

Vrije wil, verantwoordelijkheid en sociale beïnvloeding

De grenzen van het compatibilisme

Maud Beekmans

S 4645448

Begeleid door Dr. Leon de Bruin

Aantal woorden: 16.042

15 – 07 – 2017

Scriptie ter verkrijging van de graad “Master of Arts” in de filosofie. Radboud Universiteit: Nijmegen.

Hierbij verklaar en verzeker ik, Maud Beekmans, dat deze scriptie zelfstandig door mij is opgesteld, dat geen andere bronnen en hulpmiddelen zijn gebruikt dan die door mij zijn vermeld en dat de passages in het werk waarvan de woordelijke inhoud of betekenis uit andere werken – ook elektronische media – is genomen door bronvermelding als ontlening kenbaar gemaakt worden.

's-Hertogenbosch, 15 – 07 – 2017.

Inhoudsopgave

1. Inleiding	4
2. Het vrije wil debat	7
<i>Could have done otherwise</i> , determinisme en verantwoordelijkheid	8
Klassiek compatibilisme en <i>the consequence argument</i>	9
Libertarisme en hard incompatibilisme	11
Ayer en Morris	12
Frankfurt	14
3. Intentionele handelingen en bewuste keuzes	16
Bewuste deliberatieprocessen	17
Onbewuste redenen	19
<i>Priming</i>	21
4. Mate van restricties	25
Begrenzungen en beperkingen	25
Onduidelijke criteria	28
Conclusie	30
5. Milgram's experiment: I	32
Onderzoeksmethode	32
Resultaten	34
Interne restricties	36
Externe restricties	38
Verantwoordelijkheid van de proefpersonen	44
6: Milgram's experiment: II	45
Vrije wil van de proefpersonen	45
Frankfurt over verantwoordelijkheid	49
Verantwoordelijkheid van de proefpersonen	51
7: Vrije wil, verantwoordelijkheid en sociale beïnvloeding	53
8: Conclusie	56
Dankwoord	59
Bibliografie	60

1. Inleiding

Tijdens het experiment van Milgram in de jaren '60 van de vorige eeuw waren het reguliere burgers die hun 'leerlingen' (fictief) electrocuteerden met de dood tot gevolg, na aandringen van een man in een witte jas. Onder andere door de invloed van een autoritair figuur in dezelfde ruimte, handelden de proefpersonen niet in overeenstemming met hun eigen wil: sommige proefpersonen gaven in een interview aan dat ze *eigenlijk* hadden willen stoppen met het experiment.¹ Er zijn legio andere voorbeelden waarbij personen anders handelden dan dat ze eigenlijk hadden gewild. Denk hierbij aan de massamoorden in My Lai tijdens de Vietnamoorlog, waarover veel betrokken soldaten later spijt betuigden, of aan ontgroeningen bij studentenvereniging, waarbij zogenaamde 'feuten' de meest gruwelijke zaken ondernemen om ouderejaars te laten zien dat ze een waardig lid zullen zijn.

Ook kunnen onbewuste gedachten en verlangens van personen sociaal sterk worden beïnvloed. De propaganda technieken van Joseph Goebbels – die sommige frames van een *Coca-Cola* reclame liet vervangen door ratten en Joden, waardoor men onbewust Joden met ratten ging associëren – zijn hiervan voorbeelden. *Coca-Cola* zelf maakt ook gebruik van onbewuste manipulatie technieken, bijvoorbeeld door kinderstoeltjes op kleuterscholen in Amerika te financieren. Op deze rode stoeltjes staat echter wel het *Coca-Cola* logo vermeldt, waardoor kinderen *Coca-Cola* gaan relateren aan 'happiness' en 'fun' en later eerder geneigd zullen zijn *Coca-Cola* te drinken.²

Handelt een volwassene, die als kind veelvuldig op rode *Coca-Cola* stoeltjes heeft gezeten, volledig uit vrije wil als deze een blikje *Coke* opentrekt? Zijn de Amerikaanse soldaten moreel verantwoordelijk voor het executeren van een heel dorp of valt deze verantwoordelijkheid (deels) af te schuiven op omgevingsfactoren en de sociale beïnvloeding van de leidinggevende in kwestie?

¹ S. Milgram, *Obedience to Authority. An Experimental View* (New York: Harper & Row Publishers, 1974), 54, 73 en 80.

² S. Soechtig, *Fed Up* (Atlas Films, 2014), MPEG4 Video, <https://archive.org/details/FedUp2014Documentary>, 1:01:15.

Als ze niet geheel uit vrije wil de trekker overhalen, dan lijkt het immers lastig de soldaten volledig verantwoordelijk te houden voor hun gedrag. Handelt een proefpersoon in het experiment van Milgram uit vrije wil als hij de genadeklap aan een ‘leerling’ geeft en kan hij verantwoordelijk worden gehouden voor het toedienen van het hoogste voltage?

In deze scriptie staat de volgende onderzoeksvraag centraal: *In hoeverre vormt sociale beïnvloeding een bedreiging voor het begrip van vrije wil en voor morele verantwoordelijkheid?* Het experiment van Milgram zal als uitgangspunt worden genomen voor het beantwoorden van deze vraag.

De scriptie is als volgt opgebouwd: in het tweede hoofdstuk zal eerst het vrije wil debat uiteen worden gezet. De klassieke definitie van vrije wil als *could have done otherwise* zal hierin eerst worden geëxpliciteerd en het mogelijke probleem van determinisme zal worden besproken. Volgens determinisme is iedere gebeurtenis noodzakelijkerwijs het gevolg van voorgaande gebeurtenissen, waardoor de mogelijkheid om anders te handelen in het gedrang lijkt te komen. Er zijn grofweg twee kampen in het debat: de compatibilisten, die beweren dat vrije wil en determinisme zonder problemen met elkaar te verenigen zijn, en de incompatibilisten die het tegenovergestelde verdedigen.

Eerst zullen we naar de klassieke vorm van compatibilisme kijken. Dan zullen we kijken naar de ‘tegenpartij’: de incompatibilisten, die betogen dat vrije wil en determinisme niet tegelijkertijd waar kunnen zijn. Ook incompatibilisten kunnen weer worden opgesplitst in enerzijds libertaristen en anderzijds harde incompatibilisten. De eerste stellen dat determinisme niet altijd waar is, waardoor vrije wil in sommige gevallen wel degelijk bestaat. Harde incompatibilisten beweren daarentegen dat vrije wil niet waar kan zijn. Tot slot zullen de neo-compatibilistische visies van Ayer en Frankfurt worden doorgenomen.

In het derde en vierde hoofdstuk zullen de twee voorwaarden voor vrije wil, die door Morris zijn afgeleid uit het werk van Ayer, kritisch bekeken worden. In het derde hoofdstuk komt de eerste voorwaarde aan bod: een handeling dient voort te komen uit een bewuste keuze. Mogelijke bedreigingen voor deze voorwaarde zullen aan bod komen, maar uiteindelijk kan worden geconcludeerd

dat deze bedreigingen minder gevaarlijk zijn dan ze op het eerste gezicht lijken. Aan een intentionele handeling gaat namelijk *altijd* een bewuste keuze vooraf.

In het vierde hoofdstuk wordt de tweede voorwaarde besproken, waarmee wordt gesteld dat een handeling niet door interne of externe beperkingen mag worden begrensd. De vraag is echter wanneer zo'n beperking daadwerkelijk als restrictief kan worden gezien. Costa specificeert deze interne en externe factoren, maar uiteindelijk zal kunnen worden geconcludeerd dat zijn poging niet tot harde criteria leidt: ook met de specificatie van Costa kan namelijk niet worden bepaald wanneer beperkingen als *volledig* restrictief kunnen worden bestempeld.

Hoofdstuk vijf staat in het teken van experiment van Milgram. De twee behandelde criteria voor vrije wil zullen worden toegepast op het experiment en er zal een graduele schaal worden voorgesteld. Hieruit kan worden geconcludeerd dat sommige proefpersonen vrijer handelden dan anderen proefpersonen op het moment dat ze het hoogste voltage van 450 volt toedienden. Tot slot zal worden besloten dat sommige proefpersonen *meer* verantwoordelijk kunnen worden gehouden voor het toedienen van dit voltage als ze *vrijer* handelden.

Ook in het zesde hoofdstuk zal naar het experiment van Milgram worden gekeken. Echter zullen zowel vrije wil als verantwoordelijkheid vanuit de neo-compatibilistische visie van Frankfurt worden beschouwd. Het zal blijken dat dit resulteert in andere conclusies over de vrije wil en verantwoordelijkheid van de proefpersonen. In het zevende hoofdstuk zal vanuit het experiment van Milgram worden uitgezoomd en zal het begrip 'sociale beïnvloeding' breder worden getrokken. Vervolgens zal in de conclusie de scriptie in grote lijnen worden samengevat. Er zal tot slot worden geconcludeerd dat welk antwoord volgt op de vraag of sociale beïnvloeding een gevaar vormt voor vrije wil en verantwoordelijkheid, afhankelijk is van welke compatibilistische positie wordt gehonoreerd: de positie van Ayer, dan wel die van Frankfurt.

2. Het vrije wil debat

Het vrije wil debat is een discussie met veel haken en ogen: de formulering van ‘vrije wil’ staat namelijk niet vast en verschillende definitiekwesties zijn nog lang niet uitgevochten. De meest gangbare opvatting van vrije wil is als volgt: iemand is vrij als hij anders had kunnen handelen dan dat hij heeft gedaan. Er lijkt een spanning te bestaan tussen dit begrip van vrije wil enerzijds en determinisme anderzijds. Determinisme stelt dat iedere gebeurtenis noodzakelijkerwijs het gevolg is van voorgaande gebeurtenissen. Dit lijkt op het eerste gezicht een probleem voor de vrije wil te vormen: als de gehele wereld gedetermineerd – en daarmee bepaald – is, hoe zou iemand dan überhaupt anders kunnen handelen?

Als iemand niet volledig uit vrije wil handelt, lijkt het moeilijk deze persoon verantwoordelijk te houden voor zijn of haar daden. De inzet is dan ook hoog – de meeste filosofen in het debat willen de notie van verantwoordelijkheid namelijk niet opgeven. Aan de hand van deze ogenschijnlijke spanning tussen determinisme en vrije wil kunnen de basisposities in het vrije wil debat worden geschetst. Aan de ene kant zijn er compatibilisten, die beweren dat determinisme en vrije wil prima te verenigen zijn met elkaar. Een klassiek compatibilist als Hobbes betoogde bijvoorbeeld dat determinisme geen probleem vormt als je niet door externe factoren wordt beperkt. Mocht dit laatste wel het geval zijn, dan handel je niet uit vrije wil. Dit is bijvoorbeeld het geval als een discotheek wil bezoeken, maar je door de portier wordt geweigerd.

Aan de andere kant van de medaille bevinden zich de incompatibilisten. Volgens hen gaan determinisme en vrije wil niet door één deur. Grofweg bestaan incompatibilisten uit enerzijds libertaristen, die stellen dat determinisme niet waar is en vrije wil in ieder geval in bepaalde situaties wel bestaat. Anderzijds zijn er harde incompatibilisten, die beweren dat determinisme waar is en de vrije wil niet. Daarmee is de kous nog niet af: zelfs *als* determinisme niet waar zou zijn, dan zijn vrije wil en verantwoordelijkheid alsnog illusionair.

Daarna zal de neo-compatibilistische visie van Ayer worden besproken. Hij stelt dat niet determinisme als de contrastklasse van vrije wil moet worden

gezien, maar dat we moeten denken in termen van *constraints*, oftewel *beperkingen*³. Tot slot komt het neo-compatibilistische standpunt van Frankfurt aan bod. Vrije wil zou volgens hem een stuk letterlijker moeten worden opgevat: het gaat niet zozeer om vrij te kunnen *handelen* zoals je wil, maar om vrij te kunnen *willen* zoals je wil.⁴

***Could have done otherwise*, determinisme en verantwoordelijkheid**

In deze sectie zullen we eerst kijken naar de compatibilistische visie, die stelt dat vrije wil en determinisme verenigbaar zijn. Stel dat vrije wil wordt gekenmerkt door de mogelijkheid om uit meerdere opties te kunnen kiezen, oftewel: *to could have done otherwise*. Intuïtief lijkt dit een geslaagde formulering. Als een vrouw met een bivakmuts en revolver de plaatselijke bank overvalt, dan zou je zeggen dat ze dit uit vrije wil doet. De vrouw had immers ook iets anders kunnen doen: ze had thuis kunnen blijven bij de kinderen of een dagje naar het strand van Domburg kunnen gaan. Het begrip ‘determinisme’ lijkt hier echter roet in het eten te kunnen gooien. Mocht de wereld volledig gedetermineerd zijn, dan is er immers maar één uitkomst mogelijk: die uitkomst die zich ook daadwerkelijk in de wereld op een bepaald moment voordoet.⁵

In het geval van de vrouw die de bank overvalt lijkt er dan ook geen andere mogelijkheid te zijn dan het beroven van de bank. Dat is immers de enige mogelijkheid die zich op *dat* moment in de wereld voor kan doen als de wereld volledig gedetermineerd is. Dit lijkt tevens problemen op te leveren voor verantwoordelijkheid. De vrouw aanspreken op haar gedrag lijkt immers niet op zijn plek: ze kon niet anders handelen dan dat ze heeft gedaan.

Echter, zo stelt de compatibilist, is het belangrijk om hier een onderscheid duidelijk te maken tussen enerzijds de fatalist en anderzijds de compatibilist. De

³ A.J. Ayer, “Freedom and Necessity” in *Free will*. Geredigeerd door G. Watson (Oxford: Oxford University Press, 1982), 19.

⁴ H. Frankfurt, “Freedom of the will and the concept of a person” in *The Journal of Philosophy* (Volume 68, nummer 1, 1971), 90.

⁵ M. McKenna en J. Coates, “Compatibilism” in *The Stanford Encyclopedia of Philosophy*. Geredigeerd door E. Zalta (2016).

fatalist⁶ zou namelijk beweren dat de betreffende handeling van de vrouw – het overvallen van de bank – per definitie zou gebeuren en dat het daarom inderdaad zinloos voor haar is om alternatieve mogelijkheden te overwegen. Ze heeft zich er maar bij neer te leggen dat ze de bank zal overvallen: dat is immers hoe het gaat verlopen.

De compatibilist stelt daarentegen dat de vrouw wel degelijk haar keuzes af moet wegen, ook al is de wereld gedetermineerd. Ze weet namelijk helemaal niet wat er precies gaat gebeuren, maar slechts dat *wat* er zal gebeuren gedetermineerd zal zijn. Dit neemt niet weg dat ze zal moeten kiezen tussen (in dit geval) de overval, het thuisblijven of het uitje naar Domburg en dat *zijzelf* verantwoordelijk is voor die keuze. De keuze voor het overvallen van de bank kan namelijk een schakel zijn in de gedetermineerde keten van oorzaak en gevolg. Hierdoor staan het maken van de keuze en determinisme niet in contrast met elkaar: ook de keuze *zelf* is dan deterministisch bepaald, maar wel een noodzakelijk en cruciaal onderdeel om tot de uiteindelijke uitkomst van de gepleegde overval te komen.

Klassiek compatibilisme en *the consequence argument*

De deterministische visie zoals hierboven beschreven, kan als een klassiek compatibilistische visie gezien worden. Determinisme en het maken van een keuze door een actor staan in het voorbeeld immers niet op gespannen voet met elkaar. Ook bij het klassiek compatibilisme van Hobbes staat het verbinden van determinisme en vrije wil centraal. Volgens hem is iemand vrij als: “(...) he finds no stop, in doing what he has the will, desire or inclination to doe.”⁷ Deze definitie kan worden opgesplitst in twee voorwaarden⁸. Iemand is volgens Hobbes vrij als:

⁶ C. Hofer, “Causal Determinism” in *The Stanford Encyclopedia of Philosophy*. Geredigeerd door E. Zalta (2016).

⁷ T. Hobbes, *Hobbes's Leviathan*, herdruk van het origineel uit 1651 (Oxford: Clarendon Press, 1929), p. 161.

⁸ McKenna en Coates, “Compatibilism”.

1. Hij kan handelen naar zijn propositionele attitudes. Met andere woorden: Als hij kan doen wat hij verlangt of wil.
2. Hij deze handeling onbelemmerd, dat wil zeggen, zonder externe beperkingen, uit kan voeren.

Er is sprake van een externe beperking als ik bijvoorbeeld een kamer zonder ramen wil verlaten, maar de deur door iemand anders op slot wordt gedraaid voordat ik bij de klink ben. Ik kan dan niet handelen naar mijn vrije wil; de deur wordt immers dichtgedraaid waardoor de propositionele attitude ‘ik wil de kamer via de deur verlaten’ niet kan worden uitgevoerd.

Van Inwagen heeft echter een conditioneel argument tegen klassiek compatibilisme opgevoerd, dat bekend staat als *the consequence argument*.⁹ Hierin vraagt hij zich af of ‘Judge J’, die door zijn hand op een bepaald moment naar beneden te houden in de rechtbank een verdachte veroordeelt, ook anders had kunnen handelen *als* determinisme waar zou zijn. Mocht dit laatste het geval zijn, dan is het opsteken van de hand door de rechter onmogelijk: zou je namelijk een willekeurig tijdstip pakken vóór deze gebeurtenis, en voor het gemak misschien zelfs een tijdstip vóór de geboorte van ‘J’, en daar vervolgens alle causale natuurwetten op loslaten, dan had het niet optillen van de hand noodzakelijkerwijs moeten gebeuren. ‘Judge J’ had niet anders kunnen handelen: hij had niet wél zijn hand op kunnen steken en daardoor de verdachte vrij kunnen spreken.

Om het in van Inwagens eigen woorden te formuleren: “If A and B are any propositions that express the state of the world at some instants, then the conjunction of A with the laws of physics entails B.”¹⁰ ‘A’ is hierbij bijvoorbeeld de toestand van de wereld op een bepaald moment vóór de geboorte van ‘Judge J’ en ‘B’ de toestand op de wereld op het moment dat ‘Judge J’ zijn arm naar beneden houdt in de rechtbank. A + [alle deterministische natuurwetten] heeft noodzakelijkerwijs B tot gevolg. Mocht determinisme waar zijn, dan had ‘Judge J’ niets anders kunnen doen dan zijn hand naar beneden houden. Mocht determinisme waar zijn, dan vormt dit volgens *the consequence argument* wel

⁹ P. van Inwagen, “The Incompatibility of Free Will and Determinism” in *Philosophical Studies* (Volume 27, nummer 3, 1975), 185 – 199.

¹⁰ *Ibid.*, 186.

degelijk een probleem voor vrije wil als *could have done otherwise*. Van Inwagens argument vormt hiermee een sterk tegenargument tegen het klassiek compatibilisme van filosofen als Hobbes. Echter is het van belang te vermelden dat het slechts een conditioneel tegenargument vormt *als* vrije wil wordt gedefinieerd als *could have done otherwise*.

Libertarisme en hard incompatibilisme

Het klassiek compatibilisme van Hobbes wilde, zoals we hebben gezien, het bestaan van vrije wil en verantwoordelijkheid verdedigen door te stellen dat determinisme en vrije wil goed samengaan. Een andere strategie om vrije wil en verantwoordelijkheid te verdedigen, is door determinisme te ontkennen. Deze vorm van incompatibilisme wordt libertarisme genoemd. Kane is zo'n libertarist en stelt dat determinisme niet altijd waar is en dat vrije wil daardoor in sommige gevallen wel degelijk vrij spel heeft. Eerst stelt Kane dat handelingen vrij zijn als de actor de ultieme bron is van op zijn minst sommige van zijn intenties of karaktereigenschappen.¹¹ Om de uiteindelijke oorsprong van intenties en karaktereigenschappen te kunnen zijn, moet (absoluut) determinisme echter wel worden ontkend. Als determinisme waar is, dan zijn ook intenties en karaktereigenschappen immers slechts een vaststaande schakel in de lange – en door deterministische natuurwetten bepaalde – keten van causaliteit.

Kane stelt dan ook dat er wel degelijk onbepaaldheid voorkomt in de natuurlijke wereld. Er zijn immers natuurwetten die niet deterministisch maar probabilistisch van aard zijn, zoals sommige kwantummechanische wetten. Ook het brein is volgens Kane gevoelig voor kwantummechanische onbepaaldheid, die op het neurale niveau kan optreden.¹² Hierdoor kunnen sommige handelingen niet-deterministisch zijn en toch tot de natuurlijke wereld behoren.¹³ Deze handelingen hebben de actor als ultieme bron en zijn daarom volledig vrij. De

¹¹ R. Kane, "Libertarianism" in *Philosophical Studies* (Springer 144: 35 – 44, 2009), 36.

¹² *Ibid.*, 39.

¹³ M. Slors et al., *Philosophy of Mind, Brain and Behaviour* (Amsterdam: Boom Uitgevers, 2015), p. 184.

actor bepaalt in die gevallen namelijk de intenties die voorafgaan aan een bepaalde handeling.

Een harde incompatibilist als Pereboom is het hier radicaal mee oneens. Hij verwerpt niet alleen de compatibilistische visie dat vrije wil en determinisme samen zouden kunnen gaan, maar ook libertarisme:¹⁴ of determinisme nu waar is of niet, in beide gevallen zijn vrije wil en verantwoordelijkheid illusies. Zelfs *als* handelingen middels kwantummechanische processen tot stand zouden kunnen komen, dan is het nog maar de vraag in hoeverre zo'n handeling als 'vrij' kan worden bestempeld. Kwantummechanica kenmerkt zich immers door onbepaaldheid en willekeur.

Daarnaast rijst de vraag hoe onbepaalde hersenactiviteiten op het kwantumniveau bewuste en intentionele handelingen tot gevolg kunnen hebben.¹⁵ Om van een bewuste keuze tot een intentionele handeling te komen, lijkt causaliteit ten slotte noodzakelijk: het is de keuze die de intentionele handeling *tot gevolg* heeft. Hoe dit middels probabilistische wetten tot stand kan komen, is niet helder.

Ayer en Morris

Nu de twee incompatibilistische visies besproken zijn, zullen we kijken naar de neo-compatibilistische visie van Ayer en – in de volgende paragraaf – de neo-compatibilistische visie van Frankfurt. Volgens Ayer moeten we vrije wil anders definiëren dan als *could have done otherwise* en middels een tactische zet weet hij het probleem dat determinisme voor vrije wil kan vormen te vermijden. Hij stelt namelijk dat vrije wil en determinisme onterecht als tegenpolen van elkaar worden gezien: “For it is not, I think, causality that freedom is to be contrasted with, but constraint.”¹⁶ Het tegenovergesteld van vrije wil is dus niet zozeer determinisme, volgens Ayer, maar beperking [*constraint*].

¹⁴ D. Pereboom, “Living Without Free Will: The Case for Hard Incompatibilism” in de online versie van *The Oxford Handbook of Free Will*. Geredigeerd door R. Kane (2009), 1.

¹⁵ Slors, *Philosophy of Mind, Brain and Behaviour*, 185.

¹⁶ Ayer, “Freedom and Necessity”, 19.

Neem ter verheldering het volgende voorbeeld: stel, je hebt contant geld nodig om iets te kopen en loopt daarom naar de dichtstbijzijnde pinautomaat om geld te pinnen. Dit doe je zonder een interne of externe beperking: er zit geen stemmetje in je hoofd dat stelt dat je *moet* pinnen of dat anders de wereld vergaat en er is ook niemand die een pistool tegen je hoofd zet en eist dat je dat geld nu uit de muur moet trekken. Je leeft dan misschien in een determineerde wereld, maar je gaat *zelf* – uit vrije wil – dat geld pinnen, zonder dat er sprake is van een beperking.

Determinisme is volgens Ayer dan ook geen inperking van vrije wil. Een interne of externe beperking, echter, staat wel op gespannen voet met vrije wil. Op het moment dat er namelijk wél een stemmetje in je hoofd zit dat je allerlei opdrachten toefluistert, of als er iemand met een pistool op je gericht staat, word je beperkt in je handelen. In beide gevallen – het pinnen zonder interne of externe beperking en het pinnen met interne of externe beperking – is de wereld gedetermineerd. Echter, in het eerste geval is iemand vrij en in het tweede geval niet. Hierdoor is determinisme volgens Ayer geen gevaar voor vrije wil en zijn een gedetermineerde wereld en vrije wil compatibel met elkaar.

Het compatibilisme van Ayer kan worden gezien als een verfijning van het klassiek compatibilisme: enerzijds laat hij zien dat determinisme niet de contrastklasse is van vrije wil en anderzijds wordt aangetoond dat een handeling vrij is als deze niet wordt beperkt door externe én interne factoren. Morris vat deze opvatting over vrije wil als volgt helder samen:

“Since I perform certain actions that are the result of conscious choices made by me and by no one (or nothing) else, and since I am not unduly coerced in the performance of these actions, it is accurate to say that I sometimes exercise free will (...)”¹⁷

¹⁷ S. Morris, “Understanding Moral Responsibility within the Context of the Free Will Debate” in *Florida Philosophical Review* (Volume XII, nummer 1, 2012), 76.

De twee eerder gestelde voorwaarden waar aan moet worden voldaan voordat gesproken kan worden over een handeling die voortkomt uit de vrije wil, kunnen hierdoor verder als volgt worden gespecificeerd:

1. De handeling komt voort uit een bewuste keuze die de handelende persoon zelf heeft gemaakt.
2. De handelende persoon wordt niet gelimiteerd door interne of externe beperkingen tijdens het uitvoeren van de handeling.

Frankfurt

Ook Frankfurt verdedigt een vorm van compatibilisme, maar zijn aanpak verschilt behoorlijk van die van Ayer. Volgens Frankfurt ligt dat wat de mens onderscheidt van andere levende wezens namelijk in het feit dat onze wil gestructureerd is.¹⁸ Mensen *willen* iets ten aanzien van hun wil: ze staan er niet neutraal tegenover. Dit klinkt misschien wat abstract, maar het volgende voorbeeld kan verheldering bieden. Stel: je bent verslaafd aan roken, maar je zou er graag mee stoppen. Gelijktijdig zijn er hier twee verlangens actief: enerzijds het verlangen om een sigaret te roken, je bent immers verslaafd en snakt om de tien minuten naar een ‘peuk’. Anderzijds het verlangen om geen sigaret te roken, het roken gaat immers ten koste van je gezondheid en kost je daarnaast ook nog eens bergen met geld. Frankfurt zou beide verlangens eerste-orde verlangens¹⁹ noemen: dit zijn de verlangens om een bepaalde handeling wel of juist niet uit te voeren.

Echter sta je niet neutraal ten aanzien van deze twee verlangens. Je wil graag stoppen en daarom wil je graag dat het verlangen om *niet* te roken wordt verwezenlijkt in plaats van het verlangen om *wel* te roken. Dit soort tweede-orde verlangens noemt Frankfurt tweede-orde volities.²⁰ Tweede-orde volities zijn dus tweede-orde verlangens die specifiek zijn gericht op eerste-orde verlangens die handelingen produceren, zoals het wel of niet roken van een sigaret. Je hebt

¹⁸ Frankfurt, “Freedom of the will and the concept of a person”, 82.

¹⁹ Ibid., 83.

²⁰ Ibid., 86.

namelijk een bepaalde wil ten aanzien van een wil: je *wil* de wil om te roken niet hebben en je *wil* de wil om de sigaret links te laten liggen juist wel hebben.

Volgens Frankfurt handelt iemand uit *vrije* wil als diegene handelt naar de verlangens waarvan hij *wil* dat ze ook daadwerkelijk effectief zijn.²¹ Mocht je het restantje sigaretten in de prullenbak gooien, dan handel je volgens Frankfurt uit vrije wil: je doet immers waarvan je *wil* dat je het zou willen doen. Mocht je toch een sigaret opsteken, dan zou Frankfurt dit als onvrij bestempelen. Dit is namelijk niet wat je wil dat je zou willen doen. Tot slot is er de mogelijkheid om vrij noch onvrij te handelen. Dit is het geval als het je niet uitmaakt welke van de twee verlangens – het wel of het niet roken – wordt geactualiseerd. Dit kan volgens Frankfurt komen door ofwel het gebrek aan mentale capaciteiten om op het eigen gedrag te reflecteren of door achteloze onverschilligheid om eigen verlangens en motieven te evalueren.²²

Nu de positie van Frankfurt is besproken, zal in de komende twee hoofdstukken eerst worden gekeken in hoeverre de twee criteria die door Morris zijn herleid uit het werk van Ayer nader te bepalen zijn. In het volgende hoofdstuk (*Intentionele handelingen en bewuste keuzes*) zal worden gekeken in hoeverre een handeling altijd voortkomt uit een bewuste keuze. In het daaropvolgende hoofdstuk (*Mate van restricties*) zal onderzocht worden in welke mate interne of externe beperkingen restrictief kunnen zijn. Vervolgens zullen de voorwaarden worden toegepast op het experiment van Milgram. Tot slot komt de visie van Frankfurt opnieuw ter sprake en ook zijn inzichten zullen worden toegepast op het experiment van Milgram.

²¹ Slors, *Philosophy of Mind, Brain and Behaviour*, 192.

²² Frankfurt, “Freedom of the will and the concept of a person”, 89.

3. Intentionele handelingen en bewuste keuzes

Volgens het compatibilisme bestaat er niet zoiets als ‘absolute vrijheid’. Vrije wil is bij compatibilisme per definitie begrensd en daarmee niet absoluut vrij. De meest voor de hand liggende begrenzing is, zoals eerder besproken, het feit dat de wereld gedetermineerd is. Zoals Ayer heeft laten zien, vormt determinisme niet per se een probleem voor de vrije wil, maar uiteraard is niemand ‘vrij’ om de natuurkundige wetten te overschrijden. Daarnaast wordt vrijheid begrensd door de wil: de wil van een persoon komt voort uit diens gedachte en overtuigingen en deze zijn niet oneindig. Toch wordt binnen het compatibilisme gesproken over vrije handelingen die, buiten deze onvermijdelijke begrenzingen om, wel degelijk als volledig vrij kunnen worden beschouwd.

In de volgende twee hoofdstukken zal worden gekeken in hoeverre hiervan sprake is. Zijn er wel handelingen die worden gekenmerkt door volledige vrijheid of volledige onvrijheid? Misschien gaat het eerder om een *mate* van vrijheid: een graduele schaal van de vrije wil. Om dit te onderzoeken zullen beide door Morris gestelde criteria voor de vrije wil worden getoetst. In dit hoofdstuk zal gekeken worden naar het eerste criterium, waarin wordt gesteld dat een handeling altijd dient voort te komen uit een bewuste keuze die is gemaakt door de handelende persoon. Eerst zal het criterium verder worden geëxpliciteerd. Vervolgens zullen twee mogelijke bedreigingen besproken worden: onbewuste redenen en *priming*. Uiteindelijk zal duidelijk worden dat geen van de mogelijke bedreigingen ook daadwerkelijk een bedreiging vormen en kan geconcludeerd worden dat alle intentionele handelingen voortkomen uit een bewuste keuze.

Bewuste deliberatieprocessen

Nahmias heeft het expliciet over het eerste criterium dat hij de *Deliberation and Control (DC) condition*²³ noemt. Om dit toe te lichten zet hij een gedachte-experiment uiteen waarin een ‘confident agent’ centraal staat: een persoon die al haar keuzes altijd nauwkeurig afweegt en daar vervolgens naar handelt.²⁴ Voor deze denkbeeldige vrouw stelt hij een aantal criteria²⁵ vast:

1. Het is noodzakelijk dat ze deliberereert vóór iedere handeling en dat doet ze ook.
2. Ze is niet overtuigd van haar beslissing voordat ze heeft gedelibereerd.
3. Ze kan zichzelf identificeren met conflicterende verlangens, maar ze wil en hoeft daar niet per se naar te handelen.
4. Ze zal nooit tegen de uiteindelijke uitkomst van haar deliberaties in handelen.
5. Ze is geen fatalist: ze heeft de overtuiging dat haar deliberaties van invloed zijn op haar uiteindelijke handelen.

Het volgende voorbeeld zou van toepassing kunnen zijn met betrekking tot de door Nahmias gestelde criteria: Jasmijn twijfelt of ze op een donderdagavond de laatste zaken voor haar werk gaat regelen of op de uitnodiging van haar vriendinnen ingaat om een biertje te drinken in hun stamkroeg. Ze weet dat als ze de werkgerelateerde zaken niet vanavond regelt, morgen een erg drukke dag wordt en ze constant achter de feiten aan moet rennen. Haar vriendinnen heeft ze echter al een tijd niet gezien en het vooruitzicht van een koud pilsje vindt ze niet verkeerd. Uiteindelijk beslist ze dat de verminderde druk die ze op vrijdag op haar werk zal ervaren als ze al wat zaken regelt niet opweegt tegen de gezelligheid van een avondje uit met vriendinnen. Daarom kiest ze, na het afwegen van haar uiteindelijke beslissing, voor het laatste.

²³ E. Nahmias, “Close calls and the confident agent: free will, deliberation, and alternative possibilities” in *Philosophical Studies* (Springer 131, 2006), 627.

²⁴ *Ibid.*, 635.

²⁵ *Ibid.*, 636 – 637.

Het zou zo kunnen zijn dat Jasmijn op de terugweg naar huis van haar fiets valt en haar enkel breekt. Dit zegt niets over het deliberatie-proces en het maken van de uiteindelijke keuze om de deur uit te gaan en op de fiets naar de stamkroeg te stappen. Met andere woorden: het is niet zo dat ze meerdere keuzemogelijkheden had gewild vóórdát ze haar uiteindelijke keuze had gemaakt, maar dat ze had gewild dat ze beter had nagedacht voor het maken van haar beslissing. Om Nahmias' eigen woorden te gebruiken:

“In general, we want the possibility of different considerations coming to mind *not* so that we can choose otherwise in the *exact* same situation, but so that the situations would be *different* – such that we more informatively decide what to do. We wish that the situation leading to our choice had *not* been exactly the same, because we wish that *different* considerations had come to mind.”²⁶

Het hebben van meerdere keuze mogelijkheden is volgens Nahmias dus van ondergeschikt belang en het delibereren *an sich* staat centraal. Daarmee is de kous echter niet af. De vraag in hoeverre iemand de volledige controle heeft over zijn of haar eigen deliberatieprocessen komt namelijk op. In het geval van Jasmijn zou het bijvoorbeeld zo kunnen zijn dat ze, net voor het maken van haar beslissing, langs een café is gelopen of een reclame van Jupiler voorbij heeft zien komen. Dit zou vervolgens van invloed kunnen zijn op de uiteindelijke keuze om haar vriendinnen in de kroeg toch een bezoekje te brengen. De eerste voorwaarde die wordt gesteld aan vrije wil, het voortkomen van een handeling uit een *bewuste* keuze, komt hiermee mogelijk in gevaar.

²⁶ Ibid., 639.

Onbewuste redenen

Laten we kijken naar het eerste mogelijke gevaar: wat als intentionele handelingen voort kunnen komen uit onbewuste redenen? Vollmer argumenteert voor deze mogelijkheid²⁷ door een onderzoek aan te halen dat in de jaren '60 van de vorige eeuw is uitgevoerd door Latane en Darley. Tijdens dit experiment wordt een proefpersoon in de wachtkamer gezet met het verzoek om een vragenlijst in te vullen. Nadat de proefpersoon enkele minuten bezig is met het invullen van het formulier, wordt de kamer langzaam maar zeker gevuld met rook. De rook blijft de kamer in stromen totdat er na vier minuten dusdanig veel rook in de kamer aanwezig is dat het zicht wordt belemmerd en de rook op de ademhaling slaat.

Of de proefpersoon de ruimte verlaat en melding doet van de rook nadat deze door hem of haar is gedetecteerd, is sterk afhankelijk van de aanwezigheid van 'passieve derden'. Van de proefpersonen die alleen in de ruimte zaten en de rook opmerkte, heeft 85% de rook gerapporteerd. Van de proefpersonen die in de ruimte zaten met de 'passieve derden' rapporteerde echter maar 10% de rook nadat ze deze hadden ontdekt.

Hoewel de aanwezigheid van de anderen, zoals de percentages laten zien, duidelijk een reden vormt voor het intentionele gedrag (blijven zitten), confabuleren de proefpersonen hun redenen na afloop. Op de vraag waarom ze in de wachtkamer zijn blijven zitten, geven ze antwoorden als "the smoke did not represent dangerous fire, and (...) the proper course of action was to stay right where [I was]".²⁸ Daarnaast beweerden de proefpersonen dat ze geen of weinig aandacht hadden besteed aan de reacties van de andere 'proefpersonen' en deze geen invloed hadden gehad op hun handelen.

De aanwezigheid van de passieve bijstanders zorgt er dus (mede) voor dat de proefpersonen conclusies trekken als 'ik hoef niets te doen'. Deze bewuste overtuigingen zijn echter gevormd via onbewuste redenen. Om het in Vollmer's eigen woorden te formuleren:

²⁷ F. Vollmer, "Intentional action and unconscious reasons" in *Journal for the Theory of Social Behaviour* (Volume 23, nummer 3, 1993).

²⁸ Ibid., 321.

“While Ss [subjects] are aware of having these beliefs, they are not aware of having formed them *because* of the presence of others. And they are of course unaware of the nature of the connection between these beliefs and the presence of others.”²⁹

De intentionele handeling ‘blijven zitten’ komt voort uit een bewuste keuze (‘ik blijf zitten’), die weer voortkomt uit een overtuiging waar de proefpersoon zich bewust van is (‘ik hoef niets te doen’). De proefpersonen zijn zich er echter niet van bewust dat deze overtuigingen gevormd zijn *door* de aanwezigheid van anderen. De vraag is alleen of dit problemen oplevert voor het eerste criterium voor de vrije wil.

Het antwoord hierop is simpelweg: nee. Het criterium stelt immers dat de *keuze* die voorafgaat aan de handeling bewust moet zijn. In het geval van de proefpersoon die blijft zitten is niet alleen de keuze zelf bewust, zelfs de overtuiging die eraan voorafgaat is bewust. De redenen die daaraan voorafgaan, zijn in deze niet relevant. Stel dat ik onbewuste redenen heb om een man niet te mogen: het gezicht van de man is asymmetrisch, hij heeft een tic met zijn linker pink en hij komt soms net iets te dichtbij staan. Ik weet dat ik de man niet mag en daarom heb ik ook de bewuste overtuiging ‘ik mag deze man niet’.

Echter ben ik me er niet van bewust *waarom* ik deze man niet mag, ik weet slechts dat ik hem niet aardig vind. De redenen – het asymmetrische gezicht, de tic en het dichtbij komen staan – zijn onbewuste redenen voor de overtuiging ‘ik mag deze man niet’. Toch is het onzin om te zeggen dat als ik naar de man toe stap en hem in zijn gezicht sla, dit te ‘wijten’ valt aan de onbewuste redenen die eraan ten grondslag liggen. De bewuste en intentionele handeling ‘slaan’ komt immers voort uit de bewuste keuze ‘ik ga hem slaan’ die weer voortkomt uit de bewuste overtuiging ‘ik mag hem niet en daarom wil ik hem slaan’. Het slaan van

²⁹ Ibid., 322.

de man gebeurt daarom wel degelijk vanuit een bewuste keuze, ondanks de onbewuste redenen die daaraan vooraf zijn gegaan.

In eerste instantie lijkt hier iets tegenin te brengen: het slaan van de man is misschien helemaal niet zo uitvoerig uitgedacht als hierboven wordt beschreven en de bewuste overtuiging ‘ik mag hem niet en daarom wil ik hem slaan’ is wellicht minder bewust dan wordt gesuggereerd. Het slaan van man zou zelfs zo snel kunnen gebeuren dat ik er zelf – als ‘dader’ – van schrik. Het lijkt in dat geval misschien lastig om te stellen dat er aan deze intentionele handeling een bewuste keuze vooraf is gegaan.

Dit is echter niet het geval. Zelfs *als* bovenstaande het geval zou zijn, dan moet alsnog worden gesproken van een bewuste *keuze* om de man te slaan. Het is immers niet zo dat mijn hand reflexmatig uitschiet naar de man zijn gezicht, zoals een onderbeen vooruit beweegt als er net onder de knieschijf een tikje met een hamer wordt gegeven. *Ikzelf* moet de keuze *maken* om de man te slaan. Zelfs als deze keuze in een fractie van een seconde tot handelen wordt omgezet en ik ontsteld ben over mijn eigen gedrag. Uiteindelijk kan dan ook geconcludeerd worden dat het hebben van onbewuste redenen voor een uiteindelijke intentionele handeling geen gevaar vormt voor het eerste criterium.

Priming

Priming is een ander mogelijk probleem voor het eerste criterium voor vrije wil. Een helder voorbeeld van *priming* wordt gegeven door Di Nucci.³⁰ Hij haalt een onderzoek aan van de psycholoog J.A. Bargh, waarin een groep proefpersonen wordt geconfronteerd met woorden die hen onbewust aan bejaarden doen denken. Hierbij kan worden gedacht aan woorden als ‘lonely, grey, sentimental, wise, stubborn, bingo, ancient, helpless’.³¹ Opvallend genoeg liepen de proefpersonen die waren ‘geprimed’ met dit soort woorden een stuk langzamer uit de onderzoeksruijme dan de controlegroep: waar de controlegroep met een snelheid

³⁰ E. Di Nucci, “Priming Effects and Free Will” in *International Journal of Philosophical Studies* (Volume 20, nummer 5, 2012).

³¹ *Ibid.*, 726.

van 4.8 kilometer per uur de ruimte verliet, deed de onderzochte groep dit slechts met 4.2 kilometer per uur.

Bargh zelf trekt hieruit de conclusie dat de langzame proefpersonen niet, of op zijn minst minder, handelen uit vrije wil, in tegenstelling tot de controlegroep. Di Nucci trekt deze gevolgtrekking in twijfel en stelt dat de proefpersonen niet minder vrij zijn, ondanks de *priming*. Hiervoor gebruikt hij Davidson's formulering voor een bewuste, intentionele handeling:

S heeft handeling X intentioneel uitgevoerd als: (S een propositionele attitude heeft en de overtuiging dat het doen van X die propositionele attitude bevredigt) en (het verlangen plus de overtuiging zorgen ervoor dat S, X doet).³²

Voor de controlegroep is dit duidelijk het geval. Deze proefpersonen verlaten de ruimte met gemiddeld 4.8 kilometer per uur, omdat ze de propositionele attitude hebben om het gebouw binnen een bepaald tijdsbestek te verlaten en de overtuiging hebben dat dit lukt als ze in een bepaald tempo lopen. De controlegroep loopt dus bewust en intentioneel met een zekere snelheid het gebouw uit.

Dit is volgens Di Nucci niet anders bij de testgroep. Ook deze proefpersonen hebben een propositionele attitude om het gebouw binnen een bepaald tijdsbestek te verlaten en, net als de controlegroep, hebben ze de overtuiging dat ze deze propositionele attitude kunnen bevredigen door in een bepaald tempo het gebouw uit te lopen. Dus, om Di Nucci's woorden te gebruiken:

“Since primed subjects walk down the hallway at 4.2 km/h because of the relevant pro attitude plus belief, primed subjects walk down the hallway at 4.2 km/h *intentionally*.”³³

³² Ibid., 728.

³³ Ibid., 729.

Dat deze proefpersonen zijn ‘geprimed’ waardoor hun propositionele attitude anders is dan die van de controlegroep is niet relevant: het lopen *zelf* gebeurt namelijk bewust en intentioneel.

Ook hier lijkt een tegenwerping misschien op zijn plek. Mocht ik gehaast de voordeur achter me dicht trekken en tot de conclusie komen dat ik mijn portemonnee ben vergeten, dan loop ik waarschijnlijk gauw terug de trap op om deze alsnog te pakken. Hierbij wordt vermoedelijk niet iedere stap op een traprede volledig bewust gezet. Er zal namelijk geen sprake zijn van propositionele attitudes als ‘ik wil mijn voet op de volgende tree plaatsen door mijn been op te tillen en deze weer neer te zetten’: bij de meeste personen gaan handeling als traplopen immers automatisch.

Bovenstaande is echter niet problematisch om het traplopen als een bewuste en intentionele handeling te bestempelen. Als de door Davidson gestelde definitie voor het bewust en intentioneel handelen wordt gehandhaafd, dan blijft er ook bij het snel de trap oplopen sprake van een bewuste en intentionele handeling. *Zelfs* als het stappen op individuele treden en het optillen van het been vanzelf gebeurt. Allereerst heb ik namelijk het verlangen snel mijn portemonnee te pakken. Daarnaast heb ik de propositionele attitude en de overtuiging dat als ik haastig de trap opren, ik snel mijn vergeten portemonnee te pakken zal krijgen. De combinatie van de propositionele attitude en de overtuiging zorgen er vervolgens voor dat ik haastig de trap opren om mijn portemonnee te pakken.

Ook al ben ik me niet volledig bewust van mijn motoriek en de afzonderlijke treden, toch gebeurt het traplopen wel degelijk bewust en intentioneel. Overeenkomstig geldt dat het tragere lopen van de proefpersonen wel degelijk bewust en intentioneel gebeurt, ook al zijn de proefpersonen uit het experiment van Bargh zich niet bewust van hun daadwerkelijke snelheid. Dat hun propositionele attitude verschilt van die van de controlegroep – de proefpersonen willen immers in een langer tijdsbestek dezelfde afstand afleggen – neemt niets weg van het feit dat het lopen van de proefpersonen wel degelijk bewust en intentioneel gebeurt. De conclusie luidt daarom dat ook *priming* geen gevaar

vormt voor het eerste criterium van de vrije wil, waarmee wordt gesteld dat een handeling moet voortkomen uit een bewuste keuze.

4. Mate van restricties

In het vorige hoofdstuk hebben we gezien dat een intentionele handeling *altijd* voortkomt uit een bewuste keuze. In dit hoofdstuk zal het tweede criterium onder de loep worden genomen: de handelende persoon wordt niet begrensd in zijn of haar handelen. Aan de hand van een specificering van interne en externe beperkingen zal worden gekeken of de mate van vrijheid kan worden bepaald. Hiervoor zal een poging van Costa³⁴ worden beschouwd die Ayer's notie van interne en externe beperkingen nog verder opdeelt in twee vormen van restricties die zich op drie verschillende niveaus kunnen voordoen. Uiteindelijk zal blijken dat Costa's argument, anders dan hij beweert, niet tot harde maatstaven voor het tweede criterium leidt.

Begrenzingen en beperkingen

Zoals eerder besproken is een handeling volgens Ayer vrij als er geen sprake is van interne of externe beperkingen. Ook volgens Costa is er sprake van vrije wil als er een geen sprake is van restricties.³⁵ In zekere zin is dit vergelijkbaar met Ayer's standpunt: er is sprake van vrije wil als er geen sprake is van interne of externe beperkingen. Costa specificeert deze interne en externe beperkingen, deze restricties, echter nog verder dan Ayer. Hierbij toont hij eerst aan dat restricties kunnen worden onderscheiden in begrenzingen en beperkingen. Vervolgens laat hij zien dat er verschillende niveaus zijn van restrictie.

Laten we eerst kijken naar het onderscheid tussen begrenzingen en beperkingen. Stel dat ik in een lage en smalle gang sta. Deze gang is net zo hoog als dat ik lang ben: 1 meter 67. De gang is daarnaast net zo breed als dat ik breed ben, namelijk een centimeter of 40. In deze specifieke gang heb ik dus de vrijheid om me voor- en achteruit te bewegen. Omhoog kan ik gezien de gegeven

³⁴ C. Costa, "Free will and the soft constraints of reason" in *Ratio* (Volume 19, nummer 1, 2006).

³⁵ Uiteraard wordt causale determinatie door Costa niet als restrictie gezien. Zie voor verdere uitleg de paragraaf over Ayer in hoofdstuk 2.

omgeving immers niet en naar rechts of links bewegen wordt ook lastig. Deze vrijheid om me voor- en achteruit te bewegen kan worden ingeperkt door een begrenzing, namelijk als er een blokkade of belemmering in de weg staat. Dit is bijvoorbeeld het geval wanneer er een groot rotsblok vóór mij in de gang verschijnt, waardoor ik me alleen nog maar achteruit door de gang kan bewegen. Er kan ook sprake zijn van een beperking, die wordt gekenmerkt door dwang. Bijvoorbeeld wanneer een sterker persoon mij van achteren door de gang heen naar voren duwt. Hierdoor verdwijnt de mogelijkheid mij naar achteren te verplaatsen en kan ik me slechts nog naar voren voortbewegen.

Dit onderscheid tussen begrenzingen en beperkingen beperkt zich niet tot het fysieke domein. Costa geeft het volgende voorbeeld om dit te illustreren³⁶: stel dat je in een boekwinkel staat en graag een bepaald boek wil kopen. Je komt er echter achter dat je niet genoeg geld (bij je) hebt, waardoor er sprake is van een begrenzing. De gangbare keuzemogelijkheden, het wel of niet kopen van het boek, zijn er namelijk niet meer. Er kan ook sprake zijn van beperking in het boekwinkelscenario. Dit is bijvoorbeeld het geval wanneer jouw werkgever de winkel in loopt die zelf vindt dat hij een waanzinnig boek heeft geschreven, terwijl je daar zelf toch net iets anders over denkt. Het kan zijn dat je je, gezien de sociale verhoudingen, gedwongen voelt zijn boek op dat moment te kopen.

Nu de twee verschillende vormen van restricties uiteen zijn gezet, wordt het tijd om te kijken naar de drie verschillende niveaus van restricties. Deze zijn het fysieke, het intentionele en het rationele niveau. Het is hierbij belangrijk te vermelden dat deze niveaus zijn afgeleid van *the causal theory of action*³⁷, die stelt dat (rationele) handelingen eerst een causaal proces hebben doorstaan. Zo komen handelingen voort uit een rationele overweging, die bestaat uit verlangens en overtuigingen.

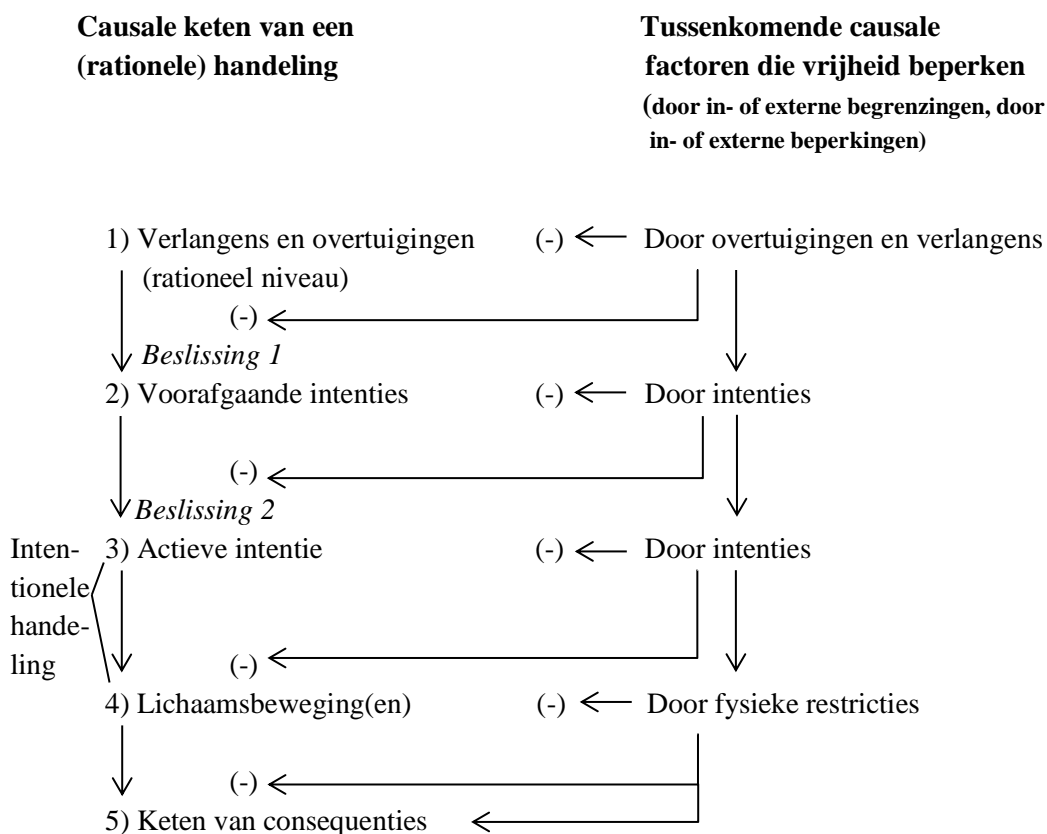
Een voorbeeld van zo'n causaal proces is het verlangen van Tom om het fortuin van zijn oom te erven en de daaraan gekoppelde overtuiging dat de snelste manier om dit geldbedrag op zijn rekening te krijgen, gerealiseerd kan worden

³⁶ Costa, "Free will and the soft constraints of reason", 6 – 7.

³⁷ Ibid., 11.

door het ombrengen van die oom. Dit resulteert in een voorafgaande intentie: Tom *beslist* om zijn oom ook daadwerkelijk te gaan vermoorden. Vervolgens wordt deze voorafgaande intentie omgezet in de actieve intentie dat Tom *probeert* zijn oom te vermoorden. Dit heeft direct lichaamsbeweging tot gevolg: Tom schiet met een vuurwapen op zijn oom. Uiteindelijk heeft dit een hele keten aan consequenties tot gevolg. Bijvoorbeeld het sterven van de oom, Tom die het fortuin erft, et cetera.³⁸

Het bijzondere van Costa's artikel is niet de uiteenzetting van deze *causal theory of action*, maar zijn standpunt dat de mogelijke restricties hetzelfde causale karakter hebben als de causale keten die zojuist is besproken. Dit klinkt misschien nog wat abstract, maar het schema³⁹ hieronder kan hierin verhelderend werken. Hierbij geven de pijlen causale relaties aan en laten de mintekens zien waar restricties zich mogelijk voor kunnen doen.



³⁸ Ibid., 11.

³⁹ Ibid., 12.

Een belangrijke notie is dat het karakter van de restrictie, zoals ook te zien is in het schema, overeen moet komen met het karakter van *wat* het beperkt.⁴⁰ Zo kunnen fysieke restricties alleen lichaamsbeweging(en) beperken en hebben ze geen (mogelijke) beperkende invloed op het rationele of intentionele niveau. Tot slot moet het volgende in acht worden genomen:

“... the levels of restrictions are interdependent: restrictions on the volitional level, coming earlier in the causal chain, compromise the freedom of action too, while restrictions on the level of reasoning compromise both of them and restrictions on the level of action compromise none of them causally...”⁴¹

Mocht er een restrictie plaatsvinden op het rationele niveau, dan heeft dit dus ook indirect invloed op de intenties, lichaamsbeweging(en) en de keten van consequenties. Mocht de restrictie zich op het fysieke niveau (lichaamsbeweging(en)) voordoen, dan heeft dit niet met terugwerkende kracht invloed op de voorafgaande verlangens, overtuigingen en intenties.

Onduidelijke criteria

Hoewel Costa de mogelijke interne en externe begrenzings- en beperkingen al verder weet te specificeren, blijven er vragen liggen over de grenzen van de vrije wil. Zo beweert Costa onterecht dat, door zijn gegeven definities, vrijwel alle situaties kunnen worden beoordeeld op het al dan niet aanwezig zijn van restricties op (ten minste) één van de niveaus. Een voorbeeld dat hij hiervoor geeft is dat van een vrouw die een groot deel van haar geld besteedt aan overbodige aankopen.⁴² Zij komt later tot de conclusie dat zij deze aankopen eigenlijk liever niet had gedaan, maar zich door haar sociale milieu had laten sturen om dit toch te

⁴⁰ Ibid., 19.

⁴¹ Ibid.

⁴² Ibid., 21.

doen. Volgens Costa kan zijn theorie wel degelijk laten zien dat deze vrouw niet vrij was toen zij haar beslissing maakte om de spullen te kopen, in tegenstelling tot meer gangbare theorieën voor compatibilisme. De vrouw zou namelijk op het rationele en intentionele niveau zijn beïnvloed, aldus Costa, met het doen van overbodige aankopen tot gevolg.

Dit geeft echter nog geen heldere scheidslijn aan. Stel dat de vrouw haar geld niet in grote getalen had uitgegeven aan overbodige aankopen als tasje, schoenen en make-up, maar aan een goed doel. Stel tevens dat zij – zoals ook van toepassing in het voorbeeld van Costa – achteraf spijt zou hebben van het uitgeven van haar geld en tot de conclusie zou komen dat zij dusdanig was beïnvloed door haar linkse omgeving dat ze een keuze heeft gemaakt die niet (volledig) is voortgekomen uit haar vrije wil. Volgens Costa zou de vrouw ook in dit geval zijn beperkt op rationeel en intentioneel niveau.

Stel nu dat deze zelfde dame over straat loopt en wordt benaderd door een vriendelijke man die haar om vijftig euro vraagt om in een restaurant een drie-gangen menu te kunnen nuttigen. De vrouw is in een bijzonder vrijgeevige bui en besluit de man de gevraagde vijftig euro te overhandigen. Later komt zij tot de conclusie dat ze spijt heeft van haar beslissing: ze was beïnvloed door het zomerse weer, de vriendelijkheid van de man en een volle portemonnee waardoor dat ene briefje meer of minder toch niet zo heel veel leek uit te maken. Kortom: op rationeel niveau werd ze in zekere mate beperkt, waardoor haar uiteindelijke beslissing om de man het briefje van vijftig te geven niet volledig uit haar eigen vrije wil was ontsproten.

Het eerdere voorbeeld met het doneren aan het goede doel verschilt in principe niets van het laatste voorbeeld: in beide gevallen wordt de vrouw beïnvloed door haar directe omgeving, waardoor ze later concludeert dat ze op rationeel niveau was beperkt en niet op basis van haar vrije wil heeft gehandeld. Toch lijkt het voor de hand te liggen om de vrouw in het laatste geval te verwijten dat ze zelf niet goed heeft nagedacht voordat ze het briefje overhandigde en dat ze zich verschuilt achter deze zogenaamde beperkingen om haar eigen handelen goed te kunnen praten.

De vraag is dus wanneer rationele restricties *daadwerkelijk* als beperking van de vrije wil kunnen worden gezien. Tot op welke hoogte kun je je beroepen op je opvoeding, standplaatsgebondenheid of (sociale) omgeving voor het maken van je keuzes? Waarom zou de persoon die geld gaat pinnen zonder ogenschijnlijke interne en externe beperkingen handelen uit vrije wil en de geld spenderende vrouw niet? Ook het voornemen om geld op te nemen is mede *bepaald* door ingebakken rationele overtuigingen en verlangens die culturele en sociale gronden hebben. Zoals ‘ik geloof dat het sociaal wenselijk is om geld uit te geven om spullen te kopen’, ‘ik verlang naar meer spullen’, ‘ik geloof dat ik, door te pinnen, geld kan vergaren waarmee ik spullen kan kopen’, et cetera.

Costa weet te benoemen in welke vormen restricties zich voor kunnen doen, maar daarmee is de kous nog niet af. *Wanneer* deze vormen als restrictief kunnen worden beschouwd, blijft in het artikel van Costa nog de vraag. De demarcatie van de tweede voorwaarde van vrije wil is geen hard en helder onderscheid, maar meer een soort vage grens. Het is immers onmogelijk om te achterhalen wat iemands wil exact heeft beïnvloed, ook voor de persoon in kwestie zelf. De mentale toestanden die op de achtergrond meespelen zijn niet totaal ‘vrij’ en in zekere mate geconditioneerd: ze zijn gebonden aan de cultuur waarin je leeft, je sociale omgeving, je opvoeding et cetera. Het is dan ook niet zo dat er sprake kan zijn van een volledig gebrek aan interne of externe beperkingen.

Conclusie

Deze twee hoofdstukken zijn begonnen met de twee voorwaarden voor een handeling die voortkomt uit vrije wil:

1. De handeling komt voort uit een bewuste keuze die de handelende persoon zelf heeft gemaakt.
2. De handelende persoon wordt niet gelimiteerd door interne of externe beperkingen tijdens het uitvoeren van de handeling.

Zoals we hebben gezien komt een intentionele handeling altijd voort uit een bewuste keuze. Hoewel de mogelijkheid van onbewuste redenen voor een

intentionele handeling en *priming* op het eerste gezicht problematisch lijken voor het eerste criterium, is aangetoond dat dit niet het geval is. Het tweede criterium is echter lastiger vast te stellen: *wanneer* factoren restrictief zijn, is moeilijk te bepalen. Niemand is immers *volledig* gevrijwaard van restricties. Met andere woorden: niemand is volledig autonoom. Niemand staat ‘op zichzelf’ of maakt bewuste keuzes die ‘op zichzelf’ staan. Ieder mens denkt namelijk binnen een bepaald kader dat voortkomt uit omgeving, opvoeding en opgedane ervaringen en dit is van invloed op de uiteindelijke bewuste keuzes die worden gemaakt. De notie van vrije wil moet dan ook worden gezien als een graduele schaal: de *mate* van restricties bepalen de *mate* van vrijheid van een bepaalde bewuste, intentionele handeling.

5. Milgram's experiment: I

In de vorige hoofdstukken is geconcludeerd dat aan een intentionele handeling altijd een bewuste keuze voorafgaat. Dit betekent nog niet dat handelingen als volledig vrij of volledig onvrij bestempeld kunnen worden. De *mate* van vrijheid wordt namelijk bepaald door de *mate* van restrictie. In dit hoofdstuk zal aan de hand van het experiment van Milgram worden geprobeerd om een graduele schaal van vrije wil nauwkeuriger op te stellen.

Allereerst zal de onderzoeksmethode van het 'klassieke' experiment uiteen worden gezet en zullen de resultaten van deze variant alsmede van andere variaties worden bekeken. Daarna wordt vastgesteld dat mogelijke interne beperkingen niet nader te bepalen zijn. Tot slot zal worden gekeken naar de externe beperkingen per variatie en zal een graduele schaal worden gepresenteerd waarin zal worden aangetoond dat sommige proefpersonen meer en andere minder vrij waren op het moment dat ze de maximale schok toedienden.

Onderzoeksmethode⁴³

Het onderzoek is op een relatief simpele manier opgesteld. In het 'klassieke' onderzoek participeren drie personen: de proefpersoon, de onderzoeker en een tweede 'proefpersoon' die in werkelijkheid een acteur is en bij het onderzoek hoort. De proefpersonen krijgen als beloning voor het participeren aan het onderzoek op de Yale Universiteit \$4,00, *ongeacht het verloop of de uitslag van het onderzoek*. Er wordt ze verteld dat Yale een theorie wil toetsen, die stelt dat mensen beter leren als ze worden gestraft voor het maken van een fout.

Vervolgens worden er twee rollen zogenaamd willekeurig verdeeld onder de twee proefpersonen: de rol van 'leerling' en de rol van 'leraar'. In werkelijkheid is deze verdeling geënceneerd: op beide briefjes staat de rol van 'leraar', waardoor de daadwerkelijke proefpersoon deze rol altijd op zich zal nemen.

⁴³ Milgram, *Obedience to Authority*, 13-26.

Hierna worden zowel leerling als leraar meegenomen naar een kamer. Daar wordt de leerling geïnstalleerd op een stoel en vakkundig ingesnoerd. Vervolgens krijgt de leerling een elektrode op zijn pols geplaatst die wordt ingesmeerd met een vettig smeersel: “to avoid blisters and burns.”⁴⁴ Tot slot wordt de proefpersoon verteld dat de geplaatste elektrode is verbonden met de *shock generator* in de aangrenzende kamer en dat de schokken erg pijnlijk kunnen zijn, maar geen permanente schade veroorzaken. De acteur-proefpersoon is natuurlijk niet werkelijk aan het apparaat aangesloten, maar de daadwerkelijke proefpersoon heeft de overtuiging dat dit wel degelijk het geval is.

Als de leerling is geïnstalleerd, wordt de proefpersoon meegenomen naar een andere ruimte. Hier is een indrukwekkende machine klaargezet met dertig schakelaars. Boven iedere schakelaar staat een voltage, oplopend van 15 volt tot en met 450 volt. Boven groepjes van vier schakelaars staan verbale aanwijzingen als *moderate shock*, *intense shock*, *Danger: severe shock* en boven de laatste twee schakelaars staat slechts *XXX*. De proefpersoon krijgt zijn rol als leraar verder toegelicht. Hij dient woordparen voor te lezen aan de leerling, bijvoorbeeld *blue box*, *nice day*, *wild duck* et cetera. Daarna leest hij een eerste woord voor, gevolgd door vier mogelijke antwoorden. Bijvoorbeeld: *blue: sky ink box lamp*.

De leerling dient via vier schakelaars, ieder corresponderend met een antwoord, het goede antwoord door te geven door één van de schakelaars om te zetten. Als de leerling niet de juiste schakelaar omzet, in dit geval de schakelaar die ‘box’ vertegenwoordigt, dan krijgt de leerling een schok toegediend. Voor ieder fout antwoord, krijgt de leerling een schok. Bovendien krijgt de leraar de instructie “to move one level higher on the shock generator each time the learner gives a wrong answer.”⁴⁵ Vervolgens krijgt de leraar zelf één schok van 45 volt toegediend.

⁴⁴ Ibid., 19.

⁴⁵ Ibid., 21.

Gedurende het experiment richtten de proefpersonen zich tot de onderzoeker om advies te vragen of om aan te geven niet door te willen gaan. De onderzoeker werd geacht de volgende antwoorden⁴⁶ te geven:

1. “Please continue” of “Please go on”.
2. “The experiment requires that you continue.”
3. “It is absolutely essential that you continue.”
4. “You have no other choice, you *must* go on.”

De onderzoeker werd geïnstrueerd deze volgorde aan te houden. Dat wil zeggen dat alleen als antwoord 1 geen effect bleek te hebben, antwoord 2 mocht volgen. Als de proefpersoon na het vierde antwoord nog steeds weigerde te gehoorzamen, dan werd het onderzoek beëindigd. Ook de acteur-proefpersoon kreeg nauwgezette instructies mee. Zo liet hij bijvoorbeeld bij een schok van 75 volt een grom horen die zijn ongemak moest illustreren, vanaf 120 volt schreeuwde hij naar de onderzoeker dat de schokken pijnlijk werden, bij 180 volt riep hij “I can’t stand the pain!” en vanaf 330 volt veinsde de leerling bewusteloos te zijn. Na het onderzoek volgde een debriefing.

Resultaten

De resultaten van het ‘klassieke’ onderzoek van Milgram waren onthutsend. Aanvankelijk werd verwacht dat ongeveer 2.5% van de proefpersonen een ‘gehoorzaam subject’ zou zijn. Hiermee worden de proefpersonen bedoeld die uiteindelijk doorgaan tot het uiterste voltage van 450 volt. Een voltage dat dodelijk kan zijn. Het kleine percentage dat tot het einde door zou gaan, zouden psychopaten zijn met geen tot weinig empathisch vermogen.⁴⁷ Veel proefpersonen, zo werd verwacht, zouden überhaupt weigeren ook maar één knopje om te zetten. De uitkomsten van het experiment lieten echter iets anders

⁴⁶ Ibid.

⁴⁷ P. Goldie, *The Emotions: A Philosophical Exploration*. (Oxford: Oxford University Press, 2002), 180.

zien: 62.5% van de proefpersonen gingen door tot het maximale voltage en de eerste proefpersoon stopte pas bij 135 volt.⁴⁸

Naast deze ‘klassieke’ versie van het experiment zijn er veel variaties uitgevoerd. Tenzij anders vermeld, zijn verder alle omstandigheden gelijk aan die van het klassieke experiment met uitzondering van de besproken variabele(n). Allereerst het *afgezonderde* experiment, waarbij de kreten van het slachtoffer niet te horen waren: de proefpersoon kreeg geen enkele (vocale) feedback van het slachtoffer. Er is ook een *nabijheids*-experiment uitgevoerd, waarbij het slachtoffer in dezelfde ruimte werd geplaatst als de proefpersoon. Daarnaast kreeg het slachtoffer alleen een schok als zijn hand op een schokplaat rustte. Vanaf 150 volt eiste het slachtoffer te worden bevrijd en werd de proefpersoon geacht de hand van het slachtoffer te pakken en deze op de plaat te leggen.

In het *afwezige onderzoeker* experiment gaf de onderzoeker vóór het experiment instructies. Tijdens het experiment kregen de proefpersonen alleen via de telefoon orders van de onderzoeker. Er is ook een onderzoek uitgevoerd waarbij er niet één, maar twee onderzoekers in de ruimte van de proefpersoon aanwezig waren die elkaar tegenspraken. De één vond dat het onderzoek gestopt moest worden, de ander dat ze door moesten gaan. Deze versie noemen we het *twee autoriteiten* experiment. Tot slot het *proefpersoon kiest* experiment: hierbij staat het de proefpersoon vrij om een voltage te kiezen bij een fout antwoord van de leerling. Het was toegestaan om gedurende het hele onderzoek hetzelfde voltage toe te dienen of om de hoogte van het voltage te variëren.

Dit zijn niet de enige variaties die zijn uitgevoerd. Voor de overzichtelijkheid is er echter voor gekozen het aantal tot deze zes experimenten te beperken. Op de volgende pagina staan de resultaten⁴⁹ van deze experimenten weergegeven.

⁴⁸ Milgram, *Obedience to Authority*, 35.

⁴⁹ Ibid., 35, 60, 61 en 95.

	Experiment	Percentage gehoorzame proefpersonen⁵⁰
1.	Afgezonderd	65%
2.	Klassiek	62.5%
3.	Nabijheid	30%
4.	Afwezige onderzoeker	20%
5.	Twee autoriteiten	10%
6.	Proefpersoon kiest	2.5%

Interne restricties

Zoals we in het tweede hoofdstuk hebben gezien gaat aan een intentionele handeling altijd een bewuste keuze vooraf. Hierdoor kunnen we concluderen dat aan de eerste voorwaarde voor de vrije wil is voldaan: alle gehoorzame proefpersonen maakten een bewuste keuze toen ze de genadeklap van 450 volt toebachten. Het is immers niet zo dat ze reflexmatig de knopjes omzetten. Toch is aan de ver uit elkaar liggende uitslagen van de experimenten te zien dat er omstandigheden zijn die de uiteindelijke intentionele handeling – het al dan niet omzetten van de 450 volt – kunnen beïnvloeden. De mate van vrijheid moet daarom worden bepaald aan de hand van mogelijke interne en externe restricties.

De mate van interne beperkingen is echter lastig te bepalen. De mentale toestanden van iemand zijn immers moeilijk te achterhalen: er bestaat geen mechaniek dat in iemands hersenen kan worden geplaatst en dat vervolgens

⁵⁰ In het geval van het zesde experiment kan niet gesproken worden van gehoorzame subjecten. De proefpersonen mogen het voltage immers zelf bepalen: er zijn geen op te volgen bevelen.

precies registreert welke overtuigingen die persoon heeft. Ook het herleiden van overtuigingen uit gedrag is geen waterdichte methode. Als mensen over het algemeen een glas water pakken *omdat* ze de overtuiging hebben dat ze graag water willen drinken, dan betekent dit niet dat als iemand een glas water pakt diegene per definitie wat wil drinken. Misschien pakt iemand wel een glas water om de inhoud in iemand anders zijn gezicht te gooien of om de cactussen water te geven. Gedrag kan een goede aanduiding zijn voor bepaalde gedachten en overtuigingen die vooraf zijn gegaan aan dat gedrag, maar die indicatie is niet altijd correct.

De interne processen zijn dus moeilijk vast te stellen. Daarnaast zijn deze processen, zoals eerder gesteld, niet volledig autonoom. Ieder mens is beïnvloed door zijn standplaatsgebondenheid, opvoeding, opleiding en ervaringen door het leven heen. Al deze zaken zijn van invloed op de mentale processen en keuzes die daaruit voortvloeien, maar kunnen onmogelijk worden meegenomen. Dit klinkt in eerste instantie misschien behoorlijk problematisch: intuïtief lijkt het noodzakelijk om interne processen vast te stellen als het om sociale beïnvloeding gaat. Het zijn in het geval van het experiment immers de gedachten en verlangens van de proefpersoon die wijzigen door (onder andere) de orders van de onderzoeker.

Dat de orders invloed hebben op de gedachten en verlangens van de proefpersoon, moet echter niet worden gezien als een interne *beperving*. Het is namelijk een *externe* beperking die de interne processen tot gevolg heeft. Neem ter verheldering het volgende voorbeeld. Stel, ik sta weer in de lange gang die net zo lang is als dat ik groot ben en net zo breed is als dat ik breed ben. Terwijl ik me door de gang voortbeweeg, verschijnt er op een bepaald moment een rotsblok. Dit rotsblok kan ik met moeite passeren, maar niet zonder er met kleerscheuren vanaf te komen. Ongetwijfeld heeft dit invloed op mijn mentale toestanden: zo zal ik gedachten hebben als ‘ik kan niet zonder pijn rechtdoor, maar ik wil niet terug achteruit. Wat zal ik nu doen?’. Vervolgens zal ik een keuze maken om toch terug de gang door te gaan of me langs het rotsblok te wurmen.

Dit soort mentale toestanden vormt echter geen *beperving*. De daadwerkelijk beperkende factor is een externe factor, namelijk het rotsblok. Het

rotsblok zorgt er immers voor dat ik me niet in alle vrijheid door de gang kan bewegen, zoals ik zou willen. Mijn interne mentale toestanden zijn in deze niet restrictief. Hetzelfde geldt voor de mentale toestanden van de proefpersonen. Door de orders van de onderzoeker worden er waarschijnlijk interne processen in gang gezet, waardoor er sprake is van mentale toestanden als: ‘ik wil eigenlijk geen schok toedienen, maar de onderzoeker blijft zo aandringen dat ik het misschien toch maar moet doen.’ Het zijn – net als bij de situatie in de gang – echter niet deze interne mentale toestanden die beperkend zijn. De beperkende factoren zijn extern: (onder andere) de aanwezigheid en orders van de onderzoeker.

Dit betekent niet dat er nooit sprake kan zijn van een interne beperking. Hier is wel degelijk sprake van als iemand kampt met psychische stoornissen, zoals schizofrenie of OCD. Op basis van de interviews met de proefpersonen en de waarnemingen van de onderzoekers kunnen we er echter vanuit gaan dat geen van de proefpersonen kampten met ernstige psychische problemen en daardoor zijn ‘ge vrijwaard’ van interne beperkingen in de sterke zin van het woord. Daarnaast zijn de experimenten in verschillende landen – verspreid over de hele wereld – veelvuldig herhaald met vergelijkbare resultaten.⁵¹

Externe restricties

De externe beperkingen zijn daarentegen wel goed te bepalen. Deze zijn immers vastgelegd door Milgram en zijn onderzoeksteam. Daarom zal nu per experiment gekeken worden naar externe factoren die de uiteindelijke keuze voor het toedienen van de 450 volt hebben kunnen beïnvloeden.

Het afgezonderde experiment

- Het onderzoek vindt plaats op Yale University; een gerenommeerde universiteit.

⁵¹ Milgram, *Obedience to Authority*, 170 – 171.

- Proefpersoon weet dat de leerling zich – net als hij – heeft gecommitteerd aan het experiment.
- De proefpersoon bevindt zich in een situatie waarin wordt gesteld dat het pijnigen van een ander legitiem is (ten behoeve van het onderzoek).
- De proefpersoon krijgt de instructie om het voltage op te laten lopen per fout antwoord.
- Er is een onderzoeker in ruimte aanwezig die de verantwoordelijkheid op zich zegt te nemen en het belang van het continueren van het onderzoek benadrukt.

Het klassieke experiment

- Alle externe factoren uit het *afgezonderde experiment*
- De proefpersoon krijgt reactie van de leerling die zegt pijn te hebben. Bij 330 volt en hoger ontbreekt iedere reactie van de leerling.

Het nabijheids-experiment

- Alle externe factoren uit het *afgezonderde experiment*
- De proefpersoon zit in dezelfde ruimte als de leerling.
- De proefpersoon dient de leerling fysiek aan te raken om de schokken toe te kunnen dienen.

Het afwezige onderzoeker experiment

- Het onderzoek vindt plaats op Yale University; een gerenommeerde universiteit.
- Proefpersoon weet dat de leerling zich – net als hij – heeft gecommitteerd aan het experiment.
- De proefpersoon bevindt zich in een situatie waarin wordt gesteld dat het pijnigen van een ander legitiem is (ten behoeve van het onderzoek).
- De proefpersoon krijgt de instructie om het voltage op te laten lopen per fout antwoord.

- De proefpersoon krijgt reactie van de leerling die zegt pijn te hebben. Bij 330 volt en hoger ontbreekt iedere reactie van de leerling.
- Via de telefoon geeft de onderzoeker verdere instructies. Hij is niet fysiek aanwezig.

Het twee autoriteiten experiment

- Het onderzoek vindt plaats op Yale University; een gerenommeerde universiteit.
- Proefpersoon weet dat de leerling zich – net als hij – heeft gecommitteerd aan het experiment.
- De proefpersoon bevindt zich in een situatie waarin – in eerste instantie – wordt gesteld dat het pijnigen van een ander legitiem is (ten behoeve van het onderzoek).
- De proefpersoon krijgt – in eerste instantie – de instructie om het voltage op te laten lopen per fout antwoord.
- De proefpersoon krijgt reactie van de leerling die zegt pijn te hebben. Bij 330 volt en hoger ontbreekt iedere reactie van de leerling.
- De onderzoekers zijn het met elkaar oneens. Ze uiten hun tegengestelde opvattingen over het voortzetten van het onderzoek in dezelfde ruimte als de proefpersoon.

Het proefpersoon kiest experiment

- Het onderzoek vindt plaats op Yale University; een gerenommeerde universiteit.
- Proefpersoon weet dat de leerling zich – net als hij – heeft gecommitteerd aan het experiment.
- De proefpersoon bevindt zich in een situatie waarin wordt gesteld dat het pijnigen van een ander legitiem is (ten behoeve van het onderzoek).
- De proefpersoon krijgt reactie van de leerling die zegt pijn te hebben. Bij 330 volt en hoger ontbreekt iedere reactie van de leerling.

In alle experimenten zijn drie dezelfde factoren aanwezig: het onderzoek vindt plaats op een gerenommeerde universiteit, de leerling heeft zich gecommitteerd aan het onderzoek en de proefpersoon wordt in een situatie gebracht waarin wordt gesteld dat het pijnigen van een ander legitiem is. Deze factoren lijken de grote verschillen tussen de uitkomsten van de experimenten daarom moeilijk te kunnen verklaren.

Een eerste bepalende externe factor kan worden gezien in het verschil tussen het *klassieke* en het *afgezonderde* experiment. Hoewel de persoon in het *afgezonderde* experiment zich misschien kan bedenken dat een schok van 450 volt misschien niet heel bevorderlijk voor de gezondheid is, krijgt hij dit op geen enkele manier terug van de leerling. Hij kan de leerling immers niet zien of horen. Hierdoor speelt de factor van de kreten van een gepijnigde leerling voor hem niet mee.

Aan de resultaten is af te lezen dat het percentage gehoorzame subjecten zakt als de proefpersoon reactie krijgt van de leerling. Het nog dichterbij halen van de leerling – waarbij de proefpersoon de leerling zelfs aan dient te raken – heeft een nog grotere afname van gehoorzame subjecten tot gevolg. Hieruit kan worden afgeleid dat, als het gaat om beperkende externe factoren, er sprake is van een grotere mate van vrijheid als hij besluit de leerling de volle laag te geven op het moment dat hij *wel* reactie krijgt van de leerling. Deze mate van vrijheid is nog groter als de proefpersoon naast hem zit en hij deze aan dient te raken.

Door deze bepalende externe factor wordt de proefpersoon immers gedwongen zich te verhouden tot de (geveinsde) pijn en kreten van de leerling en het omzetten van de volgende schakelaar te heroverwegen. Bij het *afgezonderde* experiment is dit niet het geval. De proefpersoon hoeft zich daar slechts te verhouden tot de onderzoeker: de leerling is zowel auditief als fysiek afwezig en vormt geen – of hooguit een zeer geringe – externe factor.

Ook bij het *twee autoriteiten* en het *autoriteit op afstand* experiment kan worden gesproken van een grotere mate van vrijheid. Door de aanwezigheid van twee autoriteiten die contradictoire orders geven, wordt de proefpersoon gedwongen zelf te delibereren en zich te verhouden tot één van de autoriteiten. Hij

kan niet *beiden* orders tegelijkertijd opvolgen; deze zijn immers tegengesteld aan elkaar. Bij het experiment *autoriteit op afstand* is de onderzoeker niet fysiek in dezelfde ruimte aanwezig en geeft deze slechts via telefonisch contact instructies door aan de proefpersoon. De mate van externe beperking neemt hierdoor nog verder af: de fysieke aanwezigheid van de onderzoeker kan niet meer als beperkend worden ervaren.

Ook de opdracht om het voltage op te laten lopen kan als beperkende externe factor worden beschouwd. Hierdoor schuift de proefpersoon steeds één schakelaar op. Als deze opdracht door de onderzoeker wordt weggelaten, is een gigantische afname te zien van het percentage gehoorzame proefpersonen. De proefpersonen ervaren dan niet de externe druk om het voltage op te schroeven. Aan de hand van de percentages van gehoorzame proefpersonen, kunnen sommige externe factoren als *beperkender* worden gezien dan anderen. De factoren zullen nu eerst gecategoriseerd worden in *zeer beperkende factoren*, *beperkende factoren* en *geringe beperkende factoren*. Hierna zal de graduele schaal worden voorgesteld.

Zeer beperkende factoren

- De proefpersoon krijgt geen feedback van de leerling.
- De hoogte van het voltage dient op te lopen.

Beperkende factoren

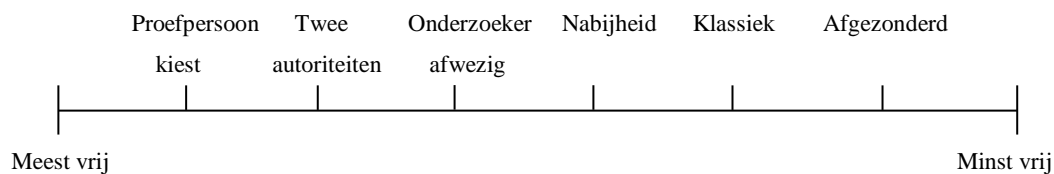
- De leerling bevindt zich in dezelfde ruimte.
- De proefpersoon dient de leerling aan te raken om hem een schok te kunnen geven.
- Er is één onderzoeker aanwezig die consistent is in het geven van het soort aanwijzingen.
- De onderzoeker bevindt zich in dezelfde ruimte als de proefpersoon.

Geringe beperkende factoren

- Het onderzoek vindt plaats op een gerenommeerde universiteit.

- De leerling heeft zich gecommitteerd aan het onderzoek.
- Er wordt gesteld dat het pijnigen van een ander legitiem is.

Deze driedeling zou kunnen leiden tot de graduele schaal zoals hieronder is te zien. Hierop kan worden afgelezen dat de proefpersonen die in het *proefpersoon kiest* experiment tot 450 volt doorgaan een stuk *vrij*er handelen dan de proefpersonen uit het *afgezonderde* experiment. In het geval van het *proefpersoon kiest* experiment is er immers geen sprake van de zeer beperkende factoren dat de proefpersoon geen feedback krijgt en dat hij het voltage op moet laten lopen. In het geval van het *afgezonderde* experiment zijn deze twee zeer beperkende factoren wel degelijk aanwezig, net als de beperkende factoren ‘de onderzoeker bevindt zich in dezelfde ruimte’ en ‘er is één onderzoeker die consistent is in het geven van aanwijzingen’.



Zoals eerder is aangetoond, bepaalt de *mate* van externe beperkingen de *mate* van vrijheid. De proefpersonen uit het *proefpersoon kiest* experiment zijn *vrij*er op het moment dat ze 450 toedienen, omdat er *minder* sprake is van externe beperkingen. Bij het afgezonderde experiment geldt dat de proefpersonen *minder* vrij zijn, omdat er *meer* externe beperkingen van kracht zijn.

Het is tot slot belangrijk de vermelden dat dit slechts *een* benadering is van het experiment van Milgram en de voorgestelde graduele schaal niet als ‘glasharde waarheid’ wordt bedoeld, maar als *een* mogelijke gradatie van vrije wil. Aan de hand van het onderzoek is geprobeerd om het inzichtelijk te maken dat de mate van (externe) beperking bepalend is voor de mate van vrijheid. De waarde die hierboven is gehecht aan verschillende externe factoren staat echter niet vast.

Verantwoordelijkheid van de proefpersonen

Het is binnen de filosofie een gangbare opvatting dat de vraag naar verantwoordelijkheid sterk samenhangt met de vraag naar vrije wil. Mocht iemand uit vrije wil handelen, dan is iemand verantwoordelijk voor zijn of haar gedrag.⁵² Zoals eerder besproken, kan er echter niet gesproken worden over volledig vrije of volledig onvrije handelingen, maar moet gesproken worden over een *mate* van vrijheid. Omdat de notie van verantwoordelijkheid is gekoppeld aan de notie van vrije wil, geldt dat slechts gesproken kan worden van een *mate* van verantwoordelijkheid. De mate van verantwoordelijkheid hangt namelijk samen met de mate van vrije wil.

In het geval van de proefpersonen in het experiment van Milgram kan daarom niet worden gesteld dat ze volledig verantwoordelijk zijn voor het toedienen van 450 volt of dat ze de verantwoordelijkheid volledig af kunnen schuiven op externe factoren. Dezelfde graduele schaal die zojuist is gebruikt om de vrije wil te bepalen, kan worden gebruikt om de verantwoordelijkheid van de proefpersonen vast te stellen. Dat wil zeggen dat een proefpersoon uit het *proefpersoon kiest* experiment meer verantwoordelijk kan worden gehouden voor het toedienen van 450 volt, dan een proefpersoon uit het afgezonderde experiment.

⁵² T. O'Connor, "Free Will" in *The Stanford Encyclopedia of Philosophy*. Geredigeerd door E. Zalta (2016).

Milgram's experiment: II

Gedurende deze scriptie is tot dusver uitgegaan van de neo-compatibilistische visie van Ayer, waarbij de twee door Morris gestelde voorwaarden centraal stonden. We zullen echter zien dat als de neo-compatibilistische visie van Frankfurt op het experiment van Milgram wordt toegepast, er hele andere conclusies met betrekking tot de vrije wil en verantwoordelijkheid van de proefpersonen getrokken kunnen worden. Vervolgens zal Frankfurts notie van verantwoordelijkheid worden toegelicht: volgens Frankfurt is het niet zo dat iemand per definitie niet verantwoordelijk is op het moment dat zijn wil niet vrij is.⁵³ Zelfs als iemand onvrij handelt, dan nog kan deze persoon (in bepaalde gevallen) verantwoordelijk worden gehouden voor zijn of haar gedrag. Tot slot wordt de verantwoordelijkheid van de proefpersonen van het experiment besproken.

Vrije wil van de proefpersonen

Zoals eerder besproken, ligt dat wat de mens onderscheid van andere levende wezens volgens Frankfurt in het feit dat onze wil gestructureerd is.⁵⁴ Mensen *willen* iets met betrekking tot hun wil: ze staan er niet neutraal tegenover. Volgens Frankfurt handelt iemand uit *vrije* wil als diegene handelt naar de verlangens waarvan hij *wil* dat ze ook daadwerkelijk effectief zijn.⁵⁵ Mocht je bijvoorbeeld als verslaafde roker die graag wil stoppen met roken je laatste 'peuken' bij het afval gooien, dan handel je volgens Frankfurt uit vrije wil: je doet immers waarvan je *wil* dat je het zou willen doen. Mocht je toch een sigaret opsteken, dan zou Frankfurt dit als onvrij karakteriseren. Dit is namelijk niet wat je wil dat je zou willen doen.

⁵³ Frankfurt, "Free will and the concept of a person", 93.

⁵⁴ Ibid., 82.

⁵⁵ Slors, *Philosophy of Mind, Brain and Behaviour*, 192.

Nu zal wederom naar het experiment van Milgram worden gekeken om te bepalen of proefpersonen die 450 volt toedienden dit uit vrije wil ondernamen. Deze keer zullen echter niet de twee door Morris gestelde voorwaarden als uitgangspunt worden genomen, maar de positie van Frankfurt. Om de vrije wil van de proefpersonen volgens Frankfurts visie te bepalen, zal naar andere zaken gekeken moeten worden: *wil* de proefpersoon het willen toedienen van de 450 volt, wil hij dit niet of maakt het hem niet zoveel uit.

In eerste instantie lijkt dit misschien moeilijk om te achterhalen. Zoals eerder gezegd zijn niet alle mentale toestanden van een persoon in kaart te brengen. Dit is echter niet noodzakelijk voor het toepassen van de theorie van Frankfurt. Wat slechts achterhaald moet worden, is of de proefpersonen handelden naar de verlangens waarvan ze graag wilden dat deze effectief zouden zijn. Dit is middels de uitvoerige verslaglegging en interviews die achteraf met de proefpersonen zijn gehouden wel degelijk vast te leggen. Twee verschillende proefpersonen, die alle twee 450 volt hebben toegediend, zullen aan bod komen: ‘Meneer Batta’ en ‘Meneer Braverman’.

Meneer Batta

Meneer Batta⁵⁶ reageert in het gehele experiment maar één keer op de leerling met de tekst “You better answer and get it over with. We can’t stay here all night”⁵⁷. Hoewel de leerling naast hem zit en hem smeekt om te stoppen, gaat hij stug – volgens de onderzoeker zijn indrukken zelfs ‘machinaal’ – door met geven van de schokken. De leerling en zijn pijnkretten negeert hij compleet en lijken hem weinig te doen. Richting de onderzoeker is hij echter erg onderdanig en beleefd.

Als Batta achteraf wordt gevraagd of hij zich gespannen of nerveus voelde gedurende het experiment, antwoordt hij dat hij zich slechts nerveus voelde op het moment dat de leerling niet mee wilde werken. Hij corrigeert vervolgens het

⁵⁶ Milgram, *Obedience to Authority*, 45 – 47.

⁵⁷ *Ibid.*, 46.

woord nerveus door “digusted”⁵⁸. Op de vraag waarom hij de schokken heeft toegediend, reageert Batta dat hij nu eenmaal orders op te volgen had. Als hem na enkele maanden een vragenlijst wordt gestuurd over het verloop van het experiment, stelt Batta dat hij er volledig van overtuigd was dat de leerling pijnlijke schokken kreeg, maar dat dit hem totaal niet stoorde.

Uit de literatuur en de specifieke uitingen van Batta valt op te maken dat hij, gedurende het experiment, voornamelijk één verlangen had: het opvolgen van de orders die door de onderzoeker werden gegeven. Van een verlangen om de leerling geen pijn te doen was absoluut geen sprake. Batta gaf zelfs aan vol walging te zijn van het feit dat de leerling op een gegeven moment niet mee wilde werken aan het experiment. Uit andere uitingen valt te concluderen dat hij ook *wilde* dat zijn wil om orders op te volgen tot stand zou komen: Batta excuseert zich zelfs dat hij geen volwaardig proefpersoon heeft kunnen zijn.

Batta heeft dus een eerste-orde verlangen – het op willen volgen van orders – en een tweede-orde volitie die in overeenstemming is met dat eerste-orde verlangen. Volgens de door Frankfurt gestelde definitie van vrije wil, handelt Batta uit vrije wil: hij handelt namelijk naar het verlangen waarvan hij *wil* dat deze ook daadwerkelijk effectief is. Batta *wil* namelijk dat hij orders op wil volgen en handelt hier ook naar.

Meneer Braverman

Meneer Braverman⁵⁹ is een totaal ander geval. Gedurende het experiment lacht hij meerdere malen hysterisch en onbedaarlijk. Braverman spreekt de leerling middels de intercom beleefd toe als hij hem mededeelt dat zijn zwijgen als fout antwoord moet worden beschouwd en er daarom een schok zal volgen. Ook vraagt hij de onderzoeker of het noodzakelijk is om de instructies letterlijk op te volgen en gaat voort met het experiment als de onderzoeker daar bevestigend op antwoordt.

⁵⁸ Ibid.

⁵⁹ Ibid., 52 – 54.

In het interview dat na het experiment wordt afgenomen, stelt Braverman dat hij zich erg nerveus voelde op het moment dat hij zich realiseerde de leerling pijn te doen. Braverman geeft ook aan zichzelf als een vriendelijk persoon te beschouwen en dat hij op een bepaald moment met de leerling wilde praten om de situatie te bespreken. In de vragenlijst die een jaar na deelname wordt opgestuurd, stelt Braverman dat het hem schokte dat hij zo vatbaar bleek voor gehoorzaamheid en inschikkelijkheid ten koste van een voor hem belangrijke waarde: het niet pijn doen van een hulpeloos persoon.

Bij Braverman zijn er duidelijk twee verlangens gelijktijdig aan de orde: enerzijds zijn verlangen om de orders van de onderzoeker op te volgen en anderzijds het verlangen om een hulpeloos persoon als de leerling geen pijn te doen. Hij heeft nadrukkelijk een voorkeur voor het laatste verlangen: op de reactie van zijn vrouw “you can call yourself Eichmann”⁶⁰ reageert hij namelijk met de bewering hij hoopt in de toekomst beter om te kunnen gaan met dit soort conflicten van waarden.

Braverman handelt – in tegenstelling tot Batta – niet in overeenstemming met zijn wil. Hij heeft twee verlangens van de eerste orde: het opvolgen van orders en het niet willen pijnigen van een hulpeloos persoon. Braverman *wil* ook iets ten aanzien van deze twee verlangens, want hij wil dat zijn verlangen om de leerling geen pijn te doen wordt verwezenlijkt. Hij handelt echter tegen deze wil over zijn verlangen in: Braverman veronderstelt de leerling wel degelijk herhaaldelijk pijn te doen door hoge voltages – inclusief de maximum schok van 450 volt – toe te dienen. Ondanks dat zowel Batta als Braverman onder dezelfde omgevingsfactoren en overige omstandigheden 450 volt geven aan de leerling, kan Batta’s handelen als vrij en Bravermans handelen als onvrij worden beschouwd als uit wordt gegaan van Frankfurts definitie van vrije wil.

⁶⁰ Ibid., 54.

Frankfurt over verantwoordelijkheid

Ook op het gebied van verantwoordelijkheid wijkt Frankfurt af van een gangbare opvatting in het vrije wil debat: de in eerdere hoofdstukken gestelde één-op-één relatie tussen vrije wil en verantwoordelijkheid trekt hij namelijk in twijfel. Een persoon kan in bepaalde gevallen volgens Frankfurt alsnog moreel verantwoordelijk worden gehouden, ook al is zijn handeling niet volledig vrij.⁶¹ Iemand wordt normaal gesproken vrijgesteld van morele verantwoordelijkheid als hij wordt gedwongen om iets te doen.⁶² Zo zijn de kindsoldaten die door Kony werden gedwongen om mensen te vermoorden niet verantwoordelijk te houden voor hun daden: als ze niet naar de bevelen van Kony zouden luisteren, moesten ze het immers zelf met de dood bekopen. Frankfurt stelt echter dat dwang niet in alle gevallen als excuus kan worden gezien om onder morele verantwoordelijkheid uit te komen. Om dit aan te tonen, komt Frankfurt met een gedachte-experiment op de proppen waarbij meneer Jones drie maal centraal staat.⁶³

Meneer Jones wordt door iemand bedreigd om een bepaalde handeling X uit te voeren. Stel dat X hier staat voor het chanteren van zijn werkgever. Mocht Jones dit niet doen, dan zijn de gevolgen voor Jones desastreus. Echter, Jones had vóór de bedreiging al plannen om zijn werkgever te chanteren. Nu is de vraag of Jones moreel verantwoordelijk is als hij zijn werkgever ook daadwerkelijk chanteert.

Laten we nu eerst Jones¹ in gedachte nemen. Deze versie van Jones is behoorlijk standvastig: als hij eenmaal een bepaalde beslissing heeft gemaakt, dan zal hij hier hoe dan ook naar handelen. Onder geen beding kan hij op andere gedachten worden gebracht. De bedreiging heeft dan ook geen enkele invloed op hem: hij was immers al van plan om zijn baas te chanteren en zou deze handeling ook uit gaan voeren, ongeacht wat er verder op Jones' pad zou komen. Ondanks

⁶¹ Frankfurt, "Freedom of the will and the concept of a person", 93.

⁶² H. Frankfurt, "Alternate Possibilities and Moral Responsibility" in *The Journal of Philosophy* (Volume 66, nummer 23, 1969), 830.

⁶³ *Ibid.*, 831 – 834.

de dwang van de bedreiging moet Jones wel degelijk moreel verantwoordelijk worden gehouden voor zijn handelen. Het is namelijk niet de dwang die Jones laat handelen zoals hij heeft gehandeld.

Jones² zit als persoon anders in elkaar. Deze versie van Jones wordt volledig overrompeld door de bedreiging en besluit hoe dan ook te handelen naar de gestelde eisen in de bedreiging, omdat hij doodsbenauwd is voor de gestelde consequenties als hij zijn baas niet chanteert. Hij zit zelfs zo vol spanning, dat hij spontaan vergeet dat hij zelf al had besloten zijn baas te chanteren. Volgens Frankfurt is het in dit geval echter niet relevant of Jones vooraf had besloten zijn werkgever wel of niet te chanteren: Jones wordt in zijn handelen volledig gestuurd door de angst voor de bedreiging, waardoor hij niet moreel verantwoordelijk kan worden gehouden.

Tot slot Jones³. Jones is niet volledig lamgeslagen door de bedreiging, maar is er zeker van onder de indruk. Mocht hij niet van plan zijn geweest om zijn baas te chanteren, dan had hij dit vanwege de bedreiging wel degelijk gedaan. Echter, Jones had al in gedachte zijn baas te chanteren. Wanneer hij zijn werkgever vervolgens ook daadwerkelijk chanteert, handelt hij vanuit zijn eigen beslissing en niet uit angst voor de mogelijke gevolgen die zijn gesteld. Deze versie van Jones is ondanks de dwang van de bedreiging wel degelijk verantwoordelijk.

Samengevat staat volgens Frankfurt het begrip 'dwang' centraal als het gaat om morele verantwoordelijkheid. Dwang alleen is echter geen goede graadmeter: als iemand *zelf* al had besloten een bepaalde handeling uit te voeren *ondanks* de dwang, dan kan die persoon alsnog verantwoordelijk worden gehouden voor zijn of haar gedrag. Slechts als door dwang tegen de eigen beslissing in wordt gehandeld, kan morele verantwoordelijkheid af worden geschoven.

Verantwoordelijkheid van de proefpersonen

De vraag is echter in hoeverre bij het experiment van Milgram überhaupt sprake is van dwang. In de sterke zin van het woord worden de proefpersonen namelijk niet *gedwongen* de schokken toe te dienen: er wordt geen pistool op ze gericht en er zijn ook geen verschrikkelijke consequenties verbonden aan het niet toedienen van de schokken. In dit geval zijn zowel meneer Batta als meneer Braverman volledig verantwoordelijk voor hun eigen gedrag: het is uiteindelijk hun eigen beslissing om de stroomstoten te geven.

Dit kan echter worden genuanceerd door te stellen dat er wel degelijk sprake is van een zachte mate van dwang. De onderzoeker stelt immers dat het noodzakelijk is dat wordt doorgedaan met het experiment en dat het onderzoek *vereist* dat de proefpersoon continueert. In zekere zin zet de onderzoeker de proefpersonen aan om hun gedrag voort te zetten en de schokken – tot de maximum van 450 volt – te blijven geven. Dit zou als een vorm van (lichte) dwang kunnen worden beschouwd.

Deze vorm van dwang is echter niet dwingend genoeg om ondermijnd te zijn voor verantwoordelijkheid. Mocht een persoon X tegen me zeggen dat het noodzakelijk is een meisje te laten struikelen, en ik laat vervolgens dat meisje ook daadwerkelijk struikelen, dan kan ik me moeilijk verschuilen achter de dwang van persoon X. Er zijn immers geen consequenties aan verbonden als ik het meisje *niet* laat struikelen: ik blijf gewoon leven en ik hoef geen bakken met geld aan X te betalen. Hoewel ik word aangespoord door X, blijf ik *zelf* verantwoordelijk voor het laten struikelen van het meisje.

Dit geldt ook voor zowel meneer Batta als meneer Braverman. Er wordt ze vooraf gemeld dat ze, ongeacht de uitkomsten van het experiment, de beloofde \$4,00 zullen krijgen. Geld kan dus niet als motief gelden om tot 450 volt te continueren. Ook de zakelijke aansporingen van de onderzoeker zijn niet meer dan dat; aansporingen. Deze kunnen – net als de opmerking van persoon X – niet als dwang worden beschouwd. Vanuit Frankfurtaans standpunt is daarom zowel

Batta als Braverman verantwoordelijk voor het toedienen van 450 volt, ondanks dat Braverman de 450 volt niet uit vrije wil toedient.

7: Vrije wil, verantwoordelijkheid en sociale beïnvloeding

In deze scriptie zijn de twee compatibilistische opvattingen van vrije wil van Ayer en Frankfurt op het experiment van Milgram toegepast, omdat het onderzoek zelf, de omstandigheden en de reacties van de proefpersonen nauwkeurig zijn gemonitord en vervolgens geregistreerd. Hierdoor zijn zaken als ‘externe factoren’ en de tweede-orde volities duidelijker vast te stellen, wat noodzakelijk is om de (mate van) vrije wil te bepalen in het geval van de visie van zowel Ayer als van Frankfurt. Vanuit het experiment kan vervolgens gemakkelijk de parallel worden getrokken naar andere situaties waar sprake is van sociale beïnvloeding.

Zo zijn in de inleiding al gevallen ter sprake gekomen als de massamoorden in My Lai of het gedrag van ‘feuten’ tijdens ontgroeningen. Ook in deze situaties kan worden gesteld dat de soldaten enerzijds en de aankomende eerstejaars anderzijds voldoen aan de eerste voorwaarde die door Morris is gesteld en er een *bewuste* keuze aan hun intentionele gedrag voorafgaat. In het derde hoofdstuk is immers aangetoond dat aan een intentionele handeling *altijd* een bewuste keuze voorafgaat. Dit is daarom ook het geval als een soldaat een trekker overhaalt of als iemand tijdens een ontgroening zes liter water wegdrinkt.

Zoals ook eerder besproken, zijn interne beperkingen lastig te bepalen. Het lijkt misschien aannemelijk dat van de pakweg honderd Amerikaanse soldaten die het bloedbad in My Lai hebben aangericht, er mogelijk één tussen heeft gezeten met ernstige psychische klachten. Dat er van de duizenden deelnemers aan ontgroeningen tenminste een paar een mentale stoornis hebben, lijkt ook waarschijnlijk. Dit neemt echter niet weg dat het gros van de soldaten en studenten in spe niet met ernstige interne beperkingen kampten.

Ook in het geval van My Lai en de ontgroeningen zijn wel degelijk (mogelijke) externe factoren te bedenken. In My Lai deden de Vietnamese trekken van de vrouwen en kinderen misschien denken aan eerdere strijders van de Vietcong die kameraden van de soldaten in kwestie hadden vermoord en ook de bevelen van de commandant zouden als een beperkende factor kunnen gelden. In het geval van de ‘feuten’ zou sprake kunnen zijn van de ‘bevelen’ van een

hogerejaars of kan het feit dat iedereen om je heen hetzelfde gedrag vertoont als een beperkende factor worden gezien.

In beide gevallen geldt echter dat er – als Ayer als uitgangspunt wordt genomen – (hooguit) van een *mate* van beperking sprake kan zijn en dat de vrije wil daardoor nooit *volledig* in het geding kan komen. Dit gaat tevens op voor verantwoordelijkheid: de soldaten en aankomende eerstejaars kunnen in bepaalde mate zijn beperkt in hun verantwoordelijkheid (afhankelijk van de mate van (externe) beperking van de vrije wil), maar ze zullen hun verantwoordelijkheid nooit *volledig* af kunnen schuiven op hun leidinggevende als ze gedrag vertonen waar ze later al dan niet spijt van hebben.

In het geval van Frankfurt is het voornamelijk de vraag of de soldaten *willen* dat ze willen handelen naar de bevelen van de commandant en in het geval van de ‘feuten’ is de vraag of ze *willen* dat dat ze willen handelen naar de vereisten om lid te worden. Dit zal per persoon verschillen en ook de vrije wil van de individuen zal daarom variëren. Echter kan wel worden gesteld dat in Frankfurtaanse zin zowel de soldaten als de deelnemers van de ontgroening verantwoordelijk kunnen worden gehouden voor hun gedrag. Ze worden immers niet in de sterke zin van het woord gedwongen om te handelen zoals ze doen.

Om antwoord te kunnen geven op de onderzoeksvraag *in hoeverre vormt sociale beïnvloeding een gevaar voor het begrip van vrije wil en voor morele verantwoordelijkheid* is het allereerst van belang welk compatibilistisch standpunt als uitgangspunt wordt genomen. De uitkomsten met betrekking tot vrij wil en verantwoordelijkheid zijn namelijk behoorlijk afhankelijk van welke visie wordt gehanteerd. Daarnaast is het van belang welke interne en externe factoren een rol spelen en of er sprake is van dwang bij het uitvoeren van de handeling.

Sociale beïnvloeding kan, zoals we hebben gezien, wel degelijk een rol spelen als het gaat om de inperking van de vrije wil. De externe factoren die door sociale beïnvloeding worden gegenereerd, zoals het geval is bij zowel het experiment van Milgram, My Lai als de ontgroening, kunnen de *mate* van beperking vergroten en de vrije wil in Ayeriaanse zin daardoor belemmeren. Hierdoor vormt sociale beïnvloeding indirect ook een gevaar voor

verantwoordelijkheid, aangezien vrije wil en verantwoordelijkheid nauw met elkaar verbonden zijn.

In Frankfurtaanse zin vormt sociale beïnvloeding een minder groot gevaar. Hoewel sociale beïnvloeding eerste-orde verlangens kan genereren die voor de sociale beïnvloeding nog niet bestonden, zoals ‘ik wil graag orders opvolgen’, heeft sociale beïnvloeding geen invloed op de twee-orde volities die bepalen of een handeling als vrij of onvrij kan worden bestempeld. Ook vormt sociale beïnvloeding geen probleem voor de Frankfurtaanse opvatting van verantwoordelijkheid: dwang, in de sterke zin van het woord, staat bij Frankfurt centraal als het gaat om het af kunnen schuiven van verantwoordelijkheid en daarvan is in het geval van het experiment van Milgram, My Lai en de ontgroeningen geen sprake.

8: Conclusie

Of de proefpersonen in het experiment van Milgram handelden uit vrije wil en verantwoordelijk kunnen worden gehouden voor het toedienen van een – mogelijk dodelijk – percentage van 450 volt, is sterk afhankelijk van welke compatibilistische visie wordt gehanteerd. Als uit wordt gegaan van de neo-compatibilistische visie van Ayer, dan dient een handeling voort te komen uit een bewuste keuze en moeten interne en externe factoren afwezig zijn om de handeling als ‘vrij’ te kunnen bestempelen. Of iemand verantwoordelijk is voor zijn acties, hangt volgens deze vorm van compatibilisme nauw samen met de vraag of de persoon in kwestie uit vrije wil handelde: handelt iemand uit vrije wil, dan is hij tevens verantwoordelijk. Handelt iemand daarentegen niet uit vrije wil, dan kan diegene ook niet verantwoordelijk worden gehouden voor zijn of haar gedrag.

Echter hebben we gezien dat niet zozeer gesproken kan worden van *volledig* vrije of *volledig* onvrije handelingen, maar dat het gaat om de *mate* van vrijheid, die afhankelijk is van de *mate* van interne en externe restricties. Dit geldt tevens voor de proefpersonen die doorgaan tot 450 volt: in welke mate ze uit vrije wil handelen en verantwoordelijk zijn, is afhankelijk van de mate van beperkingen.

Als het neo-compatibilistische standpunt van Frankfurt als uitgangspunt wordt genomen, moeten hele andere conclusies worden getrokken met betrekking tot vrije wil en verantwoordelijkheid. Volgens Frankfurt handelt iemand uit vrije wil, als diegene handelt naar *die* verlangens waarvan hij ook wil dat ze daadwerkelijk effectief zijn. Een proefpersoon die de orders van de onderzoeker wil opvolgen en ook zelf *wil* dat hij handelt naar dat verlangen om die orders op te volgen, handelt volgens Frankfurt uit vrije wil op het moment dat hij de orders ook daadwerkelijk opvolgt. Een andere proefpersoon die *niet* wil dat dit verlangen wordt geactualiseerd, maar liever zou handelen naar de propositionele attitude ‘ik wil geen hulpeloos persoon pijnigen’, handelt daarentegen onvrij als hij toch de aanwijzingen van de onderzoeker opvolgt. Wel

zijn ze volgens Frankfurt beiden verantwoordelijk: verantwoordelijkheid kenmerkt zich volgens Frankfurt in de eerste plaats door ‘dwang’ en hiervan is bij beide proefpersonen geen sprake.

In het voorgaande hoofdstuk is vanuit het experiment van Milgram uitgezoomd naar nog twee verschillende andere situaties waar sprake is van sociale beïnvloeding: de massamoorden in My Lai door Amerikaanse soldaten en de ontgroeningen bij studentenverenigingen. Vervolgens is geconcludeerd dat sociale beïnvloeding voor de Ayerianse argumentatie een groter probleem vormt voor zowel vrije wil als verantwoordelijkheid dan voor de argumentatie van Frankfurt.

De onderzoeksvraag *in hoeverre vormt sociale beïnvloeding een bedreiging voor het begrip van vrije wil en voor morele verantwoordelijkheid* moet dan ook tweeledig worden beantwoord. In het geval van Ayer vormt sociale beïnvloeding wel degelijk een gevaar voor vrije wil en verantwoordelijkheid: door sociale beïnvloeding kan de *mate* van beperking namelijk toenemen. Hierdoor wordt zowel de *mate* van vrije wil als de *mate* van verantwoordelijk beperkt door sociale beïnvloeding. In het geval van Frankfurt vormt sociale beïnvloeding geen bedreiging. Sociale beïnvloeding kan dan misschien meer eerste-orde verlangens creëren, maar het heeft geen invloed op de tweede-orde volitie die in overeenstemming met het eerste-orde verlangen moet zijn om de betreffende handeling als ‘vrij’ te laten gelden.

Tot slot: mocht deze conclusie niet een afronding maar een vertrekpunt zijn, dan zouden meer compatibilistische visies met betrekking tot vrije wil en verantwoordelijkheid verder onderzocht kunnen worden met betrekking tot sociale beïnvloeding. Hierbij kan gedacht worden aan de theorie van Strawson, die stelt dat iemand uit vrije wil handelt als diegene afziet van reactieve attitudes.⁶⁴ Ook de semi-compatibilistische visie van Fischer⁶⁵ zou in overweging kunnen worden genomen. Hij stelt dat vrije wil en determinisme niet compatibel met elkaar zijn, maar dat verantwoordelijkheid en determinisme wel degelijk hand

⁶⁴ Slors, *Philosophy of Mind, Brain and Behaviour*, 189 – 191.

⁶⁵ *Ibid.*, 193.

in hand kunnen gaan. De compatibilistische standpunten zouden vervolgens kunnen worden toegepast op het experiment van Milgram of op andere situaties waar sociale beïnvloeding aan de orde is.

|

Dankwoord

In de eerste plaats gaat mijn dank zeer uit naar dr. Leon de Bruin. Zonder zijn uitvoerige feedback op ingeleverde stukken en snelle reacties op e-mails – met waslijsten aan vragen – had ik de scriptie niet in deze hoedanigheid kunnen voltooien. Ook de mogelijkheid om op korte termijn een afspraak in te plannen op de momenten dat ik vastliep, is erg bevorderlijk geweest voor het afronden van deze scriptie.

Daarnaast ben ik Lin Roosen allereerst erg dankbaar voor het beoordelen en corrigeren van ‘vormgevingspraktijken’ van zaken als tabellen, schema’s en het voorblad. Veel meer nog ben ik haar dankbaar voor de morele ondersteuning op momenten waar het nodig was.

Bibliografie

Ayer, A.J. “Freedom and Necessity” in *Free Will*. Geredigeerd door G. Watson, 15 – 23. Oxford: Oxford University Press, 1982.

Costa C. “Free will and the soft constraints of reason” in *Ratio*. Volume 19, nummer 1 (2006): 1 – 23.

Di Nucci, E. “Priming Effects and Free Will” in *International Journal of Philosophical Studies*. Volume 20, nummer 5 (2012): 725 – 734.

Frankfurt, H. “Alternate Possibilities and Moral Responsibility” in *The Journal of Philosophy*. Volume 66, nummer 23 (1969): 829 – 839.

Frankfurt, H. “Freedom of the will and the concept of a person” in *The Journal of Philosophy*. Volume 68, nummer 1 (1971): 5 – 20.

Goldie, P. *The Emotions: A Philosophical Exploration*. Oxford: Oxford University Press, 2002.

Hobbes, T. *Hobbes’s Leviathan*. Herdruk van het origineel uit 1652. Oxford: Clarendon Press, 1929.

Hoefer, C. “Causal Determinism” in *The Stanford Encyclopedia of Philosophy*. Geredigeerd door E. Zalta, 2016.

Van Inwagen, P. “The Incompatibility of Free Will and Determinism” in *Philosophical Studies*. Volume 27, nummer 3 (1975): 185 – 199.

Kane, R. “Libertarianism” in *Philosophical Studies*. Springer 144 (2009): 35 – 44.

Kane, R. et al. *The Oxford Handbook of Free Will*. Oxford: Oxford University Press, 2005.

McKenna, M. & Coates, J.D. “Compatibilism” in *The Stanford Encyclopedia of Philosophy*. Geredigeerd door E. Zalta, 2006.

Milgram, S. *Obedience to Authority. An Experimental View*. New York: Harper & Row Publishers, 1974.

Morris, S. “Understanding Moral Responsibility within the Context of the Free Will Debate” in *Florida Philosophical Review*. Volume 12, nummer 1 (2012): 68 – 82.

Nahmias, E. “Close calls and the confident agent: free will, deliberation, and alternative possibilities” in *Philosophical Studies*. Springer 131 (2006): 627 – 667.

O’Connor, T. “Free Will” in *The Stanford Encyclopedia of Philosophy*. Geredigeerd door E. Zalta, 2016.

Pereboom, D. “Living Without Free Will: The Case for Hard Incompatibilism” in de online versie van *The Oxford Handbook of Free Will*. Geredigeerd door R. Kane, 1 – 13, 2009.

Slors, M. et al. *Philosophy of Mind, Brain and Behaviour*. Amsterdam: Boom Uitgevers, 2015.

Soechtig, S. *Fed Up*. Atlas Films, 2014.

Vollmer, F. “Intentional action and unconscious reasons” in *Journal for the Theory of Social Behaviour*. Volume 23, nummer 3 (1993): 315 – 326.