wolf 1990). Zij beweert dat een grote tekortkoming van hiërarchische visies op de vrije wil precies is dat zij geen onderscheid kunnen maken tussen mentaal gezonde 'normale' mensen en mentaal zieke 'afwijkende' mensen, noch tussen mensen met en zonder een afwijkende achtergrond en geschiedenis. Dit in tegenstelling tot haar eigen 'Reason View' die stelt dat het ons vermogen is om over goed en kwaad te reflecteren en in overeenstemming met de uitkomsten daarvan te handelen, dat bepaalt of iemand al dan niet moreel verantwoordelijk gehouden kan worden gehouden voor zijn daden.

Volgens sommigen echter is het probleem met Wolfs en vergelijkbare aanpassingen van Frankfurts positie dat het opnieuw problemen introduceert met de these van het determinisme. Stel dat iemand iets afgrijselijks doet, maar wel over alle vermogens lijkt te beschikken die haar

of hem in staat stellen het juiste te doen om de juiste redenen. Moeten we dan niet concluderen dat deze persoon ten tijde van zijn vreselijke daad anders had kunnen doen dat hij feitelijk deed? En hoe kunnen we eigenlijk betekenisvol spreken over mensen die weten wat goed en kwaad is, die uitstekend in staat zijn hun gedrag daarmee in overeenstemming te brengen en het toch niet doen? Is dat geen kenmerk van mentale ongezondheid? Hebben we naast hoe mensen zich gedragen, hoe ze handelen en hoe ze over dit handelen spreken, andere criteria om te bepalen of ze mentaal gezond zijn, of ze goed van kwaad kunnen onderscheiden en ermee in overeenstemming kunnen handelen?

Frankfurts visie is juist zo aantrekkelijk omdat hij een concrete en zinvolle betekenis geeft aan het spreken over de vrije wil, ook binnen een deterministisch wereldbeeld. Het lijkt voor ons inderdaad een verschil te maken of handelingen het gevolg zijn van gewilde en omarmde verlangens of van verlangens waarvan de actor zich distantieert, ongeacht de beschikbare alternatieven. Wat de actor *feitelijk* motiveert, *waarom* hij deed wat hij deed, dat zijn de vragen die ons doorgaans bezighouden wanneer we ons beraden over onze morele evaluatie van de handeling of actor. Het gaat niet om het gedrag op zich, maar om de reden voor het gedrag (al zijn deze natuurlijk niet strikt te scheiden). Een buurvrouw die haar belastingformulieren niet naar waarheid invult omdat ze het geld wil gebruiken voor een feestje, roept een andere morele reactie bij ons op dan een buurvrouw die, met moeite en schaamte, fraudeert om voor haar zieke dochtertje een poppenhuis te kopen. Slechte gemeenschapszin tegenover goedwillende moederliefde, dat is waar het ons om gaat. Laat me dit punt nog wat verder uitwerken.

[k3]1.2 De communicatieve functie van onze morele sentimenten

[t1]Stel dat Einstein getuige is van het gedrag van deze buurvrouw. Overspoeld door morele verontwaardiging denkt hij aan Schopenhauer en voelt een rust over zich neerdalen. Ten gevolge van deze gemoedsbeweging zal hij echter nooit achterhalen of zijn buurvrouw een feestbeest of een verscheurde moeder is. Evenmin zal hij achterhalen of gebrek aan kennis haar parten speelde, of dat zij zich met haar gedrag, zeer kien, burgerlijk ongehoorzaam verzet tegen de belastingregels in haar land.

Onze morele sentimenten hebben een belangrijke sociale en communicatieve functie. Onze morele verontwaardiging of wrok vertelt ons dat we vinden dat er wat ongeoorloofds heeft plaatsgegrepen. Als we onze morele sentimenten uitspreken, wordt bovendien duidelijk dat we menen:

[t2](1) dat de actor iets verkeerd heeft gedaan,
(2) dat zij/hij dat deed in omstandigheden die dat niet excuseren, en
(3) dat we de actor in staat achten datgene na te laten.

[t1]Daarmee communiceren we naast onze normen en waarden ook de specifieke normatieve verwachtingen waarin deze zich laten vertalen. We zeggen niet alleen dat we het gedrag in kwestie slecht en verwerpelijk vinden, maar ook dat niets in de omstandigheden dat gedrag goedpraat én dat de actor een persoon is die beter had moeten weten en anders had moeten handelen. Dit in tegenstelling tot bijvoorbeeld ons afgrijzen bij het zien van al het leed

dat wordt aangericht door een natuurramp of de wreedheid waarmee een kat een vogel verscheurt. In die gevallen kwalificeren we wat er gebeurt ook als slecht, maar niet door daar een moreel verantwoordelijke voor aan te wijzen. Doordat we bijvoorbeeld misprijzend tegen onze buurvrouw uitvallen, wordt haar duidelijk dat we haar gedrag als fraude zien, dat we dat verkeerd vinden en menen dat zij, in haar omstandigheden, dat niet had moeten doen. Als reactie daarop zou onze buurvrouw, op haar beurt, moreel verontwaardigd kunnen worden ('Waar bemoeit u zich mee!'), kunnen schrikken omdat zij zich van geen kwaad bewust was (slechts heel onhandig met cijfers), of ons kunnen uitleggen waarom zij de belastingregels in dit land bizar vindt en geen zin heeft zich daaraan te houden. Door te reageren op haar als een moreel verantwoordelijk persoon, worden wij wijzer over wat er precies aan de hand is en kan zij zich nader verklaren en ons standpunt eventueel corrigeren.

Door elkaar aan te spreken op gedrag waaraan wij aanstoet nemen en onszelf als reactie daarop te verklaren, te verdedigen of in de tegenaanval te gaan, stellen we voortdurend onze wederzijdse normatieve verwachtingen scherp, en bij. Deelname aan een dergelijke conversatie maakt ons actieve leden van een morele gemeenschap. Einsteins gemoedrust lijkt een prijs te hebben. Hij plaatst zichzelf én zijn buurvrouw ermee buiten de morele conversatie.

Let wel, wanneer wij iemand moreel verantwoordelijk houden, dan doen we dat natuurlijk niet om onze normen en waarden te communiceren. Een dusdanig pragmatische houding kenmerkt misschien onze houding ten opzichte van heel jonge kinderen of onze huisdieren, maar ten aanzien van onszelf en elkaar zitten we, zoals Strawson terecht observeerde, veel simpeler in elkaar. Onze frauderende

buurvrouw maakt bij ons onmiddellijk en automatisch morele verontwaardiging los, of als we wat rustiger en ingetogener van karakter zijn of onze buurvrouw ons dierbaar is (of juist koud laat): een milder equivalent van dat sentiment. Strawson begint zijn essay dan ook met een onderscheid tussen pessimisten en optimisten ten aanzien van onze praktijk van morele verantwoordelijkheid en niet, zoals tot dan toe gebruikelijk was, met het onderscheid tussen incompatibilisten en compatibilisten.^v Optimisten zien niet zoveel problemen aan de horizon waar het gaat om de vrije wil en onze morele praktijk: deze dient duidelijk een doel en zal zijn betekenisvolheid echt niet verliezen door een of andere theoretische discussie over de verenigbaarheid van de vrije wil met de these van het determinisme.^{vi} Pessimisten, aldus Strawson, zijn verontwaardigd over die houding. Die verontwaardiging, merkt Strawson fijntjes op, is *zelf* moreel van aard. Het verraadt dat ook de pessimist die twijfelt aan de legitimiteit van onze morele praktijk, zich in het centrum van onze morele praktijk bevindt, een praktijk waarin wij ons automatisch en vanzelfsprekend tot elkaar verhouden als moreel verantwoordelijke personen. Maar wat moet dit argument nu precies bewerkstelligen? Waarom zijn er filosofen die menen dat de implicatie van het werk van Strawson en Frankfurt is dat twijfel aan het bestaan van de vrije wil onzinnig is?

[k3]1.3 Legitimiteit van onze morele sentimenten

[t1]In zekere zin verleent het feit dat onze morele sentimenten een communicatieve functie dienen, hen een bepaalde legitimiteit. Onze omgang met elkaar is door en door gereguleerd door normen en waarden, veelal op een onbewust en automatisch niveau. Zoals al eerder

opgemerkt, reageren we onmiddellijk en direct moreel verontwaardigd wanneer iemand ons in het verkeer lomp afsnijdt. Afgezien van de verkeersregels die met een dergelijke actie overtreden worden, schendt deze ook onze morele verwachtingen ten aanzien van normale omgangsvormen. Recent onderzoek in de cognitie-, gedrags- en neurowetenschappen suggereert dat de mate waarin deze regulering zich onbewust en automatisch voltrekt, niet onderschat moet worden, terwijl we dat wel doen (zie de bijdrage van Tiemeijer in deze bundel). Indien wij grotendeels op een onbewust en automatisch niveau functioneren, is het van nog groter belang om onszelf uit te spreken en elkaar te blijven aanspreken op gedrag dat ons niet zint. Zelfs wanneer er over onze normen en waarden grote overeenstemming bestaat, valt er nog voldoende af te stemmen over wat dat impliceert voor ons gedrag (zie het voorbeeld van onze buurvrouw). Iemand geen verwijten maken voor wat hij in onze ogen verkeerd doet, heeft als gevolg dat hij als morele actor wordt uitgesloten van de morele conversatie. Hij wordt door het uitblijven van een morele reactie en morele kritiek niet in de gelegenheid gesteld om redenen aan te dragen voor het vertoonde gedrag, redenen die wellicht een heroverweging in gang zetten van de morele kwalificatie van het betreffende gedrag en de normen en waarden die onze omgang in het vervolg zullen reguleren. Als ik mijn buurvrouw erop aanspreek dat ik haar gedrag verwerpelijk vind, maar zij mij kan overtuigen van de goede gronden voor haar gedrag, dan zal dat niet alleen mijn houding ten aanzien van haar gedrag veranderen, maar wellicht ook mijn eigen toekomstige gedrag.

In sommige gevallen lijkt een dergelijk uitsluiten van de morele conversatie een aanbevelenswaardige strategie én een begrijpelijke reactie. Het toestaan van

bepaalde overwegingen in het morele discours kan destructief zijn voor onze samenleving. Neem bijvoorbeeld moord 'in koelen bloede' en pedoseksuele handelingen. Wanneer wij dat soort handelingen begrijpen als handelingen waarvoor men goede redenen kan hebben, betekent het dat wij de overwegingen (redenen) die ertoe geleid hebben serieus nemen. Maar dat willen we toch niet doen?! We willen toch vooral voorop stellen dat niets, maar dan ook niets ooit mag leiden tot moord en pedoseksuele handelingen. In deze gevallen is het veel veiliger om de daders te beschouwen als geestesziek en onze oren te sluiten voor de rechtvaardigingen die zij zelf naar voren brengen.^{vii}

Als het gaat om meer alledaagse morele overschrijdingen, lijkt uitsluiting echter altijd onverantwoord paternaliserend en ook contraproductief (Sie 2005). Waarom mensen uitsluiten van een morele discussie als daar geen concrete aanleiding voor is? Onze buurvrouw doet iets dat *door ons* als verkeerd wordt waargenomen, maar wat vindt zij daar zelf van? Telt wat zij over de kwestie denkt niet mee? Haar bij voorbaat excuseren of wegzetten als moreel incapabele gesprekspartner, heeft het effect dat zij letterlijk niet gehoord wordt. Het is contraproductief in zoverre de buurvrouw het met ons eens zou kunnen zijn, maar zich van geen kwaad bewust. Als we haar niet op haar gedrag aanspreken, zal zij er zeker niet wijzer van worden. In een samenleving die openstaat voor mogelijke onenigheid en vergissingen, zijn onze morele sentimenten vooral invitaties tot dialoog. Kortom, er pleit veel voor het strawsoniaanse adagium dat wij er goed aan doen elkaar als moreel verantwoordelijke personen te bejegenen.

Terug naar de vrije wil. Als men de vrije wil ziet als dat vermogen dat de communicatieve en sociale functie

van onze houding tot elkaar als moreel verantwoordelijke personen verklaart en legitimeert, is het begrijpelijk dat men de gedachte dat de vrije wil een illusie is, van de hand wijst. Deze gedachte bergt immers het gevaar in zich dat wij mensen op voorhand excuseren die in onze ogen moreel verwerpelijk handelen of hen niet langer serieus nemen als moreel verantwoordelijke actoren. Een dergelijke houding ondermijnt onze natuurlijke en interpersoonlijke verhoudingen, waarbinnen wij direct en onmiddellijk op elkaar reageren met de morele emoties. En een dergelijke verstoring van onze natuurlijke en interpersoonlijke verhoudingen lijkt moreel niet verdedigbaar. Wil dat zeggen dat de gedachte dat de vrije wil een illusie is onzinnig is, en dat Einstein gemoedrust en troost vond waar geen schouder is? Zoals ik in de volgende paragraaf zal laten zien, is er, ook wanneer men het enorme belang van onze alledaagse praktijk erkent, nog ruimte om de vraag naar het bestaan van de vrije wil als een maatschappelijk belangrijke vraag te beschouwen.

[k2]2 De vrije is een illusie

[t1]Wanneer wij er niet in slagen 'de golfbal te putten', zoals John Austin opmerkt, dan voelt het alsof we gefaald hebben (Austin 1956). Dat gevoel is niet gerelateerd aan de omstandigheden van onze poging. Wat door ons hoofd schiet is: ik had het beter kunnen en moeten doen, en wel in exact dezelfde omstandigheden. Vanuit het eerste-persoonspectief is het idee dat wij handelende wezens zijn, vrij om iets te doen of te laten (ongeacht de omstandigheden en gebeurtenissen in het verleden), zo goed als onvermijdelijk. In navolging van Immanuel Kant hebben velen erop gewezen dat wij niet

anders kunnen dan handelen onder de vlag van de vrijheid. Daar komt bij dat het gevoel te hebben gefaald, beter te kunnen en moeten, en het vermogen ons te oriënteren op hoe wij in de toekomst zaken kunnen vermijden of beter kunnen doen, essentiële ingrediënten zijn van menselijk handelen.^{viii} Naast een vrijwel onmogelijk op te geven gevoel lijkt het daarmee ook moreel onwenselijk om het gevoel dat wij uit vrije wil (kunnen) handelen, op te geven. Dit wordt nog versterkt door moreel psychologisch onderzoek dat claimt aangetoond te hebben dat mensen zonder twijfel aan de vrije wil moreel beter handelen (Vohs en Schooler 2008).

Let wel, dat alles wil niet zeggen dat daarmee aangetoond is dat wij op enig moment in ons leven 'onbewogen bewegers' zijn: in staat tot bewuste aansturing (mentale veroorzaking die beweging in gang zet) zonder dat wij ons in een keten van oorzaken bevinden (het bewogen worden). Evenmin wil het zeggen dat wij ooit in exact dezelfde omstandigheden 'anders hadden kunnen doen dan wij feitelijk deden', noch dat onze morele sentimenten gelegitimeerd zijn omdat en in zoverre dat het geval is. Hier rijst langzaam de schouder op die Einstein vond.

De concepten van de vrije wil en morele verantwoordelijkheid zijn nauw verbonden met het idee van persoonlijke verdienste en vooral ook schuld: verdienste en schuld die de actor, en de actor alleen, zijn aan te rekenen. Wij baseren er niet alleen onze alledaagse omgangsvormen op maar ook bijvoorbeeld onze strafrechtelijke praktijk (zie de bijdrage van Buruma in deze bundel). Het is de aanwezigheid van schuld, verwijtbare schuld, die maakt dat we de ene dader tot straf veroordelen en de andere tot behandeling in een kliniek. Maar, zoals Thomas Nagel opmerkt in zijn

adequaat getitelde essay 'Moral Luck' (ik kom op deze titel nog terug), wanneer we alle factoren in ogenschouw nemen die een rol spelen bij ons handelen maar die buiten de controle van ons als actor liggen, dan verdampt het punt van actorschap volledig (Nagel 1979). Waarin is het onderscheid dat wij maken dan gefundeerd? We behandelen de geesteszieke als een persoon die door zijn ziekte geen controle heeft op zijn slechte daden, maar waarin verschilt dit gebrek aan controle van ieders gebrek aan controle? Zijn we niet allemaal met een bepaalde emotionele en cognitieve *make-up* geboren? Heeft onze opvoeding geen sporen achtergelaten in wie wij zijn geworden? Maken opvoeding en biologische constitutie niet dat wij op de ene manier en niet op de andere reageren op wat ons overkomt? Kortom, waarop is het idee gebaseerd dat normale volwassenen controle hebben over hun (slechte) daden, terwijl heel jonge kinderen en mensen die lijden aan een geestelijke aandoening dat niet of slechts in beperkte mate hebben?

Wanneer we tegen deze achtergrond twijfelen aan het bestaan van de vrije wil, betekent dit dat we twijfelen aan het bestaan van een substantiële, individueel te dragen, persoonlijke schuld voor verricht kwaad. Een universum waarin menselijk handelen wetmatig schakelt in een ketting van oorzaak en gevolg, lijkt voor een dergelijke twijfel geen plaats te hebben:

[t2][Als] de vloed van verlangens de zwakke man verleidt te doen wat hij deed, [...] dan is hem dat niet meer aan te rekenen dan de slecht geconstrueerde dam die breekt onder de druk van het water. (Chisholm 1964)

[t1]Als de vrije wil een illusie is, dan noodzaakt een confrontatie met een moreel laakbare handeling of

actor ons ertoe te zoeken naar verklarende, excuserende en verontschuldigende omstandigheden. Het dwingt ons de vraag te stellen of de grove morele nalatigheid en persoonlijke schuld die we menen te zien en waarop we automatisch en onmiddellijk reageren, in wezen niet slechts het resultaat is van een optelsom van hele menselijke vergissingen die eenieder van ons had kunnen maken.

Uiteraard doen wij aan de lopende band nare, vreselijke, en afgrijselijke dingen, maar de vraag die relevant is wanneer we het bestaan van de vrije wil bespreken, is hoe we dat feit begrijpen en verklaren. We zijn geen van allen heilig, we hebben onze eigenaardigheden, onze emotionele tekortkomingen en onze blinde vlekken. Het mag geen wonder heten dat dit er af en toe toe leidt dat we ons moreel niet geheel voorbeeldig gedragen. De verleidingen zijn overvloedig, het vlees zwak. Daarnaast zijn er natuurlijk ook mensen die door een combinatie van factoren -- hun opvoeding, verkeerde vrienden, tekortschietende mentale reserves, gebrekkige cognitieve en emotionele vaardigheden -- totaal de weg zijn kwijt geraakt. Als zij in de verkeerde omstandigheden belanden, koersen ze regelrecht af op een moreel rampgebied waarbij agressie en gewelddadigheid tot moord en doodslag kunnen leiden. Het waarnemen van een dergelijk slagveld (moord, agressie, gewelddadigheid), maar ook van het leed en de ellende die onze minder grote tekortkomingen kunnen aanrichten (ontrouw, onoplettendheid, ongevoeligheid, achteloosheid) kan tot gevoelens van ongeduld en wanhoop leiden. Je geduld bewaren ten aanzien van een krijsend kind dat de verkeerde kleur beker heeft gekregen (of je moeder die je de verkeerde kleur beker heeft gegeven) verbleekt daarbij. 'Jullie hadden elkaar nog zo beloofd voortaan

eerlijk en open te zijn, hoe kan iemand dat vertrouwen beschamen?' 'Zo lang bevriend en dan zo'n laffe streek!' 'En dan heb je zo veel geld en zo veel geluk in het leven en dan doe je zoiets?' Als alle morele zondes, klein of groot, uitpuittend te verklaren zijn door een combinatie van factoren die de daders persoonlijk niet werkelijk aan te rekenen zijn, dan kan het besef dat dat zo is gemoedrust brengen.

Zoals uitgelegd in de vorige paragraaf, impliceert dat nog niet dat we onze praktijk volledig moeten veranderen, dat we onze onmiddellijke en natuurlijke reacties op elkaar, voorzover überhaupt mogelijk (zie de bijdrage van Anderson in deze bundel), moeten afleren en veranderen. Zoals betoogd brengen deze natuurlijke en onmiddellijke reacties op elkaar ons veel voordelen. Wel impliceert dit besef dat we een zekere terughoudendheid moeten betrachten in het toeschrijven van grove morele nalatigheid en schuld en de praktijken die daaraan verbonden zijn. Het conceptuele voorbehoud dat de vrije wil wellicht een illusie is, kan daarbij een grote dienst bewijzen, een dienst die wij wellicht wel eens harder nodig kunnen hebben dan we voorheen dachten. Uit veel recent en divers wetenschappelijk onderzoek blijkt dat onze emoties een enorme rol spelen in ons evalueren en beoordelen van de wereld. Deze emoties zijn veelal onbewuste en automatische reacties op wat ons overkomt en de situatie waarin dat het geval is. Bovendien zijn vele wetenschappers ervan overtuigd dat deze emoties ons vaak een vertekend beeld geven van wat er werkelijk 'objectief gezien' aan de hand is, en wat dus eigenlijk onze oordelen zou moeten informeren. Een dergelijke conclusie is, volgens de meeste filosofen, veel te voorbarig. In de context van onze morele houdingen ten opzichte van elkaar is er echter wel degelijk grond voor enige

terughoudendheid. Laat ik een paar voorbeelden bespreken (zie voor meer stof tot nadenken de bijdragen van Bransen en Tiemeijer).

Uit de psychologie weten we inmiddels dat we de morele misstappen die we een ander zien begaan, gemakkelijk wijten aan diens karakter en de morele fouten daarin, terwijl we met betrekking tot onze eigen misstappen vooral oog hebben voor de omstandigheden. Dit verschijnsel staat bekend als de 'fundamentele attributiefout'. Het komt erop neer dat wij bij het zien van andermans gedrag geneigd zijn de ander moreel te veroordelen voor dat gedrag en er conclusies aan te verbinden over het karakter van de persoon in kwestie: 'Zij/hij is altijd al lui geweest, nooit helemaal eerlijk, erg gemakkelijk met het geld van anderen.' Als wij ons eigen afkeurenswaardig gedrag beoordelen, hebben we vooral oog voor wat er uitzonderlijk was aan de situatie en omstandigheden: 'Ik was erg moe, had net een lelijke tegenvaller gehad, was misleid door een ander.'

Er zijn wetenschappers en filosofen die menen dat ons alledaagse toeschrijven van moreel laakbaar gedrag aan karakterfouten op een vergissing berust (Doris 2002). Deze 'situationisten' beweren dat we bij moreel laakbaar gedrag vooral oog moet hebben voor de omstandigheden van het afwijkende gedrag en geen illusies moeten koesteren omtrent ons individuele vermogen om in vergelijkbare omstandigheden wel moreel juist te handelen. Situationisten wijzen in dit verband op de schokkende uitkomsten van het gedrag dat proefpersonen vertonen in het beruchte Milgram Experiment en het Stanford-Prison Experiment. Interessant in verband met de hierboven besproken fundamentele attributiefout is dat wij, wanneer ons gevraagd wordt het gedrag van *anderen* te voorspellen in emotioneel en moreel belastend situaties zoals die van

bovengenoemde experimenten, tamelijk accuraat voorspellen wat de uitkomst is van het feitelijke experiment. Maar wanneer ons gevraagd wordt ons *eigen* gedrag te voorspellen in dergelijke situaties, menen wij vrijwel allemaal dat we ons zullen gedragen zoals slechts een zeer kleine minderheid zich feitelijk blijkt te gedragen.^{ix} Kortom, onze voorspellingen over het gedrag van anderen zijn een stuk adequater en minder rooskleurig dan onze voorspellingen over ons eigen gedrag.

Naast het onderkennen van de invloed van deze fundamentele attributiefout is er inmiddels grote overeenstemming over het feit dat onze algemene stemming en onze onderbuikgevoelens van grote invloed zijn op onze morele emoties. Ongelukkigerwijs zijn zowel onze stemming als onze onderbuikgevoelens eenvoudig te beïnvloeden en te bespelen door manipulatie van 'irrelevante' omstandigheden. Onze stemming wordt beïnvloed door bijvoorbeeld slecht weer, lollige films of een door ons gevonden euro; onze onderbuikgevoelens door vieze tafels, ongewassen handen en een stinkende geur. Als mensen bijvoorbeeld een bepaalde casus moeten beoordelen, dan zullen zij deze moreel zwaarder veroordelen wanneer zij dat doen achter een vieze tafel waarop lege pizzadozen slingeren en volle asbakken staan, dan wanneer zij dit doen van achter een schoon en opgeruimd bureau (Schnall e.a. 2008).

Deze uitkomsten zijn verontrustend omdat de toestand van het bureau waarachter wij een oordeel vellen, niets te maken heeft met de morele kwaliteit van de in de casus besproken handeling of actor. Als we ons van deze invloeden bewust zouden zijn, zouden we ze van de hand wijzen als goede grond voor ons oordeel. Dit betekent dat onze morele oordelen, als zij al een betrouwbare indicatie zijn van de morele kwaliteit van gedrag of

personen, zeker met een korrel zout genomen moeten worden. Dit soort onderzoek zegt op zichzelf niet veel over de principiële legitimeerbaarheid van onze morele houdingen ten aanzien van elkaar. Het maakt echter wel inzichtelijk hoe de filosofische vraag naar de vrije wil ook een positieve invloed kan hebben op onze morele praktijk. Dit is zo, ongeacht het gelijk van diegenen die het werk van Strawson en Frankfurt tot uitgangspunt nemen en waarschuwen tegen de neiging om onze alledaagse praktijk te ondermijnen naar aanleiding van de theoretische en conceptuele kwestie van de vrije wil. Juist wanneer onze morele houdingen ten aanzien van elkaar moeilijk op te geven zijn, vormt het probleem van de vrije wil een goede herinnering aan de mogelijke kanttekeningen bij die houdingen en de daaraan gerelateerde praktijken en instituties.

Dat wij de vraag naar de vrije wil stellen toont aan dat we, hoewel we op elkaar reageren vanuit een morele basishouding, wel degelijk in staat zijn afstand te nemen van deze emotionele betrokkenheid op elkaar. En ook dat is een onlosmakelijk ingrediënt van onze morele praktijk. De hierboven reeds aangehaalde Nagel ontkent de claim die ten grondslag ligt aan het invloedrijke essay van Strawson, namelijk dat het 'probleem van de vrije wil' een abstract, theoretisch en metafysisch probleem is dat niet zo interessant is voor onze morele praktijk.* Volgens hem ontstaat het probleem juist in onze morele praktijk door het simpelweg consequent en eerlijk doorvoeren van de alledaagse criteria die wij normaal hanteren om mensen te excuseren en verontschuldigen voor hun kleine en grotere misdragingen en daden. Wij houden iemand niet verantwoordelijk voor handelingen waarover hij geen controle had, voor handelingen waaraan hij niets kon doen. Wanneer we inzoomen op alle invloeden die maken dat

iemand doet wat hij doet, dan blijft er niets over dat onze morele verwijten aan het adres van de dader legitimeert. Als wij de dader desalniettemin zijn gedrag verwijten, dan heeft deze persoon morele pech: de pech dat diens omstandigheden, karakter, opvoeding, handelingen et cetera moreel slecht uitvallen in plaats van neutraal of goed. Vandaar de titel van Nagels essay 'Moral Luck' (het Engelse *luck* is neutraler dan het Nederlandse 'geluk' en verwijst niet noodzakelijk naar positieve zaken; *bad luck* is een gangbare uitdrukking voor 'pech').

Als we dat met Nagel eens zijn kunnen we verleid worden om een van de meer exotische posities in het debat over de vrije wil en morele verantwoordelijkheid te verdedigen, zoals bijvoorbeeld het harde determinisme. Deze positie draait niet om de hete brij heen en beaamt volmondig dat wij onze alledaagse praktijk waarin wij elkaar moreel verantwoordelijk houden, niet kunnen rechtvaardigen.^{.xi} Wij moeten onze 'morele kwaadheid', zoals bijvoorbeeld de Amerikaanse filosoof Derk Pereboom onze onmiddellijke en natuurlijke reacties ten aanzien van elkaar noemt, niet tot uitgangspunt nemen voor onze houding tot elkaar en evenmin voor de inrichting van onze samenleving (Pereboom 2001).^{.xii} Wij moeten proberen van onze heftige morele emoties af te komen en onze samenleving drastisch herinrichten.

Richard Dawkins illustreerde de misplaatstheid van onze morele emoties enige jaren geleden in de NRC door ze te vergelijken met die van Basil (John Cleese) uit *Fawlty Towers*. In een van de vele gedenkwaardige scènes uit die serie gaat Basil te keer tegen zijn auto die niet wil starten. Na de auto eerst drie keer te hebben gewaarschuwd, ranselt hij hem vervolgens hardhandig met een boomtak af omdat het 'zijn verdiende loon is'. Als

een auto dienst weigert brengen we hem naar de garage, waarom zouden we niet iets vergelijkbaars doen met de crimineel handelende medemens? (NRC Handelsblad, 14 januari 2006) Het onderscheid tussen *mad* en *bad* heeft volgens iemand als Dawkins en Pereboom zijn beste tijd gehad (zie ook de bijdrage van Meynen in deze bundel).

Als men een dergelijke visie absurd vindt, maar evenmin meent dat de door Frankfurt en Strawson geboden perspectieven erg behulpzaam zijn, dan staat er ten slotte nog een derde benadering van de hele vrije-wil- kwestie open. Wat als de vrije wil nu eens wél bestaat en onze emotionele reacties op elkaar heel goed corresponderen met handelingen die uit vrije wil verricht zijn? Ook als men de stelling onderschrijft dat vrije wil en determinisme niet met elkaar verenigbaar zijn, kan men verdedigen dat onze morele praktijk legitiem is, namelijk omdat het determinisme onjuist is. Hoewel het er niet veel zijn, zijn er enkele hedendaagse filosofen die een dergelijke positie voor hun rekening nemen. Laat ik dit hoofdstuk afsluiten met een korte paragraaf over het werk van één zo'n filosoof: de Amerikaanse libertariër Robert Kane.

[k2]3 Het bestaan van de vrije wil

[t1]Kane begint zijn boek *The Significance of Free Will* met een zorgvuldige uitleg van wat volgens hem het probleem is van de vrije wil dat rest nadat wij Frankfurts argumenten tegen het zogenaamde principe van alternatieve mogelijkheden hebben geaccepteerd (Kane 1996, Frankfurt 1969 en de bijdrage van Cuypers in deze bundel). Kane meent namelijk dat Frankfurt gelijk heeft wanneer hij argumenteert dat wij moreel verantwoordelijk

gehouden kunnen worden voor handelingen zonder beschikbaar alternatief.

Sommige handelingen, aldus Kane, vloeien voort uit het wezen en de kern van onze persoon en zijn intrinsiek verbonden met wie wij op een bepaald moment geworden zijn. 'Hier sta ik, ik kan niet anders' is een in de discussie over de vrije wil veel geciteerde uitspraak van Luther. Volgens velen een uitspraak die aantoont dat het gebrek aan vrije wil geenszins strijdig is met handelingen waarvoor wij de actor moreel verantwoordelijk houden. Voor hen is Luthers daad een toonbeeld van moreel voorbeeldig gedrag. Luther bedoelde zich met deze uitspraak overduidelijk niet te excuseren voor zijn gedrag. Kane is het met die observatie eens, zelfs wanneer wij daarbij aantekenen dat het letterlijk juist is dat Luther op het moment van deze uitspraak geen keuze meer had. Hij wijst er echter op dat Luthers verantwoordelijkheid een afgeleide is van het feit dat hij medeverantwoordelijk is voor de persoon die hij op het moment van zijn fameuze woorden geworden is. En dat geldt natuurlijk ook voor diegenen die zich moreel laakbaar gedragen, ook al hadden zij op het moment van hun daad geen keuzevrijheid.

Alleen als wij onszelf mede vormgeven, als wij een stempel kunnen drukken op wie wij zijn, op het soort mensen dat wij worden, kunnen de handelingen die daaruit noodzakelijk en onvermijdelijk voortvloeien ons uiteindelijk aangerekend worden. Dit noemt Kane *ultimate responsibility*. Een dergelijke 'uiteindelijk verantwoordelijkheid' kan alleen bestaan in een wereld die niet volledig gedetermineerd is, in een wereld die ruimte laat voor wat hij 'zelf-vormende-handelingen' noemt.

Zelf-vormende handelingen zijn letterlijk handelingen die onbepaald zijn door voorafgaande gebeurtenissen. Handelingen die wij per definitie verrichten op cruciale momenten in ons leven, omdat ze het verdere verloop van ons leven bepalen. Het besluit om te stoppen met een studie, het negeren van een concreet moreel appèl, het moedig aangaan van een confrontatie. Dat allemaal kunnen momenten zijn waarop we iets doen waarmee we, zogezegd, op onszelf vooruitgrijpen. Op die momenten twijfelen we waarschijnlijk nog, we weten niet precies wat we moeten doen, wat het goede is om te doen, en of we het soort persoon willen zijn dat zulke keuzes maakt en zulke handelingen verricht. Kane refereert aan deze handelingen ook wel met de term 'waarde-experimenten'.

Kanes gedachtegang is de volgende. Luther heeft een persoonlijke geschiedenis die gekenmerkt wordt door een strijd om de kerk te zuiveren. In die persoonlijke geschiedenis zijn allerlei momenten aan te wijzen waarop hij geconfronteerd werd met te nemen beslissingen, sommige ogenschijnlijk triviaal. Als het determinisme onjuist is, dan was in ten minste enkele van die gevallen de beslissing die Luther moest nemen geheel aan hem. Wát hij besloot vormde zijn persoonlijkheid. In allereerste instantie waren die besluiten wellicht tentatief, aarzelend, in Kanes woorden: een experiment. Bijvoorbeeld: Luther schrijft in het kader van zijn opleiding een kort en kritisch stuk en vraagt zich af of hij dat wel moet inleveren, omdat hij zijn leermeesters ermee in een lastig pakket zal brengen. Hij aarzelt, piekert, besluit het erop te wagen en gaat de confrontatie aan. Deze kleine daad van moed hoeft nog niet vol overtuiging verricht te zijn. Deze daad is hem wellicht ook nog niet volledig aan te rekenen. Hij

probeert eens wat uit en afhankelijk van het verloop van die poging leert hij of hij het soort persoon wil en kan zijn dat zulke daden verricht. En wederom geldt hetzelfde voor diegenen die van het juiste pad afraken.

Naarmate Luther uitgroeit tot de persoon die geschiedenis zal schrijven, zullen zijn beslissingen steeds onvermijdelijker worden. Tot aan het moment waarop hij, met alle gevolgen van dien, de woorden 'Hier sta ik, ik kan niet anders' moet uitspreken.^{xiii} Had hij ergens in zijn persoonlijke geschiedenis andere besluiten genomen, anders gehandeld, dan had hij wellicht op dat cruciale moment in zijn leven nog wel anders gekund en wellicht anders gedaan. Maar het feit dat hij op dat moment niet anders kon doen dan hij feitelijk deed, vermindert zijn morele verantwoordelijkheid voor die daad niet. Dat wil zeggen: wanneer Luther, zoals Kane meent, het product is van zijn eigen handelen. En dat kan alleen als niet elke handeling in dit ondermaanse bepaald is door gebeurtenissen die eraan voorafgingen (mentale gebeurtenissen inclusief), aldus Kane.

Kane laat vervolgens zien hoe het indeterminisme op microniveau ruimte biedt voor de vrije wil die onze uiteindelijke verantwoordelijkheid kan gronden. In de kern komt zijn argument erop neer dat er momenten in ons leven zijn waarop wij beslissingen kunnen en moeten nemen, omdat de redenen die we hebben voor elk alternatief niet doorslaggevend zijn. Bijvoorbeeld: ik ga door met mijn studie omdat ik er zo goed in ben, of ik stop omdat het me te weinig plezier en voldoening geeft; voor elke keuze bestaan goede redenen. Het is dus geen volstrekte onbepaaldheid of willekeur wat we kiezen -- doorgaan of stoppen met onze studie. De keuze is echter niet bepaald, wij moeten die zelf maken.

Kane visualiseert dit keuze moment als een moment waarop de redenen en het denkproces (waarvan wij ons niet noodzakelijk bewust hoeven te zijn, volgens Kane) die zouden leiden tot de ene keuze, interfereren met de redenen en het denkproces die zouden leiden tot de andere keuze. Men moet dus om een bepaalde keuze te maken een zekere interne weerstand overwinnen die bestaat uit de redenen en het denkproces die ons tot de andere keuze verleiden. Deze 'interne weerstand' kunnen we zien als een soort achtergrondruis. De actor moet als het ware 'door de ruis heen' voor de andere optie kiezen en dat kost concentratie en moeite (meer concentratie en moeite wellicht, naarmate de optie die ruis produceert aantrekkelijker is). Daarmee is er dus een inspanning vereist van de actor, is het de actor zelf die maakt dat de ene set van redenen het verliest van de andere set van redenen. En hiermee sluit Kane zich aan bij een eveneens veelgeciteerde uitspraak van Leibniz: 'Redenen verklaren onze neiging zonder ons te noodzaken' ('Reasons incline without necessitating') (zie bijvoorbeeld Chisholm 1964). Wat we uiteindelijk doen, dat zullen we zelf moeten beslissen.

Dit brengt mij terug bij compatibilisten zoals Wolf en Wallace die menen dat ons vermogen om het juiste te doen om de juiste redenen, cruciaal is voor onze alledaagse praktijk van morele verantwoordelijkheid. Wat sommige filosofen blijft terugbrengen naar de vraag hoe determinisme zich verhoudt tot vrije wil, is precies het feit dat moreel slechte handelingen bestaan uit handelingen waar het goede niet gedaan wordt, of niet op grond van de juiste redenen. Wanneer we het hebben over moreel volwaardige actoren, mensen die wel degelijk het vermogen hebben om het goede te doen om de juiste redenen, betekent dit dan niet gewoon dat zij vrij waren

om iets anders te doen in exact dezelfde omstandigheden? En als dat zo is, betekent dat dan niet toch dat wij in die gevallen aannemen dat de actor (met voldoende morele capaciteiten) die moreel verkeerd handelde (geen uitdrukking en/of gehoor gaf aan zijn morele capaciteiten) vrij was om iets anders te doen dan wat hij feitelijk deed? En als dat zo is, hoe kan dat dan waar zijn tegen de achtergrond van de these dat alles in het ondermaanse uitputtend bepaald is door voorafgaande gebeurtenissen? (Sie 2005) Het zijn vragen als deze, waarnaar filosofen keer op keer terugkeren.

De libertarische positie van iemand als Kane is op z'n zachtst gezegd niet erg populair. Veel filosofen zien niet in hoe het toeschrijven van morele verantwoordelijkheid aan legitimiteit wint door het bestaan van onbepaalde keuzevrijheid. Ook al is er in Kanes versie sprake van redenen, de uiteindelijke keuze wordt niet op basis van die redenen gemaakt. Er waren tenslotte ook redenen voor de alternatieve handelingen. Wat in het verhaal van Kane de doorslag geeft, blijft in zekere zin een 'willekeurige keuze'. De meeste hedendaagse filosofen zien gewoon niet in hoe het bestaan van een dergelijke willekeur ooit kan bijdragen aan de zelfcontrole die noodzakelijk lijkt om morele verantwoordelijkheid te gronden. Bovendien hebben zij, niettegenstaande Kanes pogingen om zijn visie gedetailleerd uit te werken, er veel moeite mee om in te zien hoe een dergelijk proces in het brein eruit zou moeten zien.^{xiv} Toch blijven wetenschappers maar ook filosofen de vraag naar de vrije wil stellen, en zeker zo belangrijk: het blijft een vraag waarvan velen onmiddellijk de betekenisvolheid inzien.

[t1]Ik ben dit hoofdstuk begonnen met een bespreking van de eerste en, in de afgelopen decennia, meest invloedrijke positie binnen de Engelstalige filosofie: het compatibilisme. Compatibilisten menen dat de these van het determinisme verenigbaar is met het bestaan van de vrije wil. Men moet zich realiseren dat de in dit hoofdstuk besproken compatibilistische positie zich onderscheidt van de meer klassieke posities die onder deze noemer gevat werden (Hobbes, zie de bijdrage van Van Ruler in deze bundel, Moore 1911; zie ook Aune 1967 en Lehrer 1968, repr. Watson 1982). In het klassieke compatibilisme verdedigde men grofweg de stelling dat de vrijheid om een bepaalde handeling te doen of te laten, tot een van de oorzaken van die handeling moest worden gerekend. Die oorzaak, aldus klassieke compatibilisten, is onderdeel van de gedetermineerde wereld en daarom is keuzevrijheid niet strijdig met, maar onderdeel van een deterministisch universum.^{xv} In de positie die in dit hoofdstuk centraal staat, laat men de gedachte varen dat wij vrij zijn ervoor te kiezen iets te doen of te laten. Ik heb uitgelegd waarom het werk van de filosofen Strawson en Frankfurt hierin een cruciale rol gespeeld heeft.

In termen van de driedeling uit de inleiding van deze bundel en het boek van Van de Laar en Voerman, *Vrije Wil*, heb ik laten zien hoe het gedachtegoed van Strawson en Frankfurt een radicale verandering markeert in de discussie over de vrije wil als speciaal vermogen dat ons op een of andere wijze buiten de gangbare causale en wetmatige orde plaatst (Van de Laar en Voerman 2011). Een verandering waarin de aandacht verlegd wordt naar het concept van de vrije wil, begrepen als voorwaarde voor morele verantwoordelijkheid en de vrije wil als

persoonlijke zelfverwerkelijking. Ik heb uitgelegd hoe deze twee laatste opvattingen samenhangen. Ik heb laten zien dat voor veel filosofen de twijfel aan het bestaan van de vrije wil voorbijgaat aan allerlei natuurlijke en onmiddellijke onderscheidingen die wij in onze alledaagse morele praktijk maken en de goede redenen die wij hiervoor hebben.

Hierna ben ik ingegaan op de reserves die men kan hebben ten aanzien van deze, zeer invloedrijke, positie. Ook al ziet en begrijpt men het enorme belang van de vrije wil in onze alledaagse morele praktijk, dan nog kan men de vraag stellen of dat argument alle zorgen omtrent het mogelijk illusoire karakter van de vrije wil wegneemt. Veel mensen associëren de vrije wil niet zozeer met onze morele praktijk in het algemeen, maar meer specifiek met het legitimeren van onze negatieve morele houdingen ten aanzien van onszelf en elkaar. Dat wil zeggen: houdingen waarin we onszelf en anderen morele verwijten maken en verontwaardigd zijn, maar ook de daaraan gerelateerde opvoedingspraktijken en institutionele praktijken (straffen en belonen, en de strafrechterlijke en penitaire instituten die daaraan verbonden zijn).

Wanneer we claimen dat iemand uit vrije wil handelde, dan wordt daarmee vaak geïmpliceerd dat deze persoon onze morele afkeuring en de daarmee gepaardgaande straf verdient (zie bijdragen van Meynen en Buruma in deze bundel). Dat wil zeggen: zij of hij heeft het aan zichzelf te danken omdat zij of hij, anders dan bijvoorbeeld een ongehoorzaam huisdier of slecht functionerend apparaat, ook anders had kunnen handelen. Het opwerpen van de vraag naar het bestaan van de vrije wil in de context van onze strafpraktijken (in brede zin begrepen) kan gezien worden als een 'conceptueel

voorbehoud' ten aanzien van de legitimatie ervan. Wat voor dit perspectief pleit, is dat het moeilijk blijft om betekenis te geven aan het idee van individuele verdienste of schuld in een universum dat geen ruimte laat voor een speciaal menselijk vermogen dat ons op een of andere wijze buiten de gangbare causale en wetmatige orde plaatst.

Tot slot, ben ik naar aanleiding van dit laatste perspectief kort ingegaan op één van de minderheidsposities in het hedendaagse debat: de positie dat wij wel degelijk zo'n speciaal vermogen bezitten. In die laatste paragraaf heb ik heel kort het werk van Kane besproken, een van de weinigen die in het hedendaagse debat volhoudt dat als de vrije wil *niet* bestaat, wij niet moreel verantwoordelijk gehouden kunnen worden voor onze handelingen (de incompatibilistische positie). Toch is dat volgens Kane geen reden tot zorg, omdat naar zijn oordeel de vrije wil *wel* bestaat (de libertarische positie).

[k2]Suggesties voor verder lezen

[t1]Dennett, D.C., *Elbow Room: The Varieties of Free Will Worth Wanting*, Cambridge: MIT Press, 1984. Of Dennetts zeer toegankelijke en in het Nederlands vertaalde *Evolutie van de vrije wil*, Amsterdam: Contact, 2004.

Kane, R., *Contemporary Introduction to Free Will*, Oxford Fundamentals of Philosophy, Oxford: Oxford University Press, 2005.

Wolf, S., *Freedom within Reason*, New York: Oxford University Press, 1990.

Mijn eigen boek *Justifying Blame* bevat een systematische samenvatting van de hedendaagse discussie

en de visie waarop het hoofdstuk in deze bundel ten dele is gebaseerd, hoofdstukken zijn ook los te verkrijgen in het elektronisch archief van de Erasmus Universiteit: Sie, M.M.S.K., *Justifying Blame: Why Free Will Matters and Why it Does Not*, Amsterdam/New York: Rodopi, 2005.

[k2]Gebruikte literatuur

[t1]Aune, B. (1967), Lehrer, K. (1968), *Cans and IFS: An Exchange*, reprinted in: Watson 1982.

Austin, J.L., 'Ifs and Cans', *Proceedings of the British Academy XLII*, 1956, 109-32.

Banuazizi, A., Movahedi, S. (1975), 'Interpersonal Dynamics in a Simulated Prison: A Methodological Analysis', *American Psychologist*, 1975, 152-60.

Chisholm, R., 'Human Freedom and the Self', *The Lindley Lecture*, 1964, 3-15.

Dijksterhuis, A., *Het slimme onbewuste: Denken met gevoel*, Amsterdam: Bert Bakker, 2007.

Doris, J.M., *Lack of Character: Personality and Moral Behavior*, Cambridge: Cambridge University Press, 2002.

Frankfurt, H.G., 'Alternative Possibilities and Moral Responsibility', *Journal of Philosophy* 56, 1969.

----- 'Freedom of the Will and the Concept of a Person', *Journal of Philosophy* 68, 1971, 5-20.

Kane, R., *The Significance of Free Will*, Oxford/New York: Oxford University Press, 1996.

Lamme, V., *De vrije wil bestaat niet: Over wie er echt de baas is in het brein*, Amsterdam: Prometheus, 2010.

Murphy, S., Haidt, J. en Björklund, F., *Moral Dumbfounding: When Intuition Finds No Reason*, University of Virginia, Unpublished Manuscript, 2000.

Laar, T. van de, Voerman, S., *De vrije wil*, Rotterdam: Lemniscaat, 2011.

Nagel, T., 'Moral Luck', *Mortal Questions*, Cambridge: Cambridge University Press, 1979, pp. 24-38.

Moore, G.E., *Ethics*, London, 1911.

Pereboom, D., *Living without Free Will*, Cambridge: Cambridge University Press, 2001, p. xv en 40.

Sie, M., 'Einsteins gemoedsrust: Over morele verantwoordelijkheid en de vrije wil', in: *Wijsgerig Perspectief*, 2010.

Schnall, S., J. Haidt e.a., 'Disgust As Embodied Moral Judgment', *Personality & Social Psychology Bulletin* 34 (8), 2008, 1096-1109.

Strawson, P.F., 'Freedom and Resentment', *Proceedings of the British Academy*, vol. XLVIII, 1962, 1-25.

Schopenhauer, A., *Die Welt als Wille und Vorstellung*, Stuttgart: Philipp Reclam Jun, 1859 (repr. 1987).

Van de Laar, T., Voerman, S., *Vrije wil. Discussies over verantwoordelijkheid, zelfverwerkelijking en bewustzijn*, Lemniscaat, 2011.

Vohs, K.D., & Schooler, J.W., The Value Of Believing In Free Will: Encouraging A Belief In Determinism Increases Cheating, *Psychological Science* 19, 2008, 49-54.

Wallace R.J., *Responsibility and the Moral Sentiments*, Cambridge/London: Harvard University Press, 1994.

Watson, G., *Free Will*, Oxford: Oxford University Press, 1982.

Wegner, D.M., *The Illusion of Conscious Will*, Cambridge, Mass.: MIT Press, 2002.

Wolfe, T., *I Am Charlotte Simmons*, New York: Picador, 2004.

ⁱ Dit hoofdstuk is een uitgebreide en aangepaste versie van een artikel dat eerder verscheen in *Wijsgerig Perspectief* onder de titel 'Einsteins gemoedsrust: Over morele verantwoordelijkheid en de vrije wil', Sie, M. (2010), *Wijsgerig Perspectief*, special issue (ed. Gerben Meynen); de tweede en derde paragraaf zijn nieuw.

ⁱⁱ Ap Dijksterhuis opent zijn boek *Het slimme onbewuste* met een opvallend vergelijkbaar citaat dat ontleend aan de schrijver Tom Wolfe: 'Let's say you pick up a rock and you throw it. And in midflight you give that rock consciousness and a rational mind. [...] That little rock will think it has free will and will give you a highly rational account of why it has decided to take the route it is taking.' Tom Wolfe, *I am Charlotte Simmons*, p. 283 (geciteerd in: Dijksterhuis, 2007, p. 9).

ⁱⁱⁱ Hetzelfde zien we in de filosofische discussie over de vrije wil gebeuren met een denker als Immanuel Kant; deze wordt door compatibilistisch filosoof Christine Korsgaard een compatibilist genoemd, en door de harde determinist Ted Honderich een incompatibilist. Zie de bijdrage van Kleingeld in deze bundel om een eigen mening hierover te vormen.

^{iv} Vergelijk dit ook met de opvatting van Daniel Dennett, zie het hoofdstuk van Wouters in deze bundel.

^v Incompatibilisten en compatibilisten geloven in de onverenigbaarheid respectievelijk verenigbaarheid van de these van het determinisme enerzijds, en het bestaan van de vrije wil en rechtvaardigheid van ons praktijk van morele verantwoordelijkheid anderzijds.

^{vi} Zie bijvoorbeeld Bert Keizers benadering in deze bundel, een duidelijke optimist in Strawsons zin.

^{vii} In sommige gevallen leidt een dergelijke houding tot overduidelijke spanningen, bijvoorbeeld in het geval van terroristische aanslagen waar mensen door de gruwelijkheid van de daad ertoe verleid worden om de daders voor 'volkomen gestoord' te verslijten, terwijl zij zich tegelijkertijd koste wat kost zouden verzetten tegen een voorstel om de mensen te behandelen in een tbs-kliniek.

^{viii} Een inzicht dat ook een grote rol speelt in het werk van Dennett (Dennett 2003). Zie ook de bijdrage van Wouters in deze bundel.

^{ix} Zie bijvoorbeeld een kritisch artikel over wat het Stanford Prison Experiment precies aantoont. Onderdeel van dat artikel is een inventarisatie van voorspellingen omtrent eigen en/of andermans gedrag in een dergelijk experiment. (Banuazizi, Ali, Movahedi, Siamak 1975)

^{*} Ook de eerderbesproken Wolf en Wallace zijn het wat dit betreft niet eens met Strawson en wijzen erop dat de metafysische vraag naar de vrije wil inherent is aan onze alledaagse praktijk. Zij menen echter eveneens dat wij op grond van die praktijk bevestigend kunnen antwoorden op de vraag naar het bestaan van de vrije wil. Zij begrijpen de vrije wil in termen van ons vermogen tot reflectieve zelfcontrole.

^{xi} Harde deterministen geloven dat determinisme niet verenigbaar is met de vrije wil, dat de vrije wil noodzakelijk is om onze toeschrijvingen van morele verantwoordelijkheid te gronden, en dat het determinisme juist is.

^{xii} Pereboom is weliswaar geen 'harde determinist' maar een 'harde incompatibilist', zie Van de Laar en Voerman 2011, § 2.6.

^{xiii} Het schijnt dat Luther zelf die woorden helemaal niet uitgesproken heeft en evenmin geloofde dat hij daar stond uit 'vrije wil'. Voor de invloed van Luther op het denken over de vrije wil, zie de bijdrage van Van Ruler in deze bundel

^{xiv} Voor een kritische bespreking van Kanes ideeën hierover, zie Dennett 2003.

^{xv} Heden ten dage geloven nog weinigen in de houdbaarheid van dit klassieke compatibilisme.