

OVER HET SAMENGAAN VAN VRIJE WIL EN DETERMINISME EEN VERKENNING MET EEN OPEN EINDE

Palmyre Oomen

In deze bijdrage zal het gaan om het vrijewilconcept zoals dat is ontwikkeld in een welbepaalde stroming in het filosofische debat: namelijk de stroming die zegt dat ‘vrije wil’ en ‘determinisme’ te verenigen (‘compatibel’) zijn,¹ en daarom als ‘compatibilisme’ wordt aangeduid. Deze positie die binnen de filosofie op dit moment het meeste weerklank vindt (maar waarheid gaat niet met meerderheid van stemmen!), botst nogal met onze intuïtieve denkwijze. De eerste denkwijze die spontaan opkomt is namelijk om het ‘vrij’ van ‘vrije wil’ te verstaan als niet-vastliggend, niet-gedetermineerd, kunnen kiezen uit alternatieve mogelijkheden. En wanneer omgekeerd gemeend wordt dat een serie van gebeurtenissen gedetermineerd is, dan is onze eerste gedachte dat dit inhoudt dat er geen ruimte is voor een vrije wil. Met andere woorden, de opvatting die het meest aansluit bij onze spontane prereflexieve opvatting is dat ‘vrije wil’ en ‘determinisme’ elkaar wederzijds uitsluiten, oftewel incompatibel zijn. De denkers die deze positie innemen worden als ‘incompatibilisten’ aangeduid. Het is duidelijk dat ‘incompatibilisme’ een verzamelterm is waar (minstens) twee groepen onder vallen die een onderling tegengestelde positie innemen. De eerste groep bevestigt de waarheid van het determinisme en concludeert dat er daarom van ‘vrije wil’ geen sprake kan zijn: de door William James zo genoemde ‘harde deterministen’.² De andere groep, die daar diametraal tegenover staat, bevestigt het bestaan van een ‘vrije wil’ en concludeert dat er daarom van determinisme geen sprake kan zijn: de libertariërs.

- 1 Met ‘determinisme’ wordt hier de these bedoeld dat de toestand van de wereld op ieder willekeurig ogenblik tezamen met de natuurwetten de toestand van de wereld op ieder ander ogenblik inhoudt. Of zeer kort: ‘*Determinism ... is the thesis that there is at any instant exactly one physically possible future*’ (Peter van Inwagen, *An essay on free will*, Oxford: Clarendon Press, 1983, p.3).
- 2 William James, ‘The dilemma of determinism’, in: *Unitarian review*, September 1884, herdrukt in: *The will to believe*, Dover Publications, [1897]1956, p. 149.

De positie die in deze bijdrage centraal staat, die van het compatibilisme, neemt stelling zowel tegen de harde deterministen als tegen de libertariërs, door te stellen dat ‘vrije wil’ en ‘determinisme’ elkaar niet uitsluiten, maar integendeel te verenigen oftewel ‘compatibel’ zijn.³

Zeer vereenvoudigd zijn de posities als volgt weer te geven:

		determinisme	vrijheid
Incompatibilisme	{	Hard determinisme	+ -
		Libertarisme	- +
Compatibilisme			+ +

Er zijn dus, dit schema volgend, twee posities die de mogelijkheid van een vrije wil bevestigen, te weten het libertarisme en het compatibilisme. Op de compatibilisten rust dan de taak om het veelal als contra-intuïtief geziene standpunt dat ‘vrije wil’ samen kan gaan met determinisme uit te leggen en te verdedigen. Deze positie is het onderwerp van kritische bespreking in het voorliggende artikel.

Vergeleken met de compatibilisten lijken de libertariërs het gemakkelijk te hebben, omdat ze het meest aan lijken te sluiten bij ons spontane zelfverstaan. Maar dat dit een gemakkelijk te verdedigen positie zou zijn is desondanks schijn. De libertariërs zitten namelijk met het omgekeerde probleem: want ook indeterminisme is moeilijk samen te denken met het begrip ‘vrije wil’ en met het daaraan gelieerde begrip ‘verantwoordelijkheid’. Hoe zouden we immers een beslissing of daad die voortkomt uit een random proces als ‘*up to us*’ kunnen beschouwen?⁴ Het compatibilistische standpunt (vrije wil kan samen met determinisme) en het libertarische standpunt (vrije wil gaat samen met indeterminisme) hebben dus *allebei* een probleem om het bestaan van een vrije wil te verdedigen. We lijken hier dus in een dilemma gevangen.

Het huidige artikel betreft een uiteenzetting en kritische reflectie op de compatibilistische positie, in haar verschillende fasen:

- Het artikel begint met een uiteenzetting hoe de klassieke versie van het compatibilisme (Thomas Hobbes, David Hume) een uitweg beproefde

3 Anders dan het harde determinisme en het libertarisme, beweert het compatibilisme op zich niets over het waar zijn van determinisme en/of van de vrije wil, enkel dat die twee samen kunnen gaan.

4 Voor een bespreking van het libertarisme zie het artikel van Thomas Müller, Niels van Miltenburg en Daan Evers in deze bundel.

uit het vrijewildilemma, door ‘vrije wil’ te begrijpen als handelingsvrijheid, en door een bijzondere (namelijk conditionele en tegenfeitelijke) invulling te geven aan de notie ‘had anders kunnen handelen’.

- In de tweede paragraaf bespreek ik welke diepgaande verfijning Harry Frankfurt's compatibilistische visie biedt door een hiërarchisch model van de wil te geven. Daarmee kan hij handelingsvrijheid van wilsvrijheid onderscheiden, en de botsing tussen verschillende willens inzichtelijk maken. Uiteindelijk kan Frankfurt aannemelijk maken dat er sprake kan zijn van morele verantwoordelijkheid zelfs bij afwezigheid van alternatieve mogelijkheden.
- Susan Wolf wordt vervolgens gepresenteerd als vertegenwoordiger van een nieuwe vorm van compatibilisme dat het compatibilisme opnieuw diepgaand verandert en verfijnt door namelijk het aspect van redelijkheid in te brengen. Dankzij de Rede als een ‘kompas’ kunnen we ons oriënteren op een niet van ons afhankelijk referentiepunt. Dat verschaft ons de mogelijkheid tot vrije wil en verantwoordelijkheid.
- Ter afsluiting benoem ik als winst van de compatibilistische zoektocht dat die (tegen de schijn in) een verdieping van het begrip vrijheid en vrije wil met zich meebrengt. En ik opper een mogelijkheid hoe, in lijn met een tweestadiamodel à la William James, het denken over vrije wil en verantwoordelijkheid nog aan adequaatheid kan winnen door binnen het model van Frankfurt en Wolf een zeker *indeterminisme* toe te laten. Een visie ‘voorbij compatibilisme en incompatibilisme’ die – zo zeg ik in het naschrift – lijkt te passen bij sommige inzichten van het moderne hersenonderzoek.

De klassiek compatibilistische positie

De klassieke compatibilistische positie, de positie dus die vrije wil verenigbaar acht met determinisme, is expliciet verwoord door Thomas Hobbes (1588-1679) en door David Hume (1711-1776). Sporen van een compatibilistische opvatting zijn echter ook al lang daarvoor aan te treffen, bijvoorbeeld in de Stoa en bij Thomas van Aquino, zij het in geheel andere kaders.

Om te begrijpen hoe een compatibilist kán beredeneren dat determinisme vrijheid niet in gevaar hoeft te brengen, moeten we beter analyseren wat in deze context met ‘vrijheid’ of ‘vrije wil’ bedoeld wordt. De vraag naar determinisme en vrijheid stelt zich met name ten aanzien van ons *handelen* (vandaar ook de vele raakvlakken met de thematiek van morele verantwoordelijkheid).

Vrije wil als handelingsvrijheid

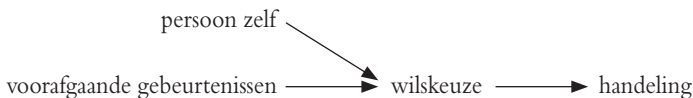
Je kunt stellen dat het een compatibilist lukt om vrije wil en determinisme als verenigbaar te zien, door van het begrip ‘vrije wil’ een beperkte definitie te hanteren, door namelijk ‘vrije wil’ te karakteriseren als het ‘ongehinderde vermogen te doen wat je wilt’. Zo schrijft Hobbes in zijn *Leviathan*: ‘Lastly, from the use of the word Free-Will, no Liberty can be inferred of the will, desire, or inclination, but the Liberty of the man; which consisteth in this, that he finds *no stop, in doing what he has the will, desire, or inclination to doe*.’⁵ Of zoals Arthur Schopenhauer (zelf geen compatibilist) het kort karakteriseert: ‘De mens kan doen wat hij wil; hij kan echter niet willen wat hij wil.’⁶

Dit wijst op twee aspecten die een rol spelen in de discussie over de vrije wil: namelijk ‘kunnen doen wat je wilt’ en ‘kunnen willen wat je wilt’. En als eerste insteek is te zeggen dat het voor de compatibilisten typerend is dat zij (in tegenstelling tot de harde deterministen en in tegenstelling tot de libertariërs) het om te kunnen spreken van ‘vrijheid’ voldoende vinden als aan het eerste aspect voldaan is (zoals getypeerd met de uitspraak van Schopenhauer).

Voor het compatibilisme bestaat vrijheid of vrije wil in het onbelemmerde vermogen van een persoon om te doen wat hij wil (omdat hij het wil). Anders gezegd, spreken van vrije wil is primair een kwalificeren van een handeling: indien gedaan ‘uit vrije wil’ is de betreffende handeling een handeling die overeenstemt met wat de persoon wil (en dat niet als coincidentie, maar op grond van die wil).

Hier komt dus het eigene van het (klassieke) compatibilisme naar voren. Voor vrijheid, zo is de gedachtegang, is het voldoende dat je willen zich vrijelijk kan doorzetten, dat je willen in het ten uitvoer komen onbelemmerd is, dat het handelen overeenstemt met je willen. Dat daarbij die wil (als determinisme waar is) zélf gedetermineerd is, namelijk het product is van voorafgaande gebeurtenissen en natuurwetten, en dat je dus over die wil zélf geen baas bent, doet volgens de compatibilist niet ter zake.

We kunnen de twee aspecten ‘kunnen doen wat je wilt’ en ‘kunnen willen wat je wilt’ verhelderen aan de hand van het volgende schema:



5 Thomas Hobbes, *Leviathan*, 1651, Ch. 21 (cursivering toegevoegd).

6 ‘Der Mensch kann tun was er will; er kann aber nicht wollen was er will’, uit: Arthur Schopenhauer, *Über die Freiheit des menschlichen Willens*, 1839.

Hierin is verbeeld dat de handeling (wel of niet) de vertaling is van de voorafgaande wilskeuze, en dat die wilskeuze zelf (wel of niet) het gevolg is van voorafgaande oorzaken (zoals genetische aanleg, hersenontwikkeling, opvoeding, reclame et cetera), dan wel ook of helemaal bepaald wordt door de persoon in kwestie zélf. Dit geeft op twee plaatsen aanleiding tot vragen die gerelateerd zijn aan de kwestie van de vrije wil:

Allereerst de meest gepraktiseerde vraag, betrekking hebbend op de tweede pijl, de vraag namelijk naar de correspondentie van de handeling met de wil ('wil' in de zin van de optie waarvoor gekozen is die te effectueren). Doe je wat je wilde doen? Was met andere woorden de wil vrij om zich te effectueren? Dit vrij zijn van de wil om zich te effectueren, zodat de handeling 'uit vrije wil' gebeurt, wordt in het Engels aangeduid met de term '*voluntariness*' (naar het Latijnse '*voluntarius*') en wordt in het Nederlands vaak weergegeven met 'vrijwilligheid', en lelijker maar beter met 'willijkheid'.⁷ 'Vrij' in deze zin van voluntair/vrijwillig/willijk betekent derhalve: de wil is in haar effectuering vrij, niet gedwarsboemd, onbelemmerd, ongedwongen; of vanuit de handeling gezien: de handeling is in overeenstemming met de wil (niet bij coincidentie, maar vanwege de wil). Het volgende citaat van David Hume verwoordt dit exemplarisch:

'Onder vrijheid kunnen we dus alleen verstaan *de macht om een daad te stellen of niet te stellen overeenkomstig de beslissing van de wil*; dat wil zeggen, als wij kiezen inactief te blijven, doen wij dat, en als wij kiezen tot actie over te gaan, dan kan dat ook. Welnu, deze hypothetische vrijheid bezit naar algemeen aanvoelen iedereen, die geen gevangene in ketens is.⁸

Er is echter ook een vraag te stellen bij de eerste pijl, een vraag naar hoe deze wilskeuze tot stand is gekomen. Is de persoon in kwestie ten aanzien van die wilskeuze vrij, in de zin van daartoe niet gedetermineerd door voorafgaande factoren, doch daar zelf onafhankelijk baas over? Zo ja, dan wordt dat aangeduid als '*oorspronkelijkheid*'. 'Vrij' in deze zin van oorspronkelijk

7 Het woorddeel 'vrij' in 'vrijwillig' kan gemakkelijk het misverstand oproepen dat de actor in kwestie in dezelfde situatie ook anders had kunnen handelen, doch dát wordt met '*voluntarius*' niet bedoeld. '*Voluntarius*' geeft enkel aan dat de handeling niet het resultaat is van dwang, maar in vrije overeenstemming is met de wil (en dus 'willijk' is, zoals 'natuurlijk' betekent in overeenstemming met de natuur).

8 David Hume, *An enquiry concerning human understanding*, [1748], hier geciteerd naar de vertaling van J. Kuin: *Het menselijk inzicht. Onderzoek naar het denken van de mens*, 2e herziene druk, Amsterdam: Boom, 2002, p.132 (cursivering origineel).

betekent derhalve: onze wil wordt door onszelf bepaald, is *'up to us'* (zoals Aristoteles dat aangaf met *'eph' hènin'*), is *'in our control'*. De wil is – in geval van oorspronkelijkheid – derhalve het resultaat van zelfbepaling en niet volledig gedetermineerd door voorafgaande gebeurtenissen en natuurwetten (noch een louter toevalseffect).

Hoe zijn beide aspecten van de notie 'vrije wil' nu verbonden met de determinismekwestie? Alle drie de genoemde richtingen (hard determinisme, libertarisme en compatibilisme) zijn het erover eens dat vrijheid van de wil als 'oorspronkelijkheid' en metafysisch determinisme elkaar wederzijds uitsluiten. Maar ze verbinden daar totaal verschillende conclusies aan.

Een libertarist redeneert: er is maar sprake van 'vrije wil' als er sprake is van willijkheid én bovenal van oorspronkelijkheid. Welnu, die oorspronkelijkheid zou uitgesloten zijn als determinisme waar was. Dus als 'vrije wil' bestaat (en dat affirmeert een libertarist) kan determinisme niet waar zijn. Een harde determinist erkent eveneens dat 'vrije wil' zou inhouden dat er sprake is van willijkheid én van oorspronkelijkheid. Derhalve als determinisme waar is (en dat affirmeert een harde determinist) kan oorspronkelijkheid en dus vrije wil niet waar zijn.

De compatibilist wijkt van beide andere richtingen af, door als enige te stellen dat oorspronkelijkheid (wat inderdaad onmogelijk is bij determinisme) níet vereist is voor de kwalificatie 'vrije wil'. Een compatibilist schrappt dus de eis van 'oorspronkelijkheid', en zet vol in op het aspect van 'willijkheid', dus op het aspect dat de wil zich vrijelijk als wil kan gedragen en zich onbelemmerd en zonder dwang van buitenaf kan vertalen in de gewilde handeling.⁹

De compatibilistische claim om van 'vrije wil' te kunnen spreken ook wanneer alles (inclusief ons handelen en de inhoud van ons willen) gedetermineerd is, is derhalve gebaseerd op twee aannames: (1) (Als determinisme waar is, is daarmee inderdaad oorspronkelijkheid onmogelijk, maar dat geeft niet, want:) voor 'vrije wil' is oorspronkelijkheid niet nodig. Van de twee genoemde aspecten volstaat 'willijkheid'; en (2) Willijkheid wordt op geen enkele manier gehinderd door determinisme.

9 Het is interessant om hier de zeer eigenzinnige positie van Immanuel Kant naast te zetten. Ook bij Kant zijn determinisme en vrije wil op een bepaalde manier compatibel, doch waar de andere compatibilisten dat voor elkaar krijgen door de omvang van het begrip vrijheid te verkleinen (tot enkel willijkheid/*voluntariness*), verkleint Kant het domein waarop het begrip determinisme van toepassing is (namelijk tot enkel de fenomenale wereld). Zie ook de uiteenzetting over Kant door Donald Loose in deze bundel.

Een deterministische uitleg van de notie 'kon anders gehandeld hebben'

Om de plausibiliteit van bovengenoemde aannames te beproeven, is het goed te beginnen met een analyse van wat we bedoelen wanneer we spreken van 'vrijelijk handelen' of van 'handelen uit vrije wil'.

Wanneer we van iemand zeggen dat hij handelde 'uit vrije wil' (en hem daarom verantwoordelijk houden voor de gevolgen) dan bedoelen we daarmee dat de persoon in kwestie (P) ook *anders had kunnen handelen*. Op dat laatste is immers ons toezwaaien van lof en zeker ook ons uiten van verwijten gebaseerd.¹⁰ Schematisch weergegeven gaan we dus uit van de equivalentie:

P handelt vrijelijk \Leftrightarrow P kon anders gehandeld hebben

Nu suggereert de uitdrukking 'kon anders gehandeld hebben' (in het Engels standaard weergegeven met: '*could have done otherwise*') dat er alternatieve handelingsmogelijkheden open liggen voor iemand. In die optie is binnen het *libertarisme* voorzien door de notie 'oorspronkelijkheid'. Immers met oorspronkelijkheid wordt aangeduid dat er iets nieuws en oorspronkelijks ontstaat, iets dat niet voortkomt uit voorgegeven feiten. Binnen het *libertarisme* is er dus ruim plaats voor alternatieve mogelijkheden.

Dergelijke alternatieve mogelijkheden kunnen echter geen plaats hebben binnen een deterministisch beeld. Vandaar dat een *harde determinist* kort en krachtig redeneert: determinisme betekent dat P's wil en P's handelen gedetermineerd zijn. Er zijn geen alternatieve mogelijkheden, dus kon P niet anders gewild of anders gehandeld hebben. Dus is P in zijn willen en handelen onvrij.

De *compatibilist* daarentegen meent dat de veronderstelling van determinisme, ondanks de schijn van het tegendeel, wél de mogelijkheid laat om 'vrijelijk handelen' in de zin van 'kon anders gehandeld hebben' te denken. De 'vondst' zit erin dat 'kon anders gehandeld hebben' gelijkgesteld wordt aan een conditionele bewering (een 'als ... dan' bewering):

P kon anders gehandeld hebben \Leftrightarrow
als P gewild had om anders te handelen, dan had P anders gehandeld

10 Er is hier te wijzen op een (overigens niet onbetwiste) asymmetrie. Wanneer we iemand schuldig achten aan een daad, dan achten we die persoon alleen verantwoordelijk als hij ook anders had kunnen handelen. Wanneer we iemand prijzen om een daad, kan het voorkomen dat we tevens ervan overtuigd zijn dat die persoon, gezien zijn goede karakter, niet anders had kunnen handelen. Zie hierover bijvoorbeeld: S. Wolf, *Freedom within reason*, New York/Oxford: Oxford University Press, 1990, p. 79-81.

De compatibilist redeneert namelijk: natuurlijk kon P, als determinisme waar is, in de feitelijke situatie niet anders gewild hebben (wat P wil is dan immers gedetermineerd), maar dit neemt de mogelijkheid niet weg om in waarheid te beweren dat *als* P anders had gewild, hij anders gehandeld zou hebben. (Omdat, als determinisme waar is, hetgeen in deze bewering volgt op *als* niet waar kán zijn, is deze conditionele bewering een zogeheten ‘tegenfeitelijke conditionele bewering’.)

Op deze manier begrepen kan de compatibilist ‘ja’ zeggen tegen de bewering ‘Als P anders gewild had, had P anders gehandeld’ en daarmee ondanks het (aangenomen) determinisme ‘ja’ zeggen tegen ‘P kon anders gehandeld hebben’, en daarmee dus tegen de uitspraak ‘P handelde vrijelijk’. Cruciaal in dezen is dus dat het determinisme een correlatie tussen wil en handeling volstrekt niet in de weg staat. Determinisme laat het willijkheidsaspect van ‘vrije wil’ dus toe, en daarenboven ook de voor het vrijewilconcept zo belangrijke ‘alternatieve mogelijkheden’, namelijk in die voor het compatibilisme kenmerkende tegenfeitelijke, conditionele zin.¹¹

William James, die overigens zelf een niet-deterministische visie voorstond, uitte zijn ongenoegen over dit in zijn ogen verdonkeremanen van de consequenties van het determinisme door de compatibilisten als ‘*soft* deterministen’ te benoemen.¹²

Enkele kanttekeningen

De vraag die we ons hier moeten stellen is hoe plausibel een dergelijke opvatting van ‘vrije wil’ is wanneer de notie ‘had anders kunnen handelen’ alleen conditioneel waar is in een (niet-bestaande) andere situatie, en bovendien het aspect van oorspronkelijkheid sowieso niet meer aan de orde komt? Met deze vraag komt ook de bovengenoemde eerste compatibilistische aanname aan de orde.

In eerste instantie klinkt de conditionele interpretatie van ‘*could have done otherwise*’ die bovendien tegenfeitelijk is omdat de wilsinhoud volledig gedetermineerd is (vanwege de afwezigheid van de mogelijkheid van oorspronkelijkheid) misschien als een enorme goocheltruc. Heb je even niet goed

11 De alternatieve mogelijkheden hebben nu een andere status. Niet langer als verschillende mogelijke gebeurtenissen volgend op één stand van zaken, maar als verschillende mogelijkheden gerelateerd aan *verschillende* standen van zaken. Dus *niet*: als X het geval is, zijn p_1 en p_2 en p_3 mogelijk, maar: als X_1 het geval is, is p_1 mogelijk, als X_2 het geval is p_2 , etc.

12 ‘Old-fashioned determinism was what we may call *hard* determinism. It did not shrink from such words as fatality, bondage of the will, necessitation, and the like. Nowadays, we have a *soft* determinism which abhors harsh words, and, repudiating fatality, necessity, and even predetermination, says its real name is freedom’ (James, ‘The dilemma of determinism’, p. 149).

opgelet ... en verdwenen is de knoop uit het touw! Toch zijn er gemakkelijk voorbeelden aan te voeren die deze in eerste instantie contra-intuïtieve stellingname van het klassieke compatibilisme in de praktijk goed invoelbaar maken.

Neem het geval dat je op een middag een bepaalde jazz-cd wilt kopen. En stel dat je niet door uitwendige of inwendige factoren belemmerd wordt deze wil in een handeling te laten effectueren. Uitwendig belemmerende factoren zouden kunnen zijn: een verbod van je ouders om dergelijke muziek te kopen, geldgebrek, de cd is uitverkocht, de winkel is die dag gesloten, et cetera. Een voorbeeld van een inwendig belemmerende factor is bijvoorbeeld besluiteloosheid, geen zin om door de regen te fietsen, bang met je muzieksmaak buiten de boot van je vriendengroep te vallen, enzovoort. Maar stel, al deze belemmeringen zijn er niet, dan ben je (wat betreft deze handeling) dus vrij. Op de vraag waarom je nu juist zoveel geeft om deze cd, of om dit type muziek, is ongetwijfeld een antwoord te geven: je hoorde de muziek en dat riep de behoefte op die cd te kopen, je ouders hebben je de liefde voor deze muziek met de paplepel ingegeven, je wilt de blits maken tegenover je vrienden, je wilt hem gebruiken op een feestje, enzovoort. Maar al die achterliggende oorzaken waardoor je wilt wat je wilt, zijn in de compatibilistische visie niet ter zake voor het oordeel dat waar deze belemmeringen er níet zijn je derhalve vrij bent, vrij bent namelijk in het effectueren van die (in zichzelf gedetermineerde, door de genoemde achterliggende oorzaken bepaalde) wil. Kunnen doen wat je wilt, dat is het springende punt. En bij een dergelijk voorbeeld voelt dat allesbehalve als een goocheltruc, doch zeer in overeenstemming met de dagelijkse taalpraktijk.

Er zijn zo zeer veel overtuigende voorbeelden te construeren, maar eigen aan de kritische reflectie is dat er tegenvoorbeelden worden verzonnen. Wat bijvoorbeeld als je wil om die cd te kopen wel als je eigen wil voelt, doch het objectief is vast te stellen dat die wil het effect is van zeer effectieve reclame. Het verhaal van de *hidden persuaders*: een geur in de winkel die je kooplustig maakt; allerlei advertentie- en mediatrucs die op verborgen wijze je gedachten en wensen beïnvloeden. Op zulk soort onbewuste beïnvloedingen wijst Victor Lamme in zijn boek *De vrije wil bestaat niet*.¹³ In hoeverre is je zo beïnvloede en gemanipuleerde willen dan jouw willen? Dat zal de compatibilist zich toch moeten afvragen.

Het tekort van de klassiek compatibilistische visie wordt nog duidelijker wanneer we als voorbeeld een drugsverslaafde nemen. Stel de verslaafde wil drugs gebruiken en die daartoe kopen. Hij gaat naar buiten en vindt op de hoek van de straat een dealer. Hij koopt drugs en gebruikt ze. De klassiek

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

compatibilistische zienswijze volgend leidt dit tot het oordeel dat de persoon in kwestie uit vrije wil handelt. Zijn handelen is immers in overeenstemming met zijn wil. Zijn wil is in deze onbelemmerd en vrij ('voluntair'). Grote vraag is dan: waarom noemen we zo iemand dan toch 'verslaafd'? Het ontbreekt het klassieke compatibilisme aan denkgereedschap om hier een antwoord op te formuleren.

Een eerste inzicht in wat er ontbreekt, is af te leiden uit de observatie dat we in zo'n geval iets zeggen als 'ja, hij ging inderdaad op zoek naar deze drugs en wilde ze maar al te graag hebben. Die drang is sterker dan hijzelf: hij kan er geen weerstand aan bieden. Als hij nuchter is, dan is hij moedeloos hoe ooit van zijn verslaving af te komen. Hij gebruikt wel, maar eigenlijk wil hij dat helemaal niet.' Wat hier gebeurt is interessant. We maken intuïtief verschil tussen iemands directe willen (het willen dat zich uitdrukt in de handeling) en iemands 'eigenlijke willen' dat daar zeer van onderscheiden kan zijn. Zo kan een vader of moeder die zijn/haar kind een forse klap heeft gegeven, naderhand zeggen: 'Ja, ik gaf mijn kind een klap, want ik was boos (het was dus niet per ongeluk, maar "gewild"). Maar toch: ik wil mijn kind helemaal niet slaan.' Kortom we maken gewag van meerdere 'willen' en van een onderscheid daarin.

Het is derhalve volstrekt ontoereikend om zoals het klassieke compatibilisme doet over willen te spreken in het enkelvoud als 'de wil'. En van 'vrije wil' als de handeling overeenstemt met deze enkelvoudige wil. We willen vele, deels ook onderling onverenigbare, zaken en er is bovendien een geïmpliciteerd laagheidsniveau van het willen. En als we het hebben over het al dan niet vrij zijn van de wil, dan moeten we goed bedenken over welke van de meerdere willen we het dan hebben. En voor het compatibilisme roept dat derhalve de vraag op hoe die meerdere willen en de eventuele vrijheid daarvan samen kunnen gaan met determinisme.

Het hiërarchische model van Harry Frankfurt

In het licht van de tekorten van het klassieke compatibilisme is de zienswijze interessant die Harry Frankfurt (*1929) heeft ontwikkeld en waarin een hiërarchie van wilsniveaus wordt gethematiseerd. Tevens hoeft hij volgens de daarin ontwikkelde zienswijze voor morele verantwoordelijkheid geen enkel beroep te doen op de notie 'had anders kunnen handelen', ook niet in de tegenfeitelijke conditionele interpretatie daarvan.¹⁴ We zullen zijn zienswijze nu onder de loep nemen.

14 H. Frankfurt, 'Alternate possibilities and moral responsibility', in: *The Journal of Philosophy*, 66, 1969, p. 829-839; H. Frankfurt, 'Freedom of the will and the concept of a person', in: *The Journal of Philosophy*, 68, 1971, p. 5-20.

Eerste-ordeverlangens en tweede-ordeverlangens

Frankfurt maakt onderscheid tussen twee soorten verlangens¹⁵: verlangens die betrekking hebben op handelingen ('*first order desires*'), en verlangens die betrekking hebben op verlangens ('*second order desires*'). Vaak zal iemand meerdere eerste-ordeverlangens hebben, verlangens dus om iets te doen (bijvoorbeeld studeren, de krant lezen, naar het café gaan). Een daarvan beweegt de persoon in kwestie feitelijk tot actie (bijvoorbeeld de krant lezen). Dát eerste-ordeverlangen dat de persoon in kwestie feitelijk tot actie beweegt, benoemt Frankfurt als de '*will*' (om duidelijkheidsredenen door mij ook weergegeven als 'de effectieve wil'). Eigen aan de menselijke persoon is nu dat die ook een verlangen heeft welke van zijn vele en deels tegenstrijdige eerste-ordeverlangens zijn effectieve wil is. Dit speciale tweede-ordeverlangen noemt Frankfurt de '*second order volition*'.¹⁶ Dus in het bovenstaande voorbeeld kan het goed zijn dat het eerste-ordeverlangen dat zich effectueert 'krant lezen' is, terwijl de persoon 'eigenlijk' wilde dat zijn verlangen te studeren hem tot handelen zou bewegen (zijn effectieve wil was). De tweede-ordevolitie komt in onze omgangstaal vaak tot uitdrukking in het gebruik van deze toevoeging 'eigenlijk'. Een mens kan dus ten aanzien van directe eerste-ordeverlangens een innerlijke distantie aannemen, door van een van hen te verlangen dat die zich in handelen vertaalt. En precies in die mogelijkheid van innerlijke distantie schuilt het wezenlijke van 'persoon zijn', aldus Frankfurt.

Vrijheid van handelen – vrijheid van willen – handelen 'uit eigen vrije wil'

Er is dus volgens deze visie van Frankfurt een hiërarchische gelaagdheid van de wil, en dat maakt het mogelijk om genuanceerder te spreken over (on)vrijheid. En het eerste waartoe Frankfurts model mogelijkheid biedt, en wat Frankfurt ook nadrukkelijk bepleit, is een helder onderscheid maken tussen 'vrijheid van handelen' en 'vrijheid van willen'.

Vrijheid van handelen betekent volgens Frankfurt dat je vrij bent om te handelen in overeenstemming met je verlangen ('the freedom to do what one wants to do'¹⁷). In het bovengenoemde voorbeeld, waarin P verlangt de

15 In het Nederlands zou je hier eerder spreken van 'willen'. Om echter de structuur van Frankfurts analyse goed te kunnen weergeven vertaal ik zijn werkwoord 'to want' hier steeds met 'verlangen' ook waar dat in het Nederlands wat vreemd klinkt, om zo dat ene specifieke verlangen dat tot handelen beweegt, met Frankfurt, terminologisch te kunnen onderscheiden als 'wil'.

16 Ik gebruik verder het woord 'volitie' in plaats van simpelweg 'wil' als weergave van Frankfurts '*volition*', om zo mét Frankfurt met deze term dat specifieke tweede-ordeverlangen te kunnen onderscheiden dat betrekking heeft op welk eerste-ordeverlangen de effectieve wil is.

17 Frankfurt, 'Freedom of the will', p. 15.

krant te lezen, is P vrij in zijn handelen als hij dat verlangen kan effectueren, en daarin niet belemmerd wordt – bijvoorbeeld doordat die dag de krant niet bezorgd is – of gedwongen wordt om iets anders te doen. Dit lijkt misschien het gewone compatibilistische verhaal, doch waar het klassieke compatibilisme in dit verband spreekt van ‘vrije wil’, verzet Frankfurt zich daartegen en benoemt dit expliciet als ‘vrijheid van handelen’.

Door de notie ‘vrijheid van de wil’ daarvan te onderscheiden en als een aparte notie in de beschouwing te betrekken is Frankfurt in staat inzicht te verschaffen waarom we een verslaafde ‘verslaafd’ noemen. Naar analogie van de vrijheid van handelen, definieert Frankfurt de *vrijheid van willen* van iemand als: ‘he is free to will what he wants to will, or to have the will he wants.’¹⁸ Dit betekent derhalve dat P vrijheid van willen heeft wanneer hij zijn (effectieve) wil in overeenstemming kan brengen met zijn tweede-ordevolitic, dus wanneer hij vrij is om te willen (eerste orde) wat hij verlangt te willen (tweede orde). Zo geredeneerd heeft de persoon die eigenlijk wil studeren, maar feitelijk de krant gaat lezen (op dat moment) géén vrijheid van de wil. Hij weet zijn effectieve wil immers niet in overeenstemming te brengen met wat hij eigenlijk wil.

Hoe is hiermee het geval van drugsverslaving te analyseren? Frankfurt onderscheidt om zijn inzichten helder voor het voetlicht te kunnen brengen drie soorten van verslaafden: als eerste de verslaafde die überhaupt geen tweede-ordevolities heeft, die dus – zoals een dier of een klein kind – wel verlangens maar geen metaverlangens heeft (*‘the wanton addict’*); als tweede de verslaafde die eigenlijk geen drugs wil gebruiken (maar ze natuurlijk wel gebruikt als hij eraan kan komen, want hij is tenslotte verslaafd) (*‘the unwilling addict’*); en als derde de verslaafde die het prima vindt om drugs te gebruiken (en dat natuurlijk ook doet als hij die kan bemachtigen, want of hij het nu wel of niet prima vindt, hij is verslaafd en dus gebruikt hij) (*‘the willing addict’*).

Omdat de *wanton addict* geen tweede-ordevolitic heeft, speelt de vraag naar de vrijheid van de wil bij deze verslaafde überhaupt niet. Er kán simpelweg geen sprake zijn van vrijheid van de wil (en zelfs niet van ‘persoon zijn’). Bij de *unwilling addict* is er wel duidelijk sprake van een tweede-ordevolitic, maar de persoon in kwestie is vanwege de afhankelijkheid van drugs niet in staat is om zijn effectieve wil in overeenstemming te brengen met die tweede-ordevolitic. En dus is de wil van deze persoon (in dit opzicht) niet vrij. Anders dan het klassieke compatibilisme biedt Frankfurts hiërarchisch model hier dus inzicht waarom we (ook al strookt zijn handelen met zijn effectieve wil) bij deze persoon van verslaving spreken.

18 Frankfurt, ‘Freedom of the will’, p. 15.

Een verdere verfijning waartoe Frankfurts model in staat is, blijkt uit zijn analyse van de *willing addict*. Ook deze heeft een tweede-ordevolentie, en is dus een ‘persoon’. In dit geval is echter zijn effectieve wil (om drugs te gebruiken) gelijk aan de wil die hij verlangt te hebben. Maar desondanks concludeert Frankfurt dat ook deze verslaafde (in dit opzicht) géén vrijheid van willen heeft. Waarom dan niet? Frankfurt redeneert aldus: voor een vrij zijn van de wil is het niet voldoende dat de eerste-ordevil feitelijk (bijvoorbeeld door coïncidentie) overeenstemt met de tweede-ordevolentie. Nee, voor vrijheid van de wil moet de persoon in kwestie in staat zijn om zijn eerste-ordevil *in overeenstemming te brengen* met zijn tweede-ordevolentie (‘securing the conformity of his will to his second-order volitions’¹⁹). En dat is nu juist wat de *willing addict* niet kan. Immers, wát zijn tweede-ordevolentie ook is, zijn eerste-ordevil is (zolang hij verslaafd is) gericht op drugs gebruiken. Dus indien deze verslaafde een week later, dankzij bijvoorbeeld een voorlichtingscampagne, een andere tweede-ordevolentie ontwikkeld zou hebben, en zo ‘eigenlijk’ geen drugs meer zou willen gebruiken, dan zal blijken (zolang hij verslaafd is) dat hij zijn effectieve wil niet in overeenstemming kan brengen met deze tweede-ordevolentie. En dus dat zijn wil in deze niet vrij is.

Het is zaak om goed te zien welke subtiele (ook terminologisch subtiele) onderscheidingen Frankfurt hier aanbrengt. Ondanks het feit dat zo’n *willing addict* ten aanzien van zijn drugsgebruik geen vrije wil heeft (want, zoals aangegeven, de verslaafde is niet in staat om zijn eerste-ordevil in overeenstemming te brengen met zijn tweede-ordevolentie), staat dat volgens Frankfurt morele verantwoordelijkheid voor zijn drugsgebruik niet in de weg. Voor morele verantwoordelijkheid moet namelijk volgens hem aan twee vereisten zijn voldaan. De eerste is dat er sprake is van vrijheid van handelen (‘he did it freely’). De tweede vereiste is *niet* dat de persoon in kwestie zijn eerste-ordevil kan aanpassen aan zijn tweede-ordevolentie (en dus een vrije wil heeft), doch *enkel* dat de effectieve wil feitelijk overeenstemt met de tweede-ordevolentie, dus een door de persoon gewilde wil is, zodat van de corresponderende handeling geldt dat die is gedaan ‘*of his own free will*’.²⁰ En dat is hier het geval: de *willing addict* doet immers de handeling vanuit zijn zelf gewilde wil. De voor veel verwarring zorgende uitdrukking ‘*of one’s own free will*’ geeft dus enkel aan dat de uitvoerende

19 Frankfurt, ‘Freedom of the will’, p.15. Of een persoon daartoe in staat is, kan best op zichzelf genetisch of anderszins bepaald zijn. Ook het vermogen de wil in overeenstemming te brengen met de tweede-ordevolentie is dus niet in tegenpraak met determinisme (p. 20).

20 Frankfurt, ‘Freedom of the will’, p. 19. Zie ook ‘Discussion with Harry G. Frankfurt’, in: *Ethical Perspectives* 5/1, 1998, 32–33.

wil de door die persoon gewenste wil is, *niet* dat die persoon baas is over die uitvoerende wil in de zin dat hij die actief in overeenstemming kan brengen met zijn tweede-ordevolting, en dus *niet* dat hij een 'vrije wil' heeft in dezen.²¹

Enkele kanttekeningen

De grote verbetering die Frankfurts model betekent ten opzichte van het klassieke compatibilisme, zit 'm in die gelaagdheid en hiërarchische structuur van het willen. Daardoor kan hij allereerst een helder onderscheid maken tussen vrijheid van handelen en vrijheid van willen, die onafhankelijk van elkaar het geval kunnen zijn. Daardoor kan ook de innerlijke spanning tussen de uitvoerende wil (de eerste-ordevil) en de metawil (wat je verlangt te willen of 'eigenlijk wilt' – de tweede-ordevolting) in beeld gebracht worden, wat de mogelijkheid biedt om verslaving als verslaving te begrijpen.

Geconfronteerd met dit genuanceerde hiërarchische model van Frankfurt, dringt zich wellicht een vraag op: waarop stoelt in zijn visie die tweede-ordevolting, waardoor wordt de inhoud van wat de persoon in kwestie verlangt te willen bepaald? Wat dat betreft staat Frankfurt stevig in de compatibilistische traditie. Mogelijkerwijs, zo zal hij zeggen, is die tweede-ordevolting geheel gedetermineerd, dat wil zeggen, afhankelijk van iemands omstandigheden (genetisch, neurologisch, milieu, persoonlijke omstandigheden et cetera.). Maar, zegt Frankfurt, dát is geen enkel probleem. De vraag waardoor de tweede-ordevolting is bepaald, is namelijk niet interessant. Het is immers wat jij verlangt te willen. Je identificeert je ermee. Hij spreekt van '*decisive identification*'.²² Dit is wat je werkelijk zélf wilt willen. Deze identificatie is een stellig en afdoend commitment, voortkomend uit een 'wilsnoodzakelijkheid'.²³ Dit commitment maakt je 'diepe, werkelijke zelf' uit. Het weersprekt dus niet de autonomie van de persoon, maar constitueert die juist.

Toch zit er precies hier ook een zwak punt in Harry Frankfurts model, dat ik naar voren breng als overgang naar het model van Susan Wolf. Volgens Frankfurts model is iemand vrij qua willen als hij zijn uitvoerende wil in overeenstemming kan brengen met zijn tweede-ordevolting, dat is, met wat hij eigenlijk zelf wil, met zijn 'waarden'. Maar hoewel dit tweede-ordevelangen voor de persoon in kwestie richtsnoer wordt om zijn uitvoerende wil naar te richten omdat hij zich daarmee identificeert, is er geen argument

21 Frankfurt, 'Freedom of the will', p. 19-20.

22 Frankfurt, 'Freedom of the will', p. 16.

23 H. Frankfurt, *The importance of what we care about*, Cambridge: Cambridge University Press, 1988, p. 86.

binnen Frankfurts compatibilistische model waarom dit specifieke verlangen niet evenzeer het resultaat kan zijn van manipulatie of verslaving als ieder ander verlangen. Spreken van een ‘*decisive identification*’ neemt deze mogelijkheid niet weg, want wat voorkomt dat die beslissende identificatie zelf niet het gevolg is van manipulatie of verslaving? Hoe zou binnen Frankfurts raamwerk überhaupt een verschil te denken zijn tussen een ‘normaal’ door de voorgeschiedenis gedetermineerde tweede-ordevolitie, en een door drugs of manipulatie gedetermineerde tweede-ordevolitie? Maar als dat verschil zich niet laat denken dan moet iemand die tot in zijn tweede-ordewil gemanipuleerd is desondanks behandeld worden als de *willing addict* en dus moreel verantwoordelijk worden gehouden.²⁴ Als dit een ongewenste uitkomst is, en dat is het mijns inziens, dan moet er echter gezocht worden naar een uitbreiding van Frankfurts theorie, dusdanig dat daarmee de ene tweede-ordevolitie kwalitatief onderscheiden kan worden van de andere. In dit licht is de *Reason View* te zien zoals die door onder meer Susan Wolf is ontwikkeld.

De Reason View van Susan Wolf: referentie aan het Ware en het Goede

Gehoorgeven aan de Rede nodig voor vrijheid en verantwoordelijkheid

Susan Wolf (*1952) bepleit in haar *Freedom within reason* (1990) om het hiërarchische compatibilistische model aan te vullen en dusdoende te modificeren. Meer dan Frankfurt is zij in haar analyse van ‘vrije wil’ primair gericht op de kwestie van morele verantwoordelijkheid, zoals uit het vervolg zal blijken. Ze verstaat

‘free will to be that relation to one’s will which is necessary in order for one’s actions (as well as one’s character and life insofar as they are governable by one’s will) to be “up to oneself” in the way that is necessary for responsibility’.²⁵

Wat Frankfurt benoemt als ‘tweede-ordevolities’ benoemt Wolf in navolging van Gary Watson als de ‘waarden’ volgens welke iemand wil handelen. En ze vat de compatibilistische visie à la Frankfurt, die door haar *Real Self View* wordt genoemd, als volgt samen: ‘According to the Real SelfView, an

24 Zie voor dit bezwaar: M. McKenna, ‘Compatibilism’, in: *Stanford Encyclopedia of Philosophy* (versie Oct. 5, 2009) onder 5.3.3 en 5.3.4.

25 S. Wolf, *Freedom within reason*, New York/Oxford: Oxford University Press, 1990, p. 3-4.

individual is responsible if and only if she is able to form her actions on the basis of her values.²⁶ Het kenmerkende voor haar *Reason View* is dat ze dit overneemt, maar daar een extra eis aan toevoegt:

‘The Reason View insists that responsibility requires something more. According to the Reason View, an individual is responsible if and only if she is able to form her actions on the basis of her values *and* she is able to form her values on the basis of what is True and Good.’²⁷

Het vermogen om onze waarden al delibererend en reflecterend af te stemmen op wat Waar en Goed is, benoemt ze als ‘Rede’. Er is in deze visie maar sprake van vrije wil en verantwoordelijkheid als we gehoor geven aan de Rede, als we ‘*reasons-responsive*’ zijn, of in meer alledaagse taal: als we onze daden doen *om de goede redenen*.

Tussen Frankfurtiaans compatibilisme en libertarisch-incompatibilisme

Door deze stellingname neemt Wolf met haar *Reason View* een interessante positie in tussen aan de ene kant compatibilisme à la Frankfurt (*Real Self View*) en aan de andere kant libertarisch-incompatibilisme (door haar aangeduid als *Autonomy View*).²⁸ Aanhangers van een compatibilisme à la Frankfurt stellen dat er voor morele verantwoordelijkheid niet méér nodig is dan dat de persoon in kwestie de vrijheid en het vermogen heeft om te handelen in overeenstemming met zijn tweede-ordevolentie, met zijn ‘*real self*’. Aanhangers van de libertarische visie leveren hier de volgende, volgens Wolf terechte, kritiek op: de persoon in kwestie mag zijn eigen waarden dan wel normerend gebruiken, maar aangezien die waarden evenzeer het onvermijdelijke gevolg zijn van voorafgaande zaken, is – hoe complex ook vervat in een hiërarchie van lagen – er toch uiteindelijk niet méér dan een causale input uit het verleden en een daardoor bepaalde output, met nergens een plaats waar de voorgegeven invloed doorbroken wordt, genormeerd wordt, onder kritiek gesteld wordt. Doch om te gelden als uiting van vrije wil van de persoon in kwestie (zodat hij ervoor verantwoordelijk is) moet het toch een handeling

26 Wolf, *Freedom*, p. 75.

27 Wolf, *Freedom*, p. 75 (cursivering origineel).

28 Hoewel ik haar analyses van dit en andere modellen zeer interessant vind, ben ik met haar naamgevingen van de verschillende modellen niet gelukkig. Mij dunkt, dat Frankfurt met zijn ‘*Real Self View*’ bij uitstek een visie presenteert waarin vrijheid van de wil betekent dat het diepe, werkelijke zelf zichzelf de wil oplegt (‘securing the conformity of his will to his second-order volitions’, p. 15) en dus als ‘*auto-nomous*’ kan worden gekwalificeerd. Met andere woorden, om van autonomie te kunnen spreken hoeft het ‘zelf’ geen ‘ultieme bron’ te zijn (zoals in het libertarisme, en bij Kant, en zoals Wolf het begrip gebruikt).

zijn die ‘*up to him*’ is. Het incompatibilistische libertarisme poogt daarin te voorzien, door de handeling uiteindelijk geheel en al neer te laten komen op de ‘autonome’ beslissing van de persoon in kwestie, die in die keuze uiteindelijk van niets afhankelijk is, door niets uitwendig bepaald wordt, en dus de ultieme bron ervan is. Hoewel Wolf deze poging van het incompatibilisme om de handeling niet het blote gevolg te laten zijn van alle voorafgaande factoren waardeert als een terechte kritiek op het compatibilisme à la Frankfurt, bestrijdt ze dat de wijze waarop het incompatibilisme zelf hieraan gestalte geeft succesvol is. Immers een beroep op een ‘zelf’ dat metafysisch volstrekt onafhankelijk zou zijn van alle andere zaken kan verantwoordelijkheid evenmin toereikend verklaren. Zoals Hume al aangaf: willekeurigheid is geen vrije wil, en geen grond voor verantwoordelijkheid.

Tussen deze twee horens van het dilemma, situeert Wolf haar *Reason View*.²⁹ Met de libertariërs is zij van mening dat het voor ‘vrije wil’ inderdaad niet voldoende is dat onze wensen en waarden het gedetermineerde effect zijn van de voorgeschiedenis (zoals frankfurtiaanse compatibilisten mogelijk achten). De uiteindelijke evaluatie behoeft een referentiepunt dat aan deze redeloze keten van oorzaak en gevolg onttrokken is. Maar dit referentiepunt kan omgekeerd ook niet een ‘zelf’ zijn dat aan niets gerelateerd uit de lucht komt vallen (zoals libertarisme suggereert). Hoe dan wel? De tweede-ordevolities of waarden die iemands handelen richting geven moeten redelijk zijn, ze moeten gewild worden om de goede redenen. Dat vergt een extern referentiepunt waar het goed zijn van redenen aan af te meten valt. Om een vrij en verantwoordelijk persoon te kunnen zijn is er met andere woorden een extern referentiepunt nodig, dat zelf niet meer te evalueren is, doch dat de grond is van alle evaluaties.³⁰ En Wolf duidt dat aan als ‘het Ware en het Goede’.

Contextualiteit en pluraliteit

Voor Wolf impliceert vrij en verantwoordelijk zijn dat we handelen om ‘de goede redenen’. Die goede redenen worden gevonden door met de rede af te stemmen op het Ware en Goede, doch dat altijd vanuit en gevoed door

29 Wolf, *Freedom*, p. 76–78.

30 Er is hier een structurele overeenkomst te zien met de visie van Thomas van Aquino op de wil. Ook Thomas zegt dat waar onze wil uiteindelijk op gericht is, niet in onze macht ligt, omdat de wil zich van nature en dus noodzakelijk hecht aan het goede, wat ook bij hem de vrijheid van de persoon niet belemmert, maar juist mogelijk maakt. (Thomas van Aquino, *Summa Theologiae*, I, q.82, a.1 ad 3m. Zie hierover ook het artikel van Rudi te Velde in deze bundel).

De spanning dat een zelf om ‘autonoom’ te zijn, een externe ultieme bron nodig heeft is onderwerp van de zeer interessante studie van M. Coeckelbergh, *The metaphysics of autonomy. The reconciliation of ancient and modern ideals of the person*, Basingstoke/New York: Palgrave Macmillan, 2004.

de kennis van de feitelijke wereld. Concreet betekent dat: dat we ons laten informeren door de gegevens van de wereld om ons heen, daaromtrent tot zo goed mogelijke kennis komen, en gewapend met de kennis van feiten en van verbanden en mogelijke consequenties ons een oordeel vormen wat in deze context als ‘goed’ mag gelden.

In de notie ‘vrijheid’ zit een idee van onafhankelijkheid, maar de wijze waarop het libertarisme deze onafhankelijkheid invult, namelijk als het komen tot een beslissing zonder enige externe of interne invloed (tot uitdrukking gebracht met termen als ‘oorspronkelijkheid’, ‘ultieme bron’ of ‘onveroorzaakte oorzaak’ et cetera), wordt door Wolf dus als volstrekt in-adequaat beschouwd. Hoe zou het immers te denken zijn dat ik bij de beslissing een huis te kopen, me niet zou laten informeren over prijzen, aanwezigheid van scholen, winkels in de buurt enzovoort. Kennis van de wereld (en dus invloed van de wereld op mij) is in zo’n geval geen vijand maar vriend van vrijheid. De onafhankelijkheid die we willen, en waar de notie ‘vrije wil’ van getuigt, is dus niet, stelt Wolf, een onafhankelijkheid van de wereld zonder meer, maar een onafhankelijkheid van krachten die mij verhinderen een levenswijze te kiezen in het licht van een redelijke evaluatie van de wereld. De vrijheid die we wensen is ‘a freedom within reason and the world’.³¹ Wolf:

‘For the Reason View requires that the agent be able to act in accordance with reason, and part of what it *is* to act in accordance with Reason is to be sensitive and responsive to relevant changes in one’s situation and environment – that is, to be flexible.’³²

Wolf bepleit dus een contextafhankelijke visie van waar en goed, en daarbij een pluralistische.³³ Evenals Thomas van Aquino geeft ze aan dat er doorgaans meerdere opties zijn die aan de norm Goed voldoen (of lijken te voldoen). Om haar wat langer te citeren:

‘[W]hen we express the belief in the objectivity of values in terms of the thesis that normative judgments are susceptible to evaluation by Reason, there seems no reason whatsoever to expect the supposition of uniqueness to be true. Why should the reasons for and against different actions, policies, ideals, ways of life, and so on converge in such a way as to point to a single optimal choice? What expectations we do have that there will be a single best set of values for everyone must be

31 Aldus de boekbeschrijving gegeven door Oxford University Press.

32 Wolf, *Freedom*, p. 69, alwaar ze zelf verwijst naar p. 55–56.

33 Zie bijv. Wolf, *Freedom*, 69, 119–120, 123–126, 138.

counted as nonrational ones, resulting, perhaps, from a combination of wish-fulfillment, unjustified presumptions based on features of ordinary language, and a tendency too readily to draw analogies between normative and perceptual and scientific discourse.³⁴

Enkele kanttekeningen

Wanneer we met dit alles in het hoofd opnieuw kijken naar de kwestie van de mogelijke manipulatie, niet alleen op het niveau van de eerste-ordeverlangens maar ook op het niveau van de tweede-ordeverlangens (waarmee we de bespreking van Frankfurt beëindigden), dan blijkt dat Wolfs visie door haar beroep op een objectief referentiepunt, theoretisch althans, manipulatie kan onderscheiden van ‘normale’ beïnvloeding. Immers, doordat er in haar verhaal een extern element binnenkomt, namelijk ‘het Ware en het Goede’, kan gedacht worden dat iemands inschatting daarvan objectief een misinterpretatie is (bijvoorbeeld op grond van manipulatie of drugs).

Naar mijn smaak komt er met Wolfs wijze van redeneren wel een ander negatief punt mee, namelijk dat de vraag naar de vrije wil enkel en alleen binnen morele kaders wordt geplaatst. Dat is mijns inziens toch een te forse verenging van de problematiek van de vrije wil.

Congruentie van de latere Frankfurt met Wolf

Opvallend is dat Frankfurt in zijn latere publicaties *The importance of what we care about* (1988) en *Necessity, volition, and love* (1999) dicht bij Wolfs positie komt.³⁵ Zo zegt hij dat ‘it seems natural and reasonable to presume that when a person is acting under his own control, he will guide his conduct with an eye to those things that he considers to be the greatest importance to him’³⁶ en dat daarbij als richtsnoer nodig is ‘a fixed point from which a self-directed volitional process can begin’.³⁷ Zo komt Frankfurt te spreken van een *volitional necessity*, waaronder hij verstaat ‘a familiar but nonetheless somewhat obscure kind of necessity, in virtue of which his caring is not altogether under his own control’,³⁸ een noodzaak (die net als in de conceptie van Thomas van Aquino)

34 Wolf, *Freedom*, p. 136.

35 H. Frankfurt, *The importance of what we care about*, Cambridge: Cambridge University Press, 1988; H. Frankfurt, *Necessity, volition, and love*, Cambridge: Cambridge University Press, 1999.

36 Frankfurt, *Necessity, volition, and love*, p. 132.

37 Frankfurt, *Necessity, volition, and love*, p. 110.

38 Frankfurt, *The importance of what we care about*, p. 86. Interessant voor Frankfurts concept van ‘*volitional necessity*’: Coeckelbergh, *The metaphysics of autonomy*, p. 105–110.

inherent is aan de eigen aard van de wil.³⁹ Zoals Stefaan Cuypers terecht constateert: ‘with [this] appeal to the notion of a deliberate and reasonable decision, Frankfurt turns his approach into a *reason view* of autonomy’.⁴⁰

Outside the box: compatibilisme à la Frankfurt–Wolf mét ‘toeval’ à la William James

De bespreking van Frankfurt en Wolf staat hier in de context van het compatibilisme. En inderdaad, in wat we boven hebben gezien van wat Wolf naar voren brengt over vrijheid en verantwoordelijkheid en over het vermogen van de Rede om ons als kompas te richten op het Ware en Goede, lijkt dat nergens te leiden tot een contradictie met een universeel determinisme. En hetzelfde geldt voor Frankfurts model en voor zijn spreken over het noodzaakaspect van de wil die inherent gecommitteerd is aan het object van zorg en liefde, waarvoor Luthers befaamde uitspraak ‘Hier sta ik, ik kan niet anders’ model staat. Er zijn ook aspecten van Wolfs *Reason View* die hier verder niet uitgewerkt zijn en die aanleiding geven te denken dat ze niet binnen een volstrekt deterministische wereld gedacht kunnen worden.⁴¹ Het is echter niet mijn bedoeling dit kritiekpunt verder uit te werken. Liever wil ik aan het slot van mijn bespreking juist de deterministische ‘box’ verlaten en de verworvenheden van de compatibilistische visie op vrije wil combineren met een *indeterministische* metafysica (een metafysica waarin bij de huidige staat van de wereld meer dan één toekomst mogelijk is, en welke toekomst werkelijkheid zal worden mede afhankelijk is van wat wij doen en laten). Daarvan hier een summier schets.

Mij dunkt dat het de *winst* is van het compatibilisme, in zijn latere gedaantes, dat het ons opnieuw ervan doordrongen heeft hoe cruciaal bepaalde aspecten van determinatie of noodzaak zijn voor inzicht in wat vrije wil betekent: dat ons handelen bepaald wordt door onze wil, dat die uitvoerende wil bepaald kan worden door onze tweede-ordevolentie, en – met Wolf – dat die tweede-ordevolentie door middel van de rede gehoor kan geven aan het Ware en Goede of – met Frankfurt – zich noodzakelijk hecht aan waar onze

39 Thomas’ uiteenzetting over in welke betekenissen noodzaak strijdig is met de wil, en in welke betekenis noodzaak het wezen van de wil uitmaakt, is te vinden in: *Summa Theologiae*, I.82.1c. Zie ook het artikel van Rudi te Velde over Thomas van Aquino (deze bundel).

40 S. Cuypers, ‘Harry Frankfurt on the will, autonomy and necessity,’ in: *Ethical Perspectives* 5/1, 1998, p. 44–52, m.n. p.48.

41 Zie M. Sie & A. Wouters, ‘The BCN challenge to compatibilist free will and personal responsibility’, in: *Neuroethics* 3, 2010, p. 121–133, m.n. p. 129. (doi 10.1007/s12152-009-9054-8).

zorg en liefde naar uitstaat.⁴² Ik beschouw dit als een zeer weldadige verdieping van wat onder persoon zijn en vrijheid te verstaan is op een wijze die mijlenver afstaat van de onhoudbare gedachte dat vrij zijn zou betekenen los van alles om het even wat kunnen doen. Mét deze winst van het opnieuw naar voren brengen van de belangrijke relaties tussen vrijheid en noodzaak ('opnieuw' omdat mensen als Thomas van Aquino, Luther, Spinoza dat ieder op geheel eigen wijze reeds hebben gethematiseerd⁴³) komt er echter volgens mij ook ruimte om de mogelijkheid en wellicht ook het aanvullend belang van een zeker indeterminisme te kunnen denken.

Hoe zie ik dat voor me? De meest gangbare (en door zeer velen afgevoerd) visie op de rol van toeval ten opzichte van de wil, plaatst toeval in de rol van (quasi)oorzaak van de wilsbeslissing. Aan David Hume wordt de volgende observatie toegeschreven waarom een dergelijke aan toeval ontsproten wil net zo min soelaas biedt als een gedetermineerde wil:

'Either our actions are determined, in which case we are not responsible for them, or they are the result of random events, in which case we are also not responsible for them.'⁴⁴

In dezelfde trant formuleert Peter van Inwagens zijn argument tegen een vrije wil die gebaseerd zou zijn op een zo begrepen rol van toeval:

'The *Mind* argument proceeds by identifying indeterminism with chance and by arguing that an act that occurs by chance, if an event that occurs by chance can be called an act, cannot be under the control of the alleged agent and hence cannot have been performed freely.'⁴⁵

Deze kritieken op een toevallige, indeterministische wil deel ik. Wat ik hier echter naar voren wil brengen is dat er ook een weinig genoemde *andere* mogelijkheid is om toeval in het verhaal een rol te laten spelen, een mogelijkheid waarvoor het door Hume en door Van Inwagen geformuleerde bezwaar niet geldt.

42 H. Frankfurt, *The reasons of love*, Princeton, 2004.

43 Zie daarvoor de betreffende artikelen in deze bundel van Rudi te Velde, Dick Akerboom en Han van Ruler.

44 Dit dilemma wordt soms aangeduid als 'Hume's fork', al slaat die uitdrukking primair op een andere zaak (Simon Blackburn, *The Oxford dictionary of philosophy*, Oxford: Oxford University Press, 2005, p. 173).

45 P. Van Inwagen, *An essay on free will*, Oxford: Oxford University Press, 1983, p.16. Dit argument wordt als '*Mind*-argument' aangeduid naar het tijdschrift *Mind*, waarin dergelijke argumenten vaak verschenen.

De visie waarop ik doel wordt wel aangeduid als het ‘tweestadiamodel’ van vrije wil.⁴⁶ In dit model is er een samenspel van zowel *toeval* als *determinering*: toeval is dan de bron van alternatieve handelingsmogelijkheden, en de determinering vindt plaats in de gedaante van een selectie daaruit op basis van de relatieve waarden van de verschillende mogelijkheden. De resulterende gewilde handeling is volgens dit model dus niet het effect van een proces met toeval in de rol van oorzaak, maar het resultaat van een determinerende selectie uit (toevallige) alternatieve mogelijkheden. Een dergelijk model is door verschillende denkers meer of minder duidelijk naar voren gebracht, van wie William James wellicht de eerste is.⁴⁷ Positief aan deze zienswijze is dat het een volstrekt andere rol geeft aan toeval, dan de onbevredigende rol van toeval als bepaler van de wil. En het mag verbazing wekken dat deze mogelijkheid zo algemeen over het hoofd wordt gezien, in feite overal waar gesproken wordt van het vrijewildilemma.

Desondanks laat het zo gepresenteerde tweestadiamodel wel een uiterst belangrijke vraag onbeantwoord, namelijk de vraag wat de grond is van de inschatting van de relatieve waarden van de diverse mogelijkheden en van de daarop gebaseerde selectie. Wil een dergelijke evaluatieve inschatting mogelijk zijn, dan vergt dat een maat, een referentiepunt, op basis waarvan de ene handelingsmogelijkheid als aantrekkelijker of waardevoller verschijnt dan het andere. Zo’n referentiepunt zagen we door Susan Wolf uitgewerkt worden in haar *Reason View*, namelijk als het Ware en Goede op grond waarvan (in principe) goede van slechte redenen te onderscheiden zijn. We zagen het ook benoemd worden door Frankfurt in zijn latere werken als het ‘vaste punt dat onze liefde richt’ (zoals bij Plato en bij Aristoteles de eerste bewegende liefde is).

Mij dunkt daarom dat Wolf en de latere Frankfurt hiermee iets te bieden hebben wat het tweestadiamodel in zichzelf mist. En omgekeerd voegt het tweestadiamodel iets toe aan het compatibilistische model van Frankfurt en Wolf, namelijk de mogelijkheid om *echte* alternatieve mogelijkheden een consistente plaats te geven, en niet enkel en hooguit de tegenfeitelijke conditionele alternatieve mogelijkheden. Door ‘Frankfurt-Wolf’ toe te voegen aan ‘James’, of ‘James’ aan ‘Frankfurt-Wolf’, wordt er volgens mij werkelijk een stap voorwaarts gezet in het doordenken van de vrije wil. Een stap die de tegenstelling compatibilisme of incompatibilisme te boven komt, omdat het een samenspel vergt van indeterminisme en

46 Zie B. Doyle, ‘Jamesian free will. The two-stage model of William James’, in: *William James Studies* 5, 2010, p. 1-28. Doyle noemt als latere denkers die een dergelijke zienswijze propageerden onder meer Karl Popper en Daniel Dennett.

47 James, ‘The dilemma of determinism’.

determinisme, ieder met een eigen rol: vrije wil vergt enerzijds *alternatieve mogelijkheden* waaruit te kiezen valt (die kunnen eventueel conditioneel en tegenfeitelijk opgevat worden, zoals Daniel Dennett laat zien,⁴⁸ maar er is geen reden meer ze daartoe te beperken en hoeveel interessanter wordt het wanneer ze als werkelijke indeterministische alternatieven gezien kunnen worden), en anderzijds *een extern referentiepunt* dat zelf niet meer te kiezen of te evalueren is, maar dat het ons mogelijk maakt in de gegeven situatie de relatieve waarde van de verschillende opties te onderscheiden en dat zo onze wil constitueert. Is dit nieuw? Nee. Reeds in de dertiende eeuw liet Thomas van Aquino beide elementen expliciet zien,⁴⁹ zo ook in de twintigste eeuw Alfred North Whitehead in zijn procesmetafysica.⁵⁰ Maar het moet kennelijk wel weer als nieuw ingebracht worden in het voortdurende debat over een deterministische of indeterministische visie op vrije wil.

Naschrift

Is hier al met al een filosofische weg ingeslagen die volledig afvoert van de neurowetenschap? Daar zou ik niet te zeker van zijn. De Nijmeegse neurowetenschapper Herman Kolk geeft in zijn boek *Vrije wil is geen illusie* aan hoe welbepaalde hersenstructuren en hersenprocessen de werking van de wil à la William James verhelderen.⁵¹ Kolk deelt met Victor Lamme (en andere representanten van het Nieuwe Onbewuste) de mening dat veel gedrag onbewust tot stand komt. Anders echter dan Lamme, die de aangeleerde reflex als fundamentele bouwsteen ziet die door eindeloze herhaling inslijpt tot menselijk gedrag, stelt Kolk dat mens en dier leren doordat ze *gedrag herhalen waar ze iets goeds van verwachten*. De zogeheten orbitofrontale schors, zo beschrijft Kolk, is een hersenstructuur die aan-

48 Volgens Daniel Dennett kan het tweestadiummodel gedacht worden binnen een volstrekt deterministische metafysica. Zie: D. Dennett, 'On giving libertarians what they say they want', in: D. Dennett, *Brainstorms. Philosophical essays on mind and psychology*, Cambridge, MA: MIT Press [1978] 1981, p. 286-299.

49 Thomas bespreekt de intrinsieke relatie tussen wil en noodzaak (*Summa Theologiae* I, q.81, a.1) én dat de wil niet in ieder aspect genoodzaakt is doch ook keuzevrijheid kent (*Summa Theologiae* I, q. 82, a.2 en I, q.83, a.1).

50 A.N. Whitehead, *Process and reality. An essay in cosmology*, New York: Free Press, [1929] 1978. Voor een inleiding op Whiteheads visie, zie bijvoorbeeld: P. Oomen, 'Gods Eros en het natuurlijke Godsverlangen in de kosmologie van Whitehead', in: R. te Velde (red.), *In het spoor van Plato's Symposium. Eros in de westerse cultuur*, Nijmegen: Valkhof Pers, 2010, p. 116-139.

51 H. Kolk, *Vrije wil is geen illusie. Hoe de hersenen ons vrijheid verschaffen*, Amsterdam: Bert Bakker, 2012, m.n. p. 44-47, 85-90.

geeft wat waardevol is, of beter: wat de verwachte meerwaarde is van een bepaalde handeling ten opzichte van alternatieven. De vlakbij gelegen mediale frontale schors beschrijft Kolk als een *aandacht*structuur die in staat is om op grond van die verwachte meerwaarde gedragspatronen *selectief* te versterken. Zo brengt het ‘*voluntary*’ gedrag tot stand: dat wil zeggen, ‘gedrag waarmee het organisme bereikt wat voor hem van waarde is’.⁵² Over de herkomst van die waarden zegt Kolk, dat die deels evolutionair bepaald en daarmee aangeboren zijn en deels cultureel verworven zijn, doch beide lijken wel in een en hetzelfde waardesysteem geïntegreerd te worden.⁵³ Hoewel veel van de beschreven processen nagenoeg onbewust kunnen verlopen, heeft bewustzijn toch een groot belang: onder veel meer dat je je kunt losmaken van het hier en nu en je je kunt voorstellen wat in de toekomst zou kunnen gebeuren. Daarmee komen mogelijke alternatieve gedragspatronen beschikbaar, waarbij je, op grond van de verwachte waarden die aan die gedragspatronen zijn gekoppeld, die handelingsmogelijkheid kiest die voor jou de meeste waarde toevoegt. Automatisch, zonder daar apart toe te hoeven besluiten, aldus Kolk.⁵⁴ Zo kun je sterk geautomatiseerde gedragspatronen overrulen met een alternatief waarvan je iets beters verwacht.⁵⁵

Het is op zijn minst interessant te noemen hoe de werking van de wil à la James, en hoe oude theologische ideeën van Thomas van Aquino en moderne filosofische ideeën als van Harry Frankfurt en Susan Wolf over waarden en waardepreferenties die onze wil constitueren het meest waardevolle na te streven, iets van een materiële onderbouwing vinden in dergelijke bevindingen vanuit de neurowetenschap. Wellicht kunnen filosofie en neurowetenschap op deze weg betekenisvolle verdere stappen zetten.⁵⁶

52 Kolk, *Vrije wil is geen illusie*, p. 86.

53 Kolk, *Vrije wil is geen illusie*, p. 152–156.

54 H. Kolk, *Vrije wil, bewustzijn en het brein*, afscheidsrede, Radboud Universiteit Nijmegen, 2009, p. 13.

55 Kolk, *Vrije wil is geen illusie*, p. 73–89, 137–139.

56 Graag wil ik Edith Brugmans, Anton de Wit en Arno Wouters bedanken voor hun commentaren op een eerdere versie van deze tekst, en Arno vooral ook voor zijn gedegen en kritische inbreng in onze vele stimulerende discussies over het compatibilisme in zijn verschillende varianten.