



Erasmus Student Journal of Philosophy

#12 | June 2017

Editorial

The twelfth edition of the Erasmus Student Journal of Philosophy started off with a record amount of nominations from both the Faculty of Philosophy and the Erasmus University College. A wide variety of high quality papers were nominated, three of which are featured in this issue. We would like to thank the academic staff for nominating the papers and encourage them to keep doing so in the years to come.

The papers published in this issue, ‘The Case of a Fear of Flying’, ‘Compatibilism and the Basic Argument’ and ‘Een licht dat schijnt op *Licht*’ (‘A light that shines on *Light*’) show both great skill and diversity of interests of the students of our faculty. It is our pleasure to share the results with you and we want to congratulate the authors and thank them for all their work that improved their already impressive papers even further.

As always, the editorial board went through some changes since our last edition. Fabian van Dijk stepped down as editor-in-chief after the previous edition. I would like to thank him for all the hard work he put into the ESJP over the years. Zoé Evrard left the editorial board and after this edition, Daniël Zevenhuizen and Joy Dijkman will leave the board as well. We thank them for their work and wish them all the best.

However sad it is to say goodbye to some of our editors, luckily our call for new editors was met with great response from talented and motivated students from both the Faculty of Philosophy and the Erasmus University College. Though the selection procedure is still going, we look forward to welcoming new talent to our board for the next edition.

It is also with great pleasure that I announce that EIPE student and current member of the editorial board Manon Dillen will take over the position of editor-in-chief from me. I am very happy Manon applied and considering the work she has done so far, I have great confidence

she has what it takes to make the coming editions of ESJP a success. I am grateful for the chance to lead this incredible team and for all it has taught me.

Let me end this editorial by thanking the jury of this year’s Pierre Bayle Trophy, prof. dr. Wiep van Bunge, dr. Sjoerd van Tuinen and dr. Stefan Wintein, for investing their valuable time in reading and evaluating the papers from the previous two editions; the members of the Advisory Board of the ESJP for always being eager to help out whenever needed; and academic staff and PhD-students who helped us by writing reviews. Lastly, I would like to congratulate Dennis Prooi for winning this year’s Pierre Bayle Trophy.

Matthijs Geleijnse
Editor-in-chief

About the Erasmus Student Journal of Philosophy

The Erasmus Student Journal of Philosophy (ESJP) is a double-blind peer-reviewed student journal that publishes the best philosophical papers written by students from the Faculty of Philosophy, Erasmus University Rotterdam and from the Humanities Program of the Erasmus University College. Its aims are to further enrich the philosophical environment in which Rotterdam's philosophy students develop their thinking and bring their best work to the attention of a wider intellectual audience. A new issue of the ESJP appears on our website every July and December.

To ensure the highest possible quality, the ESJP only accepts papers that (a) have been written for a course that is part of the Faculty of Philosophy's curriculum and (b) nominated for publication in the ESJP by the teacher of that course. Each paper that is published in the ESJP is subjected to a double-blind peer review process in which at least one other teacher and two student editors act as referees.

The ESJP encourages students to keep in mind the possibility of publishing their course papers in our journal, and to write papers that appeal to a wider intellectual audience.

More information about the ESJP can be found on our website:

www.eur.nl/fw/esjp

Contact: esjp@fwb.eur.nl

ISSN: 2212-9677

Editorial Board

Matthijs Geleijnse (Editor-in-chief)

Joy Dijksman (Secretary)

Manon Dillen

Dyonne Hoogendoorn

Art van Houwelingen

Merel van de Poel

Jan-Philipp Siebold

Daniël Zevenhuizen

Advisory Board

Huub Brouwer

Thijs Heijmeskamp

Jasper van den Herik

Julien Kloeg

Supervisory Board

dr. B. Leeuwenburgh

prof. dr. F.A. Muller

prof. dr. J.J. Vromen



All work in this issue of the Erasmus Student Journal of Philosophy is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported License. For more information, visit <http://creativecommons.org/licenses/by-nc/3.0/>

Disclaimer

Although the editors of the Erasmus Student Journal of Philosophy have taken the utmost care in reviewing the papers in this issue, we cannot exclude the possibility that they contain inaccuracies or violate the proper use of academic referencing or copyright in general. The responsibility for these matters therefore remains with the authors of these papers and third parties that choose to make use of them entirely. In no event can the editorial board of the Erasmus Student Journal of Philosophy or the Faculty of Philosophy of the Erasmus University Rotterdam be held accountable for the contents of these papers.

In this issue

Is Adriaan Koerbagh an original and independent thinker or is he merely a derivative of Spinoza? In 'Een licht dat schijnt op *Licht*', Stephan Rijdsdijk explores the relation between Koerbagh and Spinoza. If Koerbagh can be acknowledged as an independent thinker of the Enlightenment, he should be reinstated as a brave martyr of the free word. Rijdsdijk argues that Koerbagh is, in fact, not merely an amateur Spinozist and should be acknowledged as an original social, political and religious reformer.

'Compatibilism and the Basic Argument' is Lennart Ackermans's reaction to Daniel Dennett's rebuttal of two versions of the Basic Argument, presented to Dennett by Fischer. Ackermans proposes new ways to address both versions of the Basic Argument, making Dennett's position stronger. Using the concept of the Epistemic World, he shows that Fischer's version of the Basic Argument does not work, putting the compatibilist position back on par with the incompatibilist position.

In 'The Case of a Fear of Flying', Thijs Heijmeskamp discusses the rationality of recalcitrant emotions. Stating that among many theories of emotion there is a certain bias to be found, he shows that conflicts between emotions and judgements are typically seen as irrational, but can in fact be very productive and perfectly rational. Using the case of a fear of flying, Heijmeskamp argues that the (ir)rationality of an emotion is dependent on the evaluative aspect of the emotion and the action the emotion is oriented towards.

Jury report Pierre Bayle Trophy 2017

The task of the Jury is easy to state: compare seven papers that appeared in the ESJP, select the three most deserving ones and come up with a single winner. So then, the Jury had to come up with a collective ranking of the papers based on the individual rankings of those papers by its members. But which aggregation method should be adopted in order to transform the individual rankings of Wiep, Sjoerd and Stefan into the collective ranking of the Jury? We know from Arrow's impossibility theorem that an ideal aggregation method simply does not exist. In Social Choice Theory, a wide variety of aggregation methods are studied that, for the same individual rankings, typically result in different collective rankings. The individual assessments of the papers by the members of the Jury and the aggregation method thus jointly determine the winner of the Bayle Trophy. This Jury adopted a Borda-style method to determine their collective ranking, but we would like to ask the editorial board of the ESJP to communicate their preferred aggregation method to the members of a future Jury.

Although the topics of the seven papers were widely diverging, as were the backgrounds of the members of the Jury, the Jury had no trouble in selecting a top three. All papers were well written and interesting to read, but three papers stood out in terms of originality, precision and style.

The third place is for Edwin Lokker, who highlights the recent rediscovery of Franciscus van den Enden as the author of two important Dutch pamphlets from the early 1660s, arguing compellingly that it makes little sense to regard Van den Enden as a major source of inspiration to Spinoza. During the late 1650s, Van den Enden was Spinoza's Latin teacher, but Lokker demonstrates that the publication of the pamphlets in question occurred too late, and what is more: the similarities between Van den Enden's and Spinoza's political views are 'too little' to assume any profound impact on Spinoza's political philosophy. Lokker's *Too Little, Too Late* is clearly a well-argued paper.

The second place is for Bas Peeters, who deals with Frankfurt's account of personhood. In order to be a person on Frankfurt's account, one needs to have second-order desires about which first-order desires make up one's will.

Frankfurt's account faces the so-called infinite regress problem: very roughly, it has to rely on an infinite hierarchy of desires. Peeters convincingly and carefully argues that Frankfurt's solutions to the infinite regress problem, in terms of the notions of identification and satisfaction, result in faulty accounts of personhood. Along the way, Peeters explains and criticizes how Frankfurt has dealt with the problem of infinite regress at various places in his extensive oeuvre on personhood. Frankfurt on Identification and Satisfaction is a well-written, nicely structured paper with a clear argument.

This year's Pierre Bayle Trophy goes to Dennis Prooi! Prooi provides a knowledgeable account of the rise of Western Philosophy in Japan and its historical roots in the Kyoto School. Inspired by the hermeneutical philosophy of Schleiermacher and Schopenhauer, this school initially set out to study old texts that describe the birth of the Japanese state. But very quickly its aim became to establish an independent Japanese identity. At the same time that Japan modernizes - or rather: militarizes - in order to compete with Western powers, there emerges a new philosophy, in which Zen Buddhism is reconciled with modern science. An important question then is whether the interest in this new philosophy can be separated from the rise of ethnocentrism, nationalism, and ultimately fascism? Dennis Prooi convincingly argues that we cannot and does not only address Kyoto School philosophers such as Nishida and his pupils Suzuki, Kuki and Tetsuro, but also the 'philologism' of the hermeneutical method itself. At the same time he nuances this critique: for while the thought of the Kyoto School was popular under the military regime, Nishida himself very much resisted the militarism of his time. Perhaps the hermeneutical method is easy to instrumentalize, but it is also the philosophy par excellence that is capable of transcending cultural differences. This makes the contribution of Prooi a sober and balanced, but for that matter not less original and spirited contribution to intercultural philosophy.

L. van Bunge, S. van Tuinen and S. Wintein
Jury of the Pierre Bayle Trophy 2017



Table of contents

Een licht dat schijnt op *Licht*

Over de intellectuele zelfstandigheid van Koerbagh ten opzichte van Spinoza

Stephan Rijsdijk

7-25

6

Compatibilism and the Basic Argument

Lennart Ackermans

26-33

The Case of a Fear of Flying

An enquiry into the rationality of recalcitrant emotions

Thijs Heijmeskamp

34-43

Een licht dat schijnt op *Licht*

Over de intellectuele zelfstandigheid van Koerbagh ten opzichte van Spinoza

Stephan Rijsdijk

1. Inleiding

In de zeventiende eeuw waren door religie gevoede conflicten aan de orde van de dag. De Republiek der Zeven Verenigde Nederlanden (hierna te noemen: de Republiek) kende een grote mate van economische voorspoed vanwege de toegenomen handel, desondanks was zij intern conflictueus. Men kon het in de jonge Republiek niet eens worden over niet van elkaar los te denken religieuze en politieke opvattingen (Leeuwenburg, 2013, p. 12). Te denken valt aan de religieus en politiek geladen debatten tussen de remonstranten en de contraremonstranten, uiteindelijk uitmondend in de onthoofding van Johan van Oldenbarnevelt. De Republiek bevond zich tevens in de periode van de verlichting, waarin de filosofie gerevolutioniseerd werd. Dat gold in het bijzonder voor de natuurfilosofie, die door de opkomst van de wetenschappelijke methode gedurende de wetenschappelijke revolutie sterk zou worden uitgehold.

In deze turbulente periode leefden de radicale denkers en gebroeders Adriaan (1633 – 1669) en Johannes (1634 – 1672) Koerbagh.¹ Adriaan probeerde met zijn werk een eind te maken aan de godsdienstige conflicten van zijn tijd. Ondanks deze pacifistische motivatie, werd hij door zijn tijdgenoten verketterd. Zijn filosofie werd namelijk zodanig radicaal gevonden dat nota bene in de relatief tolerante Republiek, op 27 juli 1668 een gruwelijke lijfstraf en een gevangenisstraf tegen hem werden geëist: de sancties die stonden op de door hem geschreven “godslasterlijke” werken *Bloemhof* en *Licht*.² Johannes Koerbagh heeft zelf geen boeken gepubliceerd. Wel wordt vermoed dat hij medeauteur was van de werken van zijn broer Adriaan.

Van de gebroeders Koerbagh was bekend dat zij deel uitmaakten van de collegianten; een groep ondergrondse vrijdenkers die op geheime locaties vrijelijk met elkaar delibereerden over theologie en filosofie. Voor hen waren traditionele kerkelijke opvattingen over religie, zoals de triniteit, niet heilig en stond de individuele beleving van het geloof centraal (Ibid., p. 146). Onder de collegianten bevond zich ook Spinoza. Het is aannemelijk dat de gebroeders Koerbagh goed met hem waren bevriend (Ibid., p. 115). Vanwege hun beroemde tijdgenoot zijn de gebroeders Koerbagh altijd in de schaduw van Spinoza blijven staan. Zij worden beschouwd als zijn volgelingen of zelfs als amateurspinozisten. Maar is dit wel terecht?

In dit essay onderzoek ik of Adriaan Koerbagh op basis van zijn werk *Licht* beschouwd moet worden als origineel en zelfstandig denker ten opzichte van Spinoza. Is Koerbagh een afgeleide van Spinoza en is hij zelfs met recht een amateurspinozist te noemen? Zouden we Koerbagh intellectueel moeten emanciperen van Spinoza? Als Koerbagh erkend kan worden als zelfstandig verlichtingsdenker met originele ideeën, dan kan hij vanuit de vergetelheid in ere hersteld worden als moedige martelaar van het vrije woord. Daarnaast worden in dit onderzoek de leemtes in het nog relatief onontgonnen gebied van de filosofische verhouding tussen Spinoza en Koerbagh meer opgehelderd en worden de aannames omtrent die verhouding zichtbaar gemaakt. En zelfs als deze pogingen tevergeefs blijken te zijn, dan kan men in tijden van religieuze intolerantie van *Licht* leren.

1.1 Methodologische verantwoording en opbouw

In dit essay zal de onderzoeksvraag worden beantwoord op basis van een filosofische vergelijking tussen de *Ethica* en *Licht*, die in een breder historisch en filosofisch perspectief zal worden geplaatst. Een essentiële vraag die in dit kader beantwoord moet worden, is hoe men kan bepalen of de ene filosoof in intellectueel opzicht van een andere denker geëmancipieerd moet worden. In mijn optiek is een brede oriëntatie noodzakelijk voor de beantwoording van deze vraag. Denkers kunnen immers contrasterende ideeën hebben, maar als een leerling tegen zijn leermeester ageert, dan is dit een aanwijzing voor verwantschap en mogelijk afhankelijkheid. Belangrijk is ook wat de verschillen in motivaties zijn van twee filosofen en of zij ideeën huldigen die origineel zijn. Als een filosoof geïnspireerd wordt door een ander – en inspiratie impliceert niet onzelfstandigheid – maar zijn ideeën voor een ander doel inzet, dan is dit weer een aanwijzing voor zelfstandigheid. In een vergelijking tussen twee filosofen is een idee origineel als deze afkomstig is van de denker zelf of terug te voeren is tot een andere bron dan de denker waarmee wordt vergeleken. Deze niet-limitatieve opsomming als toetsingskader voor de emancipatie van denkers ten opzichte van elkaar rechtvaardigt de door mij gekozen methodologie van de systematische vergelijking, geplaatst in een bredere filosofische en historische context.

Gelet op de omvang van dit essay is ervoor gekozen slechts een werk per filosoof in acht te nemen. De keuze voor het werk *Licht* van Koerbagh ligt op het eerste gezicht meer voor de hand dan de keuze voor de *Ethica* van Spinoza. Dit omdat de *Ethica* meer metafysisch, maar uiteindelijk ethisch, georiënteerd is, terwijl op basis van een ander beroemd werk van Spinoza, de *Tractatus Theologico-Politicus*, mogelijk ook een belangrijk filosofisch contrast met Koerbagh ontstaat als het gaat om de onderwerpen theologie, godsdienst en politiek.

Concepties van Koerbagh als amateurspinozist, een slechte imitator van Spinoza of zijn leerling, zijn ingegeven door de gelijkenissen in de metafysica van de beide denkers. En het is de *Ethica* waarin de metafysica van Spinoza het meest uitgebreid wordt behandeld. In hoofdstuk 4 van dit essay zal echter blijken dat deze metafysische gelijkenissen genuanceerd dienen te worden. Bovendien is de *Ethica* het meest volledige werk van

Spinoza, omdat hij zowel zijn metafysica als epistemologische en ethische onderwerpen behandelt. Dit werk is ook zijn *magnum opus*.³

Licht is het meest systematische werk van Koerbagh. Omdat zijn andere werken op een lexicografische wijze zijn opgebouwd, zijn ideeën daar meer versplinterd weergegeven. Bovendien kan *Licht* beschouwd worden als een voortzetting van de andere werken van Koerbagh.

Ter beantwoording van de onderzoeksvraag zal dit essay als volgt zijn opgebouwd. In de hierop volgende hoofdstukken 2 en 3 zullen respectievelijk de wijsgerige werken *Ethica* en *Licht* worden besproken. Daarbij zullen ook het leven en de historische context van de denkers aangestipt worden. In hoofdstuk 4 worden de beide werken en filosofen met elkaar vergeleken. Deze vergelijking wordt vervolgens in ditzelfde hoofdstuk in een breder filosofisch en historisch perspectief geplaatst. Tot slot zal, op basis van de gemaakte vergelijking, de bronnen van de hier centraal staande werken en de originaliteit van de filosofische ideeën in *Licht*, de conclusie worden getrokken dat men Koerbagh intellectueel zou moeten verzelfstandigen van Spinoza.

2. Vrijheid en geluk in een strikt deterministische wereld

2.1 Het leven van Benedictus de Spinoza

Spinoza werd in 1632 als Bento de Spinoza geboren binnen de Portugees-Joodse gemeenschap van Amsterdam. Zijn ouders waren vanwege de inquisitie geëmigreerd vanuit Portugal naar de Republiek. Zijn familie was bij lange na niet zo vermogend als de grote kooplieden van de Republiek, maar desalniettemin was zij welgesteld (Nadler, 2006, p. 2). Al op jonge leeftijd bleek dat Spinoza intellectueel begaafd was. Op zeventienjarige leeftijd werd hij echter geacht zijn studie stop te zetten om de plaats van zijn overleden broer in het familiebedrijf over te nemen. Het leven als koopman was echter niet aan hem besteed: Spinoza was veel meer een denker dan een handelaar. Hij besloot Latijn en klassieke studies te bestuderen in de Latijnse school van Franciscus van den Enden (1602 – 1674). Ook nam hij een Latijnse naam aan: Benedictus de Spinoza. In 1656 werd hij geëxcommuniceerd uit de joodse gemeenschap. Tot op de dag van van-

daag is het onduidelijk wat de concrete aanleiding hiervoor was, al zal niemand zich over de excommunicatie verbazen, gezien de radicaliteit van Spinoza's denken (Ibid., p. 6).

Zoals al beschreven in de inleiding, behoorde Spinoza tot de collegianten, waarmee hij tevens was bevriend. Tot deze vriendenkring behoorden mensen als Lodewijk Meijer (1629 – 1681), Jarig Jellesz (1620 – 1683), Johan Bouwmeester (1630 – 1680) en vermoedelijk ook de gebroeders Koerbagh. Zijn *magnum opus*, de *Ethica*, werd pas postuum gepubliceerd door zijn vrienden in 1677. Vanaf 1663 leefde hij in Voorburg, waar hij verder zou werken aan zijn *Ethica*. Rond 1670 besloot hij naar Den Haag te verhuizen. In datzelfde jaar publiceerde hij anoniem zijn *Tractatus Theologico-Politicus*. Hij stierf ten slotte in 1677, vermoedelijk ten gevolge van tuberculose.

2.2 De methodologie van de *Ethica*

In 1662 begon Spinoza al aan zijn *Ethica* te werken. Wie voor de eerste maal een blik werpt op zijn levenswerk, zal zich meteen verwonderen over de opmerkelijke stijl waarin het is geschreven. In de *Ethica*, die volledig in het Latijn is geschreven, is het denken van Spinoza volgens meetkundige ordening uiteengezet, zoals blijkt uit de ondertitel *Ordine Geometrico Demonstrata* (Spinoza, 1677, pp. 54-55).⁴ De *Ethica* bestaat uit vijf delen, die direct de globale inhoud van die hoofdstukken weerspiegelen. De titels van deze delen luiden als volgt:

- I) God;
- II) De natuur en de oorsprong van de geest;
- III) De oorsprong en de natuur van de hartstochten;
- IV) De menselijke slavernij, met andere woorden de krachten van de hartstochten;
- V) De macht van het verstand, met andere woorden de menselijke vrijheid.

Ieder hoofdstuk is overeenkomstig de genoemde geometrische methode opgebouwd, wat *in concreto* betekent dat Spinoza zijn denken onderbouwt met verzamelingen van gelieerde definities, axioma's, stellingen, bewijsvoeringen en commentaren die in het kader van de bewijsvoering niet

strikt noodzakelijk zijn, maar wel verhelderen wat Spinoza tracht te betogen. Alle stellingen van de *Ethica* worden onderbouwd door de definities, axioma's en stellingen die eraan vooraf zijn gegaan.

Spinoza had bewust voor de geometrische methode gekozen vanwege de grote successen die wetenschappers als Galilei en Huygens gedurende de wetenschappelijke revolutie hadden geboekt (Nadler, 2006, p. 37). Spinoza breidde het toepassingsbereik uit van deze methodologie, die reeds in het jaar 300 voor Christus door de Hellenistische wiskundige Euclides was toegepast in zijn *Elementen*, met als doel de constructie van een absoluut en zeker filosofisch systeem. Bovendien vormt de structuur van de *Ethica* een weerspiegeling van de werkingwijze van de door Spinoza onderzochte werkelijkheid: inhoud en vorm zijn in de *Ethica* sterk met elkaar verweven (Ibid., p. 41). Spinoza meent namelijk dat de fysieke en de mentale werkelijkheid parallel geordend zijn: een mathematisch werkende natuur vergt een overeenkomstige rangschikking van ideeën (Ibid., p. 42). De door Spinoza gehanteerde methode heeft ook een sterke argumentatieve functie: wie immers de definities en axioma's van de *Ethica* aanvaardt, kan niet anders dan zich aansluiten bij de in totaal 259 stellingen die in het werk zijn opgenomen (Ibid., p. 40).

Daarnaast geldt dat men de meetkundige methode dient te beschouwen als een wijze waarop Spinoza zijn filosofische ontdekkingen wenste te presenteren en niet noodzakelijk als methode voor de discursieve vergaring van dergelijke kennis (Ibid., p. 39). Het is dus belangrijk te beseffen dat Spinoza de in de *Ethica* gepresenteerde stellingen in zijn denken niet *a priori* heeft afgeleid van definities en axioma's.

2.3 Over de ethische doelstelling van de *Ethica*

Om de filosofie van de *Ethica* te kunnen vatten, is het belangrijk de titel en de daarmee samenhangende ultieme doelstelling van dit werk toe te lichten. Spinoza staat met name bekend als ontkenner van de vrije wil, determinist, substantiemonist en bijbelcriticus: met name dus als metafysisch systeembouwer. Waarom luidt de naam van zijn *magnum opus* dan *Ethica*? Hoewel Spinoza inderdaad een groot metafysicus was, staat dit deel van zijn filosofie in dienst van zijn epistemologie, die op haar beurt weer de grondslag vormt van zijn ethiek. De metafysica van Spinoza resulteert in

een wereldbeeld waarin wij als mens een bepaalde plaats hebben. Het roept vragen op van ethische aard, die hij in de *Ethica* wenst te beantwoorden. De ultieme vraag die Spinoza van beantwoording tracht te voorzien, luidt: hoe kan de mens gelukkig worden in een strikt deterministische wereld waarin geen vrije wil in traditionele zin van het woord bestaat?

2.4 De vereenzelviging van God, substantie en de natuur

Het eerste deel van de *Ethica* is gewijd aan dit deterministische wereldbeeld. Het spinozistische godsbeeld was radicaal en vormde een van de belangrijkste argumenten voor tijdgenoten om Spinoza te betichten van atheïsme en ketterij. Zijn godsbeeld week namelijk in grote mate af van het traditionele – of meer spinozistisch verwoord: antropomorfe – beeld van God. In traditionele zin wordt God namelijk voorgesteld als transcendente entiteit, die *ex nihilo* ons heelal heeft geschapen vanuit een contingente, vrije handeling (Nadler, 2002, p. 228). Het is een God die op antropomorfe wijze beslist over goed en kwaad en die teleologisch gemotiveerd is. Spinoza zette zich af tegen deze traditionele conceptie van God. Zo stelt hij: “Sommige mensen beelden zich in dat God, net als de mens, uit een lichaam en een geest bestaat en aan aandoeningen onderworpen is. Hoever zij echter van de ware kennis van God verwijderd zijn, blijkt afdoende uit hetgeen reeds is bewezen” (Spinoza, 1677, E1P15s).

In het kader van deze missie tracht hij een bewijs te leveren voor een monistische substantieleer, waarin de substantie wordt vereenzelvigd met zowel God als de natuur. “*Deus, sive natura*”, een van de meest bekende aforismen die de westerse wijsbegeerte rijk is, vormt een beknopte weergave van het denken dat zich in dit deel van de *Ethica* voltrekt. Echter, zoals gewoonlijk bij aforismen, vraagt deze voor Spinoza karakteristieke uitspraak om nadere toelichting. Hoe kan Spinoza zowel de substantie als God en de natuur met elkaar identificeren?

De *Ethica* vangt in dat kader aan met een metafysische uiteenzetting, waarin het noodzakelijke bestaan wordt aangetoond van één ondeelbare, niet-teleologische en onveroorzaakte substantie, die God wordt genoemd. De deantropomorfisering van God en Zijn vereenzelviging met de ene noodzakelijke substantie vindt plaats in de eerste vijftien stellingen van het eerste deel van de *Ethica*. Om de metafysica van Spinoza te kunnen

begrijpen, dient allereerst een aantal cruciale begrippen de revue te passeren. Met behulp van deze begrippen kan de ontologie van Spinoza al op een elementair niveau worden begrepen.

Te beginnen bij de substantie zelf, die Spinoza definieert als “iets wat in zichzelf is en door zichzelf wordt begrepen, dat wil zeggen een zaak waarvan het begrip niet het begrip van iets anders nodig heeft op grond waarvan het gevormd moet worden (Ibid., E1D3).” Dit komt overeen met de cartesiaanse opvatting van een substantie als ontologisch en epistemologisch onafhankelijk: iets dat voor wat betreft zijn bestaan en begrip onafhankelijk is van andere zijnden. Een belangrijk verschil met de cartesiaanse substantieleer vormt echter de invulling van het begrip substantie: waar Descartes een primaire substantie, namelijk God, en twee secundaire substanties kent, te weten denken (*res cogitans*) en uitgebreidheid (*res extensa*), daar neemt Spinoza slechts het bestaan van één substantie aan.

Het begrip “attribuut” is tevens ontleend aan het cartesiaans begrip penapparaat. Het wordt door Spinoza gedefinieerd als “een zaak die het verstand kent als iets wat het wezen van een substantie vormt” (Ibid., E1D4). Attributen, dat wat de substantie constitueert, vormen dus de wezenlijke aard der dingen. Denken en uitgebreidheid, bij Descartes nog substanties, vormen bij Spinoza de enige door de mens kenbare attributen binnen de verzameling van een oneindig aantal attributen.⁵

Een concrete particuliere uitdrukking van de substantie en de attributen, bijvoorbeeld een steen of een idee, wordt door Spinoza aangeduid met het begrip “modus”, gedefinieerd als “de aandoeningen van een substantie, met andere woorden een zaak die bestaat in iets anders, waardoor men haar ook begrijpt” (Ibid., E1D5). Omdat de substantie vooraf gaat aan de modi die door haar worden uitgedrukt, zijn de modi voor wat betreft ons begrip ervan afhankelijk van de substantie.

Hoe hangen deze begrippen nu onderling met elkaar samen? Spinoza stelt dat oneindig veel dingen op oneindig veel wijzen op een noodzakelijke manier uit God voortvloeien (Ibid., E1P16). Zo vormt het bestaan van bijvoorbeeld een steen een particuliere modus die onder het attribuut extensie een logisch noodzakelijke uitdrukking vormt van de substantie.

Omdat alle dingen een uitdrukking zijn van één en dezelfde substantie, zijn de attributen de verschillende bestaanswijzen van God. Lichaam en geest zijn bestaanswijzen van de mens, respectievelijk onder de attributen extensie en denken. Omdat de attributen in het algemeen, en dus ook extensie en denken in het bijzonder, ontologisch niet onafhankelijk ten opzichte van elkaar zijn, doemt het cartesiaanse probleem van de causale interactie tussen lichaam en geest bij Spinoza niet op.

In het kader van het voormelde aspect van noodzakelijkheid wijs ik op de derde en vierde axioma's. Het derde axioma luidt: "uit een gegeven, bepaalde oorzaak vloeit noodzakelijk een gevolg voort en omgekeerd: zonder een bepaalde oorzaak kan het gevolg onmogelijk optreden" (Ibid., E1A3). Het drukt het sterk deterministische karakter van de spinozistische filosofie uit. Naast het feit dat alle zaken in de wereld veroorzaakt zijn, hetzij door iets externs, hetzij door zichzelf, geldt ook dat gevolgen noodzakelijk voortvloeien uit hun oorzaken. Opvallend is dat Spinoza in dit deterministische wereldbeeld toch het bestaan van goddelijke vrijheid erkent (Ibid., E1P17). Deze vorm van vrijheid is er een van spontaniteit als tegendeel van alternativiteit: vrijheid als afwezigheid van externe causale beïnvloeding, omdat de activiteit van God uitsluitend voortkomt uit Zijn interne principes. En het vierde axioma luidt: "de kennis van een gevolg hangt af van de kennis van de oorzaak en sluit die in" (Ibid., E1A4). Dit impliceert dat kennis van de werkelijkheid opgedaan kan worden door de oorzaken der dingen te bestuderen.

Spinoza stelt niet alleen dat God de ultieme efficiënte oorzaak is van alles wat bestaat, wat de gedachte behelst dat alle zijnden zich binnen de reikwijdte van de causaliteit bevinden, waardoor niets eraan ontkomt. Hij pleit ook voor de ontkenning van enige vorm van contingentie. De gehele werkelijkheid is dus ook noodzakelijk vanuit een logisch perspectief. De werkelijkheid is zoals zij is en zij had niet anders kunnen zijn dan zij is. Ook bij Spinoza's invulling van causaliteit blijkt duidelijk dat hij is geïnspireerd door de wiskunde. Zo hanteert hij de metafoer dat de wijze waarop de modi uit de attributen voortvloeien langs dezelfde logisch noodzakelijke weg verloopt als eeuwige meetkundige waarheden uit geometrische figuren, zoals uit de aard van een driehoek voortvloeit dat de som van zijn drie hoeken gelijk is aan die van twee rechte hoeken (Ibid., E1P17s).

Denken in termen van contingentie of logische mogelijkheden komt voort uit een onwetendheid over de ware oorzaken der dingen.

Het aldus geschetste wereldbeeld heeft belangrijke filosofische en theologische consequenties. Een aspect van Spinoza's hyperdeterminisme behelst de ontkenning van het bestaan van mogelijke werelden. In tegenstelling tot Spinoza meende Leibniz dat God op basis van een willekeurige, keuzevrije handeling tot schepping van de beste van alle mogelijke werelden was overgegaan. Uit de *Ethica* vloeit echter voort dat onze wereld de enige mogelijke wereld is (Nadler, 2006, p. 87). Spinoza deantropomorfiseert God in nog grotere mate door elke vorm van aristotelische finale oorzakelijkheid in de werkelijkheid uit te sluiten (Spinoza, 1677, E1P36A). De natuur wordt geantropomorfiseerd als de mens de fout maakt zijn doelmatige denken op haar te projecteren. De gedachte dat de natuur voor de mens zou zijn ingericht, berust op een menselijke onwetendheid aangaande de natuur en haar ware oorzaken. Bovendien is de actuele wereld absoluut volmaakt zoals zij op dit moment is. Het is een antropomorfe fout van de mens om de volmaaktheid van de natuur te toetsen op grond van zijn eigen ideeën. We dienen de dingen dus te beoordelen zoals ze op zichzelf zijn. En Spinoza stelt: "de dingen zijn dus niet meer of minder volmaakt omdat ze de menselijke zintuigen al of niet aangenaam zijn of de menselijke natuur al of niet van nut" (Ibid.). Dit geldt bijvoorbeeld voor misvormingen, overstromingen, ziektes en de zonde (Ibid.). Het zal nu ook niemand meer verbazen dat ook voor wonderen in traditionele zin, dat wil zeggen een vrije ingreep van God in de werkelijkheid, de wetmatigheden van de natuur schendend, geen plaats is in dit denken.

Ook de tot het attribuut van het denken behorende menselijke wil bevindt zich in deze causale en logische keten. Spinoza stelt: "in de geest is er geen absolute of vrije wil, maar de geest wordt tot het willen van iets aangezet door een oorzaak die op zijn beurt ook door een andere is aangezet, en deze weer door een andere, en zo tot in het oneindige" (Ibid., E2P48). Toch meent Spinoza vanuit zijn epistemologie dat de mens een zekere mate van vrijheid kan verkrijgen. Het is hiervoor volgens hem noodzakelijk dat de mens participeert in het oneindige en absoluut ware intellect van God. Maar hoe is dit mogelijk voor de eindige en onvolmaakte mens?

2.5 Kennis: vertrouwen in de menselijke rede

De centrale vraag die Spinoza in zijn epistemologie tracht te beantwoorden luidt dan ook: hoe kan de mens de ware ordening van de oorzaken en dingen in het intellect van God kennen?

Het begrip “idee”, gedefinieerd als “een voorstelling van de geest, die hij vormt omdat hij een zaak is die denkt”, heeft hierin een centrale plaats (Ibid., E2D3). Spinoza onderscheidt adequate van inadequate ideeën. Onder een adequaat idee wordt verstaan “een idee, die [...] alle wezenlijke eigenschappen of intrinsieke benoemingen van een waar idee heeft” (Ibid., E2D4). Alle ware ideeën zijn tevens adequate ideeën en alle onware ideeën zijn inadequate ideeën. De oneindige, complete verzameling van alle adequate ideeën bevindt zich in het oneindige intellect van God. Wanneer onze ideeën overeenstemmen met de ideeën binnen dit oneindige intellect, dan kunnen wij derhalve ook spreken van adequate ideeën.

Hoe is het nu mogelijk dat we adequate in plaats van inadequate ideeën van de werkelijkheid verkrijgen en onze geest dus kunnen vereenzelvigen met het oneindige intellect van God? De *Ethica* is in dit kader in ieder geval buitengewoon pessimistisch ten aanzien van ideeën die zich in de laagste hiërarchische positie bevinden, namelijk de verbeelding. Als op grond van deze vorm van kennis ideeën worden gevormd over externe lichamen, beschouwen we ze uitsluitend vanuit het perspectief van ons eigen menselijk kenapparaat, waardoor we de dingen noodzakelijk vanuit een temporeel en ruimtelijk perspectief bezien. We nemen de dingen dan waar zoals ze ons voorkomen en niet zoals ze op zichzelf bezien zijn.

Spinoza meent dat het toch mogelijk is voor de mens om kennis te verkrijgen via de tweede en derde wijzen van kennen, namelijk dat wat hij respectievelijk discursief inzicht en intuïtief weten noemt (Ibid., E2P40S2). Kennis die uit deze kenwijzen voortkomt, is noodzakelijk waar, in tegenstelling tot de voornoemde eerste wijze van de verbeelding. Onder het discursieve inzicht verstaat Spinoza de “aspecten die alle lichamen gemeen hebben en [die] op een gelijke wijze zowel in een deel als in het geheel zijn” (Ibid., E2P38). Het gaat hier niet om universalialia, zoals “mens, paard of hond”, maar om de eigenschappen die alle lichamen gemeenschappelijk hebben omdat ze modi onder het attribuut extensie zijn, dat wil zeggen:

hun wezen (Ibid., E2P40S1). Essentieel voor het begrip van deze categorie vormt het inzicht dat deze gemeenschappelijke aard niet extern begrepen kan worden, dat wil zeggen gemedieerd door andere lichamen of het eigen lichaam, maar intern, oftewel vanuit een logische reeks in het denken (Ibid., E2P29S).

Het ontbreekt in de *Ethica* aan een duidelijk analyse van intuïtieve kennis. Op grond van een voorbeeld dat Spinoza ter verduidelijking in de *Ethica* aanlevert, kunnen we ons er toch wel een idee van vormen. Spinoza beschrijft namelijk een getallenreeks bestaande uit drie getallen, gekenmerkt door een zekere logische verhouding, waarvan wij in één oogopslag, direct dus, het vierde afleiden. Hij lijkt hiermee te doelen op een directe toegang tot de absolute waarheid in het geval van logische of wiskundige reeksen, zoals de natuur volgens Spinoza ook geordend is. Het lijkt er dus op dat de mens intuïtief inzicht kan verkrijgen vanuit een volledig deductieve en transitieve kenwijze die in het menselijke kenapparaat geïnternaliseerd is. Op grond van discursief inzicht en intuïtief weten kunnen wij onze adequate ideeën uitbreiden en meer kennis van God of de natuur verkrijgen, omdat alle ideeën die zijn afgeleid van adequate ideeën, tevens zelf adequate ideeën zijn. De hiërarchische superioriteit van intuïtieve kennis ten opzichte van discursief inzicht is vervolgens niet gelegen in de vormen van kennis waartoe zij leiden: in beide gevallen zijn dit immers noodzakelijk adequate ideeën. Het hiërarchische verschil ligt eerder in de wijze waarop we die kennis kunnen verkrijgen en in de ethische consequenties van de kenwijzen. Waarom is het voor Spinoza zo belangrijk een adequate kennis van de natuur te verkrijgen in het licht van zijn ethische missie? Oftewel, hoe leidt deze kennis nu tot gelukzaligheid voor de mens?

2.6 De weg naar gelukzaligheid en eeuwigheid

Het antwoord op deze vraag is gelegen in de deugdzaamheid als ideaaltype voor de mens. Dit doet direct denken aan de antieke Griekse wijsbegeerte, waarin de deugd begrepen werd als het hebben van karaktereigenschappen die maken dat je als mens excelleert (Parry, 2014). Dit is echter niet waar Spinoza met dit begrip op doelt. In de *Ethica* is de deugd immers niets anders dan de *conatus*, oftewel het streven naar zelfbehoud, wanneer deze uitsluitend is voortgekomen uit de mens zelf. De deugdzaam

mens handelt zoveel mogelijk op grond van zijn innerlijke principes en wordt zo weinig mogelijk beïnvloed door externe zaken: hij leeft onder de leiding van de rede, en niet op grond van de hartstochten waardoor hij wordt aangedaan. Tegelijkertijd maakt de mens – net als alle andere eindige modi – onderdeel uit van de natuur en kan hij nooit volledig actief zijn. Als de mens namelijk volledig onafhankelijk van externe invloeden zou kunnen handelen, dan zou hij nooit vergaan (Spinoza, 1677, E4P4D). We kunnen het ideaaltype van de deugdzame mens dus benaderen, maar er nooit volledig mee samenvallen.

Deugdzaam handelen omvat het verkrijgen van adequate ideeën, of vrij vertaald: wetenschappelijke, wetmatige, universele, eeuwige kennis van de werkelijkheid. De deugd draait bij Spinoza in feite om het op discursieve of intuïtieve wijze verkrijgen van kennis van God of de natuur: “Het hoogste goed van de geest is de kennis van God, en de hoogste deugd is het kennen van God” (Ibid., E4P28). De deugdzame mens is dus actief en leeft onder de leiding van de rede, waardoor hij zijn streven naar gelukzaligheid niet af laat hangen van willekeurige, aan hem externe zaken, maar van de eeuwige wetenschappelijke kennis van de werkelijkheid, die hij vanuit zichzelf discursief dan wel intuïtief verkrijgt. De hartstochten waardoor de deugdzame mens wordt aangedaan, worden door hem gematigd. Hij streeft datgene na wat daadwerkelijk goed voor hem is, in tegenstelling tot de slaafse mens, die slechts op een oppervlakkige, hedonistische wijze streeft naar datgene wat hem goed voorkomt.

De deugdzame mens is bovendien vrij. Dit is geen vrijheid in alternatieve zin, oftewel een vorm van vrijheid waarin de menselijke wil onveroorzaakt is, zoals bij Descartes, die de *res cogitans* had onttrokken aan het bereik van het mechanisme. Bij Spinoza is de deugdzame mens vrij, omdat hij vrij is in spontane zin. Dit houdt in dat ook zijn wil binnen de keten der causaliteit valt, maar hij zelfgedetermineerd is, dat wil zeggen dat de vrije mens autonoom is en enkel hij zelf de adequate oorzaak is van zijn activiteit of handelen.

Het vrije leven onder de leiding van de rede impliceert geen teruggetrokken bestaan, bijvoorbeeld als kluizenaar, om zoveel mogelijk externe impulsen te vermijden. Integendeel, Spinoza meent hierover dat het kenmerkend is “voor de wijze dat hij de dingen gebruikt en daarin zoveel

mogelijk genoeg scheidt. [...] Een wijze geniet met mate smakelijk voedsel, aangename dranken, geuren, de lieflijkheid van uitbottende planten, sieraden en ontspant zich met muziek, spel, theater en dergelijke dingen waar ieder mens van kan genieten zonder een ander schade te doen” (Ibid., E4P45S).

De mens kan volgens Spinoza nooit de volledige controle over zijn hartstochten krijgen, iets wat Spinoza poneert in het voorwoord op deel vijf van de *Ethica*. Hij keert zich tegen Descartes, wiens gedachtegoed door Spinoza wordt geïnterpreteerd op een zodanige manier dat wij “een volstrekte macht over onze passies” zouden hebben (Ibid., E5P).

In het laatste deel van de *Ethica* laat Spinoza zien waarin de eeuwigheid van de geest schuilt en hoe we de grootst mogelijke vorm van blijdschap kunnen bereiken, te weten de gelukzaligheid. In het kader van de eeuwigheid van de geest doet Spinoza in het commentaar op stelling 20 een op het eerste gezicht opmerkelijke uitspraak: “Hiermee is alles wat ons huidige leven betreft behandeld [...] Het wordt dus tijd om verder te gaan met de duur van de geest los van het lichaam” (Ibid., E5P20S). Deze zin is opmerkelijk, aangezien Spinoza zijn metafysica juist ageert tegen de traditionele theologische en cartesische doctrine van de onsterfelijkheid van de menselijke geest. Bovendien zijn lichaam en geest bij Spinoza weliswaar causaal autonoom, maar synchroon geordend, omdat zij als attributen op ontologisch niveau uitdrukkingen zijn van dezelfde substantie. Wanneer het lichaam teniet gaat, zou dan ook niet de geest vernietigd moeten worden?

In de eerste plaats dient opgemerkt te worden dat Spinoza niet pleit voor een onsterfelijkheid van de geest in persoonlijke zin. Met andere woorden: als persoon met inbegrip van cognitieve vermogens als de verbeelding of het geheugen. Hij betoogt echter voor de eeuwigheid van ideële aspecten van de menselijke geest in het intellect van God. Wanneer een mens sterft, vergaat niet de gehele geest, maar de eeuwige aspecten daarvan zullen als idee in het intellect van God blijven bestaan.

Spinoza lijkt vervolgens twee verschillende wijzen van eeuwigheid te onderkennen. In de eerste plaats bepleit Spinoza een vorm van eeuwigheid die niet heel relevant lijkt te zijn in zijn denken, omdat we er geen invloed op uit kunnen oefenen en deze vorm van eeuwigheid op alle modi van

toepassing is. Omdat het wezen der dingen – zowel het lichamelijke als geestelijke wezen – *sub specie aeternitatis* in God na de vernietiging ervan blijft behouden, blijft van ieder ding het eeuwige wezen ideëel bestaan, zelfs wanneer dat ding is gestorven of vergaan. Alle dingen bestaan na de dood of vernietiging ervan slechts als “lege huls” in God.

Maar Spinoza gaat vervolgens verder op eeuwige aspecten die uitsluitend de mens aan kunnen gaan en die wij met behulp van de rede wel kunnen beïnvloeden. Hij meent namelijk dat we de eeuwigheid van de geest kunnen beïnvloeden, afhankelijk van de intellectuele activiteit in ons leven. Naarmate een mens meer adequate ideeën heeft, blijft een groter deel van zijn geest eeuwig bestaan wanneer hij sterft. De intellectuele activiteit gedurende het leven is dus niet alleen nastrevenswaardig omdat het je gedurende je leven gelukkig maakt, maar ook omdat een groter deel van de geest dan eeuwig is. Het vergroten van de eeuwigheid van de geest kan worden gezien als een beloning voor een deugzaam leven.

Wat Spinoza uiteindelijk van ons vraagt, is om onder de leiding van de rede te leven. Als we deugzaam zijn, worden we minder aangedaan door externe oorzaken en leven we onder de leiding van hartstochten waarvan alleen wij zelf de oorzaak zijn. Door op intuïtieve wijze de werkelijkheid te kennen, ontwikkelen we uiteindelijk een verstandelijke liefde tot God, waarvan de hoogste vorm van geluk, namelijk de gelukzaligheid, een uitvloeisel is. Door onszelf en de wereld om ons heen te begrijpen zoals deze absoluut en eeuwig geordend zijn in het intellect van God, zijn we in staat absolute kennis te verkrijgen, op grond waarvan we op een zodanige manier zullen handelen dat we doen wat werkelijk – dus op zichzelf bezien – goed voor ons is.

Hoewel Spinoza erkent dat we nooit volledig kunnen samenvallen met het geschetste ideaalbeeld van de deugdzaam mens, heeft hij vanzelfsprekend wel vertrouwen in zijn ethische opvattingen: “Ook al zou de door mij gewezen weg tot deze dingen zeer moeilijk begaanbaar zijn, toch kan men haar vinden. En werkelijk: wat zelden wordt bereikt, moet moeilijk zijn” (Ibid., E5P42S). Het lijkt voor Spinoza niet noodzakelijk te zijn als mens volledig met het geschetste ideaalbeeld samen te vallen. Hij erkent namelijk dat de mens onmogelijk volledig zelf zijn activiteit kan bepalen. Spinoza bedoelt in mijn optiek eerder dat we een zekere mate

van gelukzaligheid kunnen bereiken. Gelukzaligheid is in dat kader een gradueel begrip. Spinoza sluit zijn *Ethica* af met de volgende beroemde woorden: “Alles wat voortreffelijk is, is echter even moeilijk als zeldzaam” (Ibid.).

Omdat duidelijk is op welke manier de mens volgens Spinoza in een volstrekt mechanistisch wereldbeeld gelukzaligheid kan bereiken, wordt nu overgegaan tot een bespreking van het leven van Koerbagh en van zijn meest systematische werk *Licht*.

3. Vrede door verlichting van het volk

3.1 Het leven van de gebroeders Koerbagh

De gebroeders Adriaan en Johannes Koerbagh werden respectievelijk in 1633 en 1634 beiden te Amsterdam geboren in een gezin van in totaal elf kinderen. De familie Koerbagh was vermogend, maar niet zeer rijk. Het vermogen van de familie bestond met name uit geld verkregen uit erfenissen en vastgoed, waarin door de familie werd belegd (Leeuwenburgh, 2013, p. 20). Dit vormt een belangrijk argument om aan te nemen dat de familie connecties had met de handels- en regentenelite van de toenmalige staat Holland (Ibid., pp. 19-20). Daarnaast hadden de gebroeders Koerbagh een omvangrijk intellectueel netwerk.

Voordat Adriaan en Johannes naar de universiteit gingen, doorliepen zij eerst de Latijnse school. Het Latijn vormde in de zeventiende eeuw de primaire taal van de universiteiten en de wetenschap. In 1653 gingen de gebroeders studeren aan de universiteit van Utrecht; Adriaan studeerde medicijnen en rechten en Johannes theologie. Hier maakten zij niet alleen kennis met het mechanistische, revolutionaire gedachtegoed van Descartes, maar ook met het door veel tijdgenoten als ketterij bestempelde socinianisme (Ibid., p. 41). De theoloog Socinus (1539 – 1604), geestelijk vader van deze christelijke doctrine, keerde de verhouding tussen geloof en rede om zoals deze geldt in het fideïsme: juist de waarheden die door de rede worden voortgebracht, prevaleren volgens hem als geloof en verstand strijdig zijn. Deze epistemologische houding werd door Socinus toegepast op het interpreteren van de Bijbel, met als gevolg dat

enkele belangrijke kerkelijke dogma's, zoals de triniteit, de goddelijkheid van Jezus, zijn wederopstanding en verlossing van de zonden door zijn dood, werden opgeslokt door de rede. Het socinianisme werd in de zeventiende eeuw populair onder vrijdenkers in de Republiek, zoals Jan Knol (? – 1672) en Frans Kuypers (1629 – 1691), die bevriend waren met de gebroeders Koerbagh (Ibid., p. 72). Onlangs bleek uit een inventarisatie van de boekencollectie van de gebroeders Koerbagh dat zij een fors aantal socinianistische werken bezaten (Van Bunge, 2011, p. 14). Ze waren dus goed bekend met de opvattingen van deze antitrinitaristische beweging.

In 1656 verhuisden Johannes en Adriaan naar Leiden om aldaar te studeren. De beweegredenen van het overstappen naar een andere universiteit waren waarschijnlijk tweeledig: in de eerste plaats omdat de medische faculteit van de Leidse universiteit in de zeventiende eeuw van zeer hoog niveau was en in de tweede plaats vanwege de grotere ontvankelijkheid voor het cartesians gedachtegoed. In 1659 promoveerde Adriaan in de medicijnen op basis van een onderzoek naar tuberculose. Hij promoveerde in de rechten in 1661, waarna hij naar Amsterdam verhuisde. Johannes voltooide uiteindelijk zijn studie theologie. Op dit moment waren Adriaan en Johannes beiden woonachtig in Amsterdam. De broers begonnen hier steeds meer te radicaliseren, met als gevolg dat Johannes zich meerdere malen voor de kerkenraad diende te verantwoorden (Leeuwenburgh, 2013, pp. 136-137). Uit de manier waarop hij zich voor de kerkenraad verdedigde, bleek al dat Johannes afwijkende – soms spinozistische – opvattingen huldigde: bijvoorbeeld dat de dingen om ons heen modificaties zijn van God als de ene substantie en dat een *creatio ex nihilo* een onmogelijkheid vormt (Van Bunge, 2011, p. 18).

Over het werk van de gebroeders Koerbagh kan het volgende worden opgemerkt. Van Johannes zijn geen overgeleverde werken bekend. Wel wordt vermoed dat hij co-auteur was van enkele weldra te noemen werken van zijn broer Adriaan (Koerbagh, 1668a, p. 239). Het meest noemenswaardige werk uit de vroegere periode van Adriaan Koerbagh is het *Nieuw Woorden-boek der Regten*, gepubliceerd in 1664. In dit werk worden - uiteraard doorgaans Latijnse - juridische begrippen op lexicografische wijze gedefinieerd. Hij trachtte hiermee elitaire, academische kennis te ontdoen van haar Latijnse huls om het gewone

volk binnen de Republiek niet langer in wetenschappelijke onwetendheid te laten. In 1668 publiceerde Koerbagh zijn werk *Een bloemhof van allerley lieflijkheyd sonder verdriet*; een tevens op lexicografische wijze in spottende stijl uiteengezette kritiek op de religieuze status quo. Maar ook medische en wederom juridische begrippen werden gedefinieerd. *Bloemhof* werd – opmerkelijk genoeg – zowel onder zijn echte naam als onder het pseudoniem Vreederijk Waarmond gepubliceerd. Het verkreeg direct het predicaat “godslasterlijk” en werd door tijdgenoten ervaren als zeer schokkend. Johannes hielp zijn broer bij de publicatie van zijn tweede werk: *Licht* (1668). *Licht* kan beschouwd worden als systematische voortzetting van ideeën die al in *Bloemhof* zijn beschreven (Leeuwenburgh, 2013, p. 182).

Toen de beoogde drukker echter notie kreeg van de buitengewoon radicale inhoud van *Licht*, besloot hij het werk niet verder te drukken en de autoriteiten in te lichten over dit godslasterlijke werk. Inmiddels was Adriaan Koerbagh gevlucht en ondergedoken. Hij werd echter verraden en vervolgens gearresteerd door de schout van Leiden. Ondanks het feit dat Koerbagh een netwerk had bestaande uit invloedrijke personen in de Republiek, werd hij toch veroordeeld tot een gevangenisstraf van tien jaar die hij moest uitzitten in het Rasphuis, gevolgd door een verbanning uit de stad Amsterdam voor tien jaar. Bovendien werd er een geldboete van 6000 gulden aan hem opgelegd. Omdat van Johannes niet kon worden bewezen dat hij medeauteur was van de werken van Adriaan, werd hij ontslagen van alle rechtsvervolging. Adriaan stierf na slechts negen maanden gevangenschap in het Rasphuis. Johannes bleef echter deelnemen aan de bijeenkomsten van de collegianten totdat hij in 1672 kwam te overlijden (Van Bunge, 2011, p. 33).

3.2 *Licht*: een algemene analyse

Omdat een goed begrip van *Licht* afhangt van het begrip van de titel van dit werk, de uiteindelijke motivatie en doelstelling van Koerbagh, zijn epistemologische houding en methodologie, zal ik nu eerst op deze zaken ingaan.

De titel van *Licht* luidt in volledigheid als volgt: *Een licht dat schijnt in duistere plaatsen. Een verheldering van de voornaamste kwesties van theologie*

en *godsdiens*t (Koerbagh, 1668a, p. 5). Voor wat betreft de metaforische wijze waarop “licht” en “duisternis” hier gebruikt worden, was Koerbagh geïnspireerd door passages uit de Bijbel. Hiervan getuigt één van de motto’s van het werk, afkomstig uit Efeziërs 5:13: “Al wat openbaar maakt is licht” (Ibid., p. 6). Hier voegt Koerbagh aan toe: “De waarheid maakt openbaar. Dus de waarheid is licht” (Ibid.). Het licht staat bij Koerbagh symbool voor de rede, waarmee de mens de absolute waarheid kan ontdekken. Hij laat dit licht vervolgens schijnen op een groot aantal “duistere” kwesties, waarmee hij doelt op allerlei vormen van bijgeloof en verzinsels die deel uitmaken van de gevestigde religieuze orde. Deze duisternis heeft volgens Koerbagh meerdere oorzaken. In de eerste plaats de “duisterheid en verwardheid” van de Heilige Schrift (Ibid., p. 8). De Bijbel bevat naar de opvatting van Koerbagh namelijk inconsistenties, contradicties en dus onwaarheden. Theologen en geestelijken hebben de Bijbel vanuit geldbejag en machtslust op een onjuiste manier geïnterpreteerd en onware conclusies getrokken. Door deze bewuste misinterpretaties en de duisternis van de Bijbel wordt juist een onderlinge haat gepropageerd. De vele interpretaties van de Heilige Schrift zijn uitgemond in een veelheid aan religieuze stromingen, van rooms-katholieken, die door Koerbagh het meest worden veracht, tot calvinisten, lutheranen, quakers, socinianen, Joden en brownisten. De intense onderlinge haat die voortvloeit uit al deze gedifferentieerde claims op de ware religie of de waarheid, leidt tot oorlogen en de dood, wat nu juist in strijd is met datgene wat de Schrift en Jezus ons leren.

Koerbagh hoopt in *Licht* tot een redelijke godsdiens

t te komen om op die manier een einde te maken aan de onderlinge haat en de godsdienstige conflicten die hieruit voortvloeien. Het pseudoniem waarmee Koerbagh zijn voorwoord eindigt, Vreederijk Waarmond, duidt er op dat hij door middel van het spreken van de waarheid vrede wil bewerkstelligen. Godsdienst heeft bij Koerbagh echter absoluut niet uitsluitend een pacifistische functie: zij is omwille van zichzelf nastrevenswaardig en leidt ook tot ware gelukzaligheid en het eeuwige leven. In het licht van deze motivaties is het voor hem noodzakelijk niet alleen te ageren tegen geleerde geestelijken en theologen. Hij wenst namelijk het gehele Nederlandse volk op grond van de rede de absurditeit en tegenstrijdigheid van voorname theologische noties bij te brengen.⁶ Om die reden besloot Koerbagh zijn werk in het Nederlands te schrijven in plaats van

het academische Latijn.⁷ Hij hanteerde de Nederlandse taal als een purist. Dit was voor hem een manier om de taal als communicatief en politiek medium te zuiveren van obscure, dubbelzinnige leenwoorden.

Wat betreft zijn epistemologie geldt dat Koerbagh een rationalist is in hart en nieren. Hij meent dat geloof gegrond is op vertrouwen: het kan je bedriegen (Ibid., p. 35). Wie de rede niet gebruikt, vervalt tot bijgeloof, dat de bron is van al het kwaad (Ibid., p. 14). En alleen op grond van de universele, eeuwige en onveranderlijke rede kan men met elkaar overeenstemmen.

Wat opvalt aan de methodologie in het gefragmenteerde *Licht*, is dat zij gekenmerkt wordt door een in heldere taal en met temperament geschreven betoog, soms in spottende stijl, gegrond op een etymologisch onderzoek van theologische begrippen. De heldere taal vormt een aanklacht tegen theologen en hun bovennatuurlijke, duistere verzinsels. Bovendien sluit dit taalgebruik aan op het publiek van Koerbagh. Ook weerspiegelt het de epistemologische idee dat we uitsluitend datgene voor waar aan moeten nemen wat we helder en welonderscheiden kennen (Ibid., p. 34). De door Koerbagh uitgevoerde rationeel-etymologische herinterpretatie van theologische begrippen verloopt in veel gevallen volgens een zeker stramien. Ten eerste laat Koerbagh zien wat de gevestigde theologische orde onder een bepaald begrip verstaat. Vervolgens toont hij de etymologische achtergrond van het desbetreffende concept en legt hij dus de daadwerkelijke taalkundige betekenis van het begrip bloot. Hij laat hiermee zien dat de theologen het begrip verkeerd hebben geïnterpreteerd. Na deze deconstructie stelt Koerbagh vast of de gezuiverde Bijbelse notie al dan niet strookt met de rede. Indien het begrip strookt met de rede, dan besluit Koerbagh dat dit de ware aard is van de notie. Indien dit niet het geval is, dan poogt hij het begrip op een constructieve wijze ermee in overeenstemming brengen, bijvoorbeeld door Bijbelse passages te hertalen.⁸

Ten slotte hanteert Koerbagh tweemaal de geometrische methode in *Licht*. Hij past de methode toe in de context van de triniteit en het bestaan van boven- of tegennatuurlijke wonderen. Koerbagh hanteert de geometrische methode niet zozeer om constructieve theses te vormen, maar eerder om de posities van theologen te vernietigen (Wielema, 2001, p. 66). Zo verloopt zijn argumentatie *more geometrico* als Koerbagh stelt dat de triniteit een rationele tegenstrijdigheid betreft. Voor zijn actieve ideeën wordt de geometrische methode niet gehanteerd.

3.3 De herinterpretatie van God en substantie

In het eerste hoofdstuk van *Licht* zet Koerbagh zijn godsbeeld uiteen dat als fundament geldt voor andere belangrijke thema's in dit werk, zoals de Heilige Schrift, de triniteit en de positie van Jezus. De door Koerbagh uiteengezette substantieleer wordt voornamelijk genoemd binnen een specifieke socinianistische context, namelijk wanneer hij de goddelijke drie-eenheid tracht te deconstrueren (Van Bunge, 2011, pp. 23-24). Het is daarom noodzakelijk om eerst dit concept te behandelen, voordat wordt toegekomen aan andere noties die in *Licht* worden geherinterpreteerd.

De ene substantie wordt door Koerbagh geïdentificeerd met God of Wezen. De etymologische houding van Koerbagh wordt hier al goed duidelijk. In het cartesiaanse begrippenapparaat is het gebruikelijk het Latijnse *substantia* te vertalen met substantie of datgene wat ontologisch en epistemologisch onafhankelijk of zelfstandig is. Koerbagh meent echter dat dit een onjuiste vertaling vormt van *substantia*: het Nederlandse onderstandigheid betreft een betere vertaling. De daadwerkelijke etymologische betekenis van het begrip *substantia*, dat wil dus zeggen: onafhankelijkheid, wordt gekoppeld aan een niet werkelijk bestaande Latijnse term, namelijk *ipstantia*. God of Wezen worden dus in het denken van Koerbagh geïdentificeerd met de ipstantie, waaruit oneindig veel substanties of onderstandigheden voortvloeien, die allen bestaanswijzen zijn van die ene ipstantie. Dat er maar één ipstantie kan zijn, blijkt volgens Koerbagh uit de dagelijkse ervaring, die onderstreept dat er buiten Wezen niets volledig op zichzelf bestaat en kan bestaan. Iets wat volledig onafhankelijk is, is namelijk ook eeuwig onafhankelijk. Als wij de wereld om ons heen beschouwen, dan merken we echter dat alles afhankelijk is van God, wat door de Schrift wordt geverifieerd.

3.4 Natuurkunde of filosofie als ware theologie

Deze monistische substantieleer bepaalt de verhouding tussen de verschillende wetenschappen in het denken van Koerbagh. Omdat er maar één substantie is, omvatten de natuurkunde of filosofie, die in het denken van Koerbagh worden vereenzelvigd, alle wetenschappen (Wielema, 2001, p. 59). De werkelijke, etymologische betekenis van “metafysica” is voor Koerbagh “bovennatuurkunde”, waarvan hij vaststelt dat het een onzinnige

wetenschap is, gevuld met onwaarheden, omdat er niets buiten God of de natuur bestaat (Koerbagh, 1668a, p. 220). Metafysica wordt echter – in tegenstelling tot de werkelijke vertaling daarvan met “bovennatuurkunde” - doorgaans begrepen als bestudering van de aard van de realiteit, waarmee “hyperfysica”, oftewel “na-natuurkunde” of “achternatuurkunde”, dus een taalkundig meer zuivere term zou zijn. Deze werkelijke betekenis van de metafysica, zoals wij deze in de filosofie kennen, of voor Koerbagh de hyperfysica, vereenzelvigt hij met de natuurkunde, omdat het niets anders betreft dan de bestudering van de aard van de natuur. Men zou God middels de natuurkunde of filosofie moeten bestuderen en niet met de traditionele theologie. Koerbagh meent namelijk verderop dat de natuurkunde “de echte en ware godgeleerdheid is” (Ibid., p. 226). De filosofie omvat vervolgens “alle waarheid, wijsheid, wetenschap en kennis die er maar is” (Ibid., p. 17). Hieruit volgt dat alles wat waar is in het domein van de theologie, ook waar moet zijn in de filosofie, en vice versa.

3.5 Het Woord van God

Dat Koerbagh de Bijbel niet als theoloog maar als filosoof interpreteert, wordt goed duidelijk wanneer hij ingaat op de verhouding tussen dit werk en het werkelijke Woord van God. Hij interpreteert de Bijbel op een eigenaardige wijze. Zo meent hij dat de schrijvers van de Heilige Schrift klaarblijkelijk niet wilden dat de lezers zouden kunnen begrijpen wat ze hadden geschreven, of dat ze onwetend waren en niet erg geleerd of dat zij de waarheid niet in de openbaarheid uit wilden drukken (Ibid.). Deze idiosyncratische lezing van de Bijbel doet de vraag rijzen wat nu precies de status van dit werk is volgens Koerbagh. De Bijbel lijkt aan de ene kant een waardevol werk voor hem te zijn, want hij beroept zich er in *Licht* regelmatig op. Aan de andere kant schuwt Koerbagh niet een zeer kritische houding tegenover de “ware” aard van de Bijbel aan te nemen: hij spreekt immers over de “duisterheid” (Ibid., pp. 128-129) van de Bijbel, ontkent dat de zondeval iets is wat letterlijk moet worden opgevat en hetzelfde geldt voor de wonderen die in de Schrift worden genoemd (Ibid., p. 229), zoals de splinging van de Rode Zee door Mozes (Ibid., p. 227). De Bijbel bevat dus ook onwaarheden, contradicties en metaforen. Vormt de Bijbel voor Koerbagh dan niet een uitdrukking van het Woord van God? En is de voornoemde lezing van Koerbagh een opportunistische interpretatie van de Heilige Schrift (Van Bunge, 2011, p. 25)?

Gelukkig wijdt Koerbagh in *Licht* een separaat hoofdstuk aan de status van de Bijbel in het algemeen en de relatie van dit werk tot het Woord van God. Het woord Bijbel betekent voor Koerbagh niets anders dan “boek”. In *Bloembhof* had Koerbagh de Bijbel ook al in dezelfde categorie geplaatst als reguliere boeken, afkomstig van mensen, zoals van “reyntje de vos of uylen-spiegel” (Koerbagh, 1668b, p. 95). In *Licht* wordt de Bijbel wederom onttrokken aan de sfeer van absolute waarheid en heiligheid.

Ook hier oppert Koerbagh weer een opmerkelijk idee: de gedachte dat de Schrift voor verbetering vatbaar is. Het is volgens hem juist strijdig met de almacht van God indien een verbeterde versie van de Bijbel onmogelijk zou zijn. Koerbagh stelt hierover: “Alsof het een ijzeren wet is dat men iets niet mag verbeteren, maar voortdurend in de oude drek van onwetendheid moet blijven liggen wentelen” (Koerbagh, 1668a, p. 129).

Deze opvatting dient in het licht te worden beschouwd van zijn opvatting aangaande het werkelijke Woord van God, de rede, waarin God zich tegenover de mens heeft geopenbaard. Daardoor is de rede als goddelijk te beschouwen: de mens kan middels de goddelijke, voor alle volkeren universele rede kennis van God verkrijgen (Ibid., p. 14). En deze rede wordt bij Koerbagh vereenzelvigd met het Woord van God, omdat het gezien de aard van God, zoals eerder beschreven, niet anders kan zijn dan dat Zijn Woord ook eeuwig, onveranderlijk en waar is.

De Bijbel kan dan alleen het Woord van God vormen voor zover de Schrift elementen bevat die onveranderlijk, eeuwig en waar zijn en dus samenvallen met het ware Woord van God. In dit kader wordt duidelijk hoe Koerbagh kan stellen dat de Bijbel voor verbetering vatbaar is. En dit verklaart ook waarom Koerbagh het zich kan veroorloven om de duistere of tegenstrijdige passages van de Schrift te hertalen en in overeenstemming te brengen met de rede. Er mag dan ook geen misvatting over bestaan dat de Bijbel bij Koerbagh niet geheel zonder waarde is: het vormt een “fragment of gedeelte van Gods Woord” (Ibid., p. 135). Wel staat deze waarde in schril contrast tot de absolute wijze waarop de theologen de Schrift als equivalent van het Woord van God beschouwen.

3.6 Gelukzaligheid en eeuwigheid

Het grote vertrouwen dat Koerbagh in de menselijke rede heeft, komt voort uit de zojuist besproken opvattingen aangaande de verhouding tussen de menselijke en de goddelijke rede. Maar de rede Gods moet anders opgevat worden dan die van de mens: het Woord van God omvat namelijk een oneindige verzameling van eeuwige, ware kennis en wijsheid. Een cruciale epistemologische vraag die hier opdoemt blijft in *Licht* echter onbeantwoord: hoe kan de mens met zijn beperkte rede nu samenvallen met de volgens Koerbagh niet naar de gelijkenis van de mens op te vatten (Ibid., p. 134) volmaakte rede van God (Wielema, 2003, p. 66)?

Deze cruciale vraag is juist problematisch omdat Koerbagh stelt dat de ware, onveranderlijke gelukzaligheid of de grootste vorm van geluk te bereiken is door middel van de rede, waarmee we kennis van God kunnen verkrijgen. In deze vereenzelviging met het Woord van God bestaat volgens hem het eeuwige leven (Koerbagh, 1668a, pp. 45, 151-152). Dit doet de vraag rijzen hoe deze eeuwigheid van de ziel kan worden begrepen. Wie geen volmaakte kennis van God bezit, “sterft zoals de beesten” (Ibid., p. 57), waarmee Koerbagh lijkt te geloven dat alleen de mens een bezielde, maar ook redelijk wezen is. Dit duidt echter alleen nog maar op een mentale werkelijkheid na de dood. Hij acht ook een leven na de dood mogelijk. De mens kan op een bepaalde manier blijven voortbestaan nadat het lichaam teniet is gegaan: “Het is vast en zeker dat als we uit de dood zullen zijn opgestaan, wij een goede en geschikte woonplaats zullen hebben” (Ibid., p. 152). Koerbagh spreekt bovendien over “het eeuwige leven” van de rechtvaardige mens in tegenstelling tot “de eeuwige dood” van de beesten (Ibid., p. 57). Wanneer Koerbagh het begrip “hemel” bespreekt, lijkt hij te aanvaarden dat niet alleen een mentaal aspect, maar ook een lichamelijk aspect na onze dood zal blijven voortbestaan, en dat beide aspecten een bepaalde fysieke plaats hebben in de werkelijkheid na de dood, waarvan hij vermoedt dat dit de aarde is (Ibid., p. 151). En als Koerbagh het begrip “hel” analyseert, meent hij dat de werkelijke hel de wroeging van de ziel is over het mislopen van het eeuwige leven (Ibid., p. 161).

Deze mogelijkheid tot een wederopstanding na de dood vormt bij Koerbagh een problematisch concept: hoe is zij immers verenigbaar met zijn godsbeeld en substantieel? Hoe is het mogelijk dat de lichamelijke van de mens na de dood behouden blijft? Het wordt in *Licht* niet goed duidelijk hoe Koerbagh deze andere gesteldheid na de dood mogelijk acht.

3.7 Naar een universele godsdienst

Het verstand is voor Koerbagh niet alleen noodzakelijk voor het bereiken van de gelukzaligheid. In het licht van zijn meer belangrijke pacifistische missie is de rede ook het instrument waarmee de mens vrede kan bewerkstelligen. Op grond van de rede kan men tot een ware godsdienst komen, waar er maar één van kan bestaan. Alle versplinterde godsdienstopvattingen kunnen middels de rede gezuiverd worden van hun onwaarheden zodat de opvattingen van gelovigen overeenstemmen. Omdat Koerbagh meent dat alle gelovigen de eigen godsdienst als de enige ware beschouwen, pleit hij in *Licht* voor een onpartijdige godsdienstrecht, die zich weet te abstraheren van de georganiseerde godsdiensten. En hij meent dat deze onpartijdige rechter in dit kader vooral afstand dient te nemen van het bestaan van wonderen: een belangrijke bron van verschillen tussen de diverse geloofsovertuigingen is immers gelegen in hun ideeën over wonderen. Religies aanvaarden de waarheid van elkanders wonderen niet; dit doen zij op grond van geloof, niet op grond van de rede (Ibid., pp. 96-97).

Koerbagh vervolgt zijn betoog door zelf als rechter de diverse geloofsovertuigingen binnen het christendom te analyseren, waaronder het katholicisme, protestantisme en socinianisme, waarbij vooral de rooms-katholieken het moeten ontgelden. Het zijn de socinianiërs die van alle christelijke stromingen uiteindelijk het dichtst bij de ware godsdienst staan, omdat zij de triniteit ontkennen en het verrichten van zelfstandig onderzoek niet alleen toestaan, maar dit ook stimuleren (Ibid., p. 122): “Als de drie-eenheid waar was, zouden ze [d.w.z.: de theologen en predikanten] niet zeggen: verricht geen onderzoek maar geloof ons. Ze zouden juist zeggen: onderzoek en u zult het ontdekken. Maar omdat de drie-eenheid onjuist is, en een tegenstrijdigheid, verbieden zij de mensen die te onderzoeken en bevelen hun die te geloven” (Ibid., pp. 121-122). De socinianiërs gaan echter niet ver genoeg voor wat betreft hun redelijke interpretatie van

het christelijk geloof: hoewel zij de triniteit verwerpen, aanbidden zij Jezus nog, wat niet alleen strijdig is met de Schrift, omdat die volgens Koerbagh leert dat we Jezus alleen mogen verheerlijken en niet aanbidden, maar ook met de rede, omdat alleen God zelf aanbeden mag worden.

Op grond hiervan is Koerbagh dus niet te beschouwen als atheïst, maar juist als diep religieus mens, wiens theologische opvattingen sterk afwijken van de gevestigde orde. We moeten Koerbagh met andere woorden niet verkeerd interpreteren door te denken dat zijn pleidooi voor religieuze eenheid uitsluitend een sociaal-politieke functie van religie omvat (Wielema, 2001, p. 74).

Licht eindigt abrupt. Koerbagh stelt aan het einde van dit werk dat hij tot God bidt om een goede gezondheid, zodat hij het tweede deel van *Licht* zal kunnen publiceren. Koerbagh stierf echter kort na zijn poging tot publicatie van *Licht*, nadat hij negen maanden in gevangenschap had doorgebracht. Een van de meest radicale werken uit de periode van de vroege verlichting heeft zijn vervolg dus nooit gehad.

4. De intellectuele zelfstandigheid van Adriaan Koerbagh

De *Ethica* en *Licht* zullen nu inhoudelijk met elkaar worden vergeleken.⁹ Vervolgens zal ik in paragraaf 4.5 ingaan op de relaties tussen Koerbagh en Spinoza, op grond waarvan ik in paragraaf 4.6 de onderzoeksvraag van dit essay zal beantwoorden.

4.1 Overeenkomsten

4.1.1. Metafysica

Het staat vast dat het denken van Spinoza belangrijke overeenkomsten vertoont met het denken van Koerbagh. De meest opvallende en belangrijke overeenkomst tussen de denkers bestaat uit hun metafysica; hun substantiëmonistische beeld van God. Zowel in de *Ethica* als in *Licht* wordt God volledig genaturaliseerd en vereenzelvigd met de substantie en de natuur. Beide denkers ontkennen op grond hiervan het bestaan van wonderen als bovennatuurlijke fenomenen: er is immers niets boven of tegen de natuur.

Er bestaat ook een sterke parallel tussen de eigenschappen van God in het werk van de beide denkers: Hij is immanent, oneindig, eeuwig, nooit begonnen, nooit eindigend, eenvoudig, enig, onveranderlijk, volmaakt, almachtig en alwetend. Er moet bij Spinoza en Koerbagh voor worden gewaakt dat God en dus ook zijn voornoemde eigenschappen niet op een antropomorfe manier begrepen worden, al is deze missie in de *Ethica* meer prangend en expliciet dan in *Licht*.

Zowel Spinoza als Koerbagh reduceren al hetgeen bestaat op ontologisch niveau niet uitsluitend tot materie. Bij hen is God namelijk een substantie, die zich op een oneindig aantal wijzen uitdrukt. Dit doet Hij niet alleen op materiële, maar ook op ideële wijze. Beide denkers geloven bovendien dat de ziel eeuwige aspecten heeft die na het tenietgaan van het menselijke lichaam behouden blijven.

4.1.2. Epistemologie

Een andere belangrijke overeenkomst bestaat uit het extreem rationalistische karakter van de wijsgerige werken *Ethica* en *Licht*. Spinoza en Koerbagh menen allebei namelijk dat het verstand van God op een rationele manier gestructureerd is: Hij kent een ideëel aspect waarin de volledige en absolute waarheid op een rationele wijze is vastgelegd. Dit ideële aspect bestaat uit de oneindige verzameling van ware proposities over de werkelijkheid. Het grote vertrouwen dat beide denkers in de menselijke rede hebben vloeit uit deze rationele geordendheid van God voort. Bij Koerbagh is dit vertrouwen zelfs zo groot dat hij de rede een goddelijke status toekent. Uiteindelijk gaat het erom dat beide denkers menen dat de mens (gedeeltelijk) samen kan vallen met dit goddelijke intellect.

4.1.3. Ethiek

Ook opvallend is dat bij zowel Spinoza als Koerbagh ethiek, deugdzaamheid en de zojuist genoemde kennistheoretische eenheid van de mens met God intrinsiek met elkaar samenhangen. De grootste vorm van gelukzaligheid is niet gelegen in het nastreven van aardse zaken. Deze zaken zijn immers vergankelijk, terwijl het geluk dat God bij de mens teweeg kan brengen van eeuwige aard is. Noch Spinoza, noch Koerbagh meent echter dat we volledig afstand van aardse zaken moeten nemen.

Wel moeten we de (passieve) hartstochten matigen, omdat deze niet tot werkelijke gelukzaligheid leiden. Bij beide denkers vormt het menselijke verstand het middel waarmee we deze passies kunnen beheersen.

De rede vormt het middel waarmee we het grootste geluk kunnen bereiken; het meest werkelijke en hoogst mogelijke geluk ervaren wij wanneer we intellectueel overeenstemmen met het absoluut ware intellect van God. Hiermee is de gelukzaligheid in feite gelegen in een adequate kennis van de natuur. Wanneer wij kennis van de natuur verkrijgen, zien we pas hoe de dingen werkelijk zijn en dus wat werkelijk goed voor ons is. Ook heeft in dit kader het aardse leven in zowel *Licht* als de *Ethica* een essentiële plaats voor wat betreft de eeuwigheid van de ziel, waar beide denkers zwaar aan tillen. Het is niet – voor zover een leven na de dood bestaat – zonder meer zo dat wij in dat hiernamaals het grootste geluk zullen ervaren. Het is juist het aardse leven waarin wij ons geluk zullen moeten vinden en het zijn juist de gelukzaligheid en kennis in het aardse leven die onze positie na het vergaan van het lichaam bepalen.

4.2 Verschillen

4.2.1. Filosofische agenda's

Het meest essentiële verschil tussen de beide denkers bestaat erin dat de doelstelling van de *Ethica* en de filosofische agenda van Spinoza aanzienlijk verschillen ten opzichte van de doelstelling van *Licht* en van de agenda van Koerbagh. Dit verschil in doelstellingen vormt niet alleen op zichzelf gezien een belangrijk verschil, maar het heeft ook belangrijke consequenties voor de gehanteerde methodologieën, de filosofische onderwerpen die voor Spinoza en Koerbagh meer of minder relevant zijn en de wijze waarop dit allemaal uitgedrukt wordt in hun concrete ideeën.

De *Ethica* heeft uiteindelijk een ethische doelstelling. Spinoza was zeer pessimistisch over de verlichting van de gewone mens: de meeste mensen zijn niet in staat hun rationele capaciteiten voldoende te ontwikkelen en toe te passen. Onder de mensen die de gelukzaligheid mogelijk zouden kunnen bereiken, valt dus beslist niet de gewone man uit de zeventiende eeuw (Leeuwenburgh, 2013, pp. 130-131). De *Ethica* is om die reden in het academische Latijn geschreven. Kritiek op de gevestigde theologische orde ligt in de *Ethica* verscholen achter een complex,

op meetkundige wijze uiteengezet filosofisch systeem (Wielema, 2001, p. 59). Bij Koerbagh is dit volstrekt anders. Zijn doel is politiek en hervormend van aard. Ethiek en geluk zijn voor Koerbagh belangrijke onderwerpen, maar zij vormen zeker niet het ultieme doel van *Licht*. Hij wil het gewone volk verlichten en voor eens en altijd de religieuze conflicten van zijn tijd, gevoed door gedifferentieerde claims op de juiste religie, opheffen, door een directe aanval te lanceren op de geestelijken en hun corrumperende begrippen.

Gezien dit verschil in filosofische doelstellingen, vestigen Spinoza en Koerbagh hun aandacht op andere filosofische deelgebieden, wat het contrast tussen hen vergroot. In het licht van datgene wat Spinoza en Koerbagh betogen, laten ze ook onderwerpen buiten beschouwing die zij voor hun theses niet noodzakelijk achten. Waar Spinoza bijvoorbeeld veel aandacht besteedt aan de (deterministische) werkingswijze van de natuur, teleologie en de mogelijkheid van vrijheid, daar laat Koerbagh deze onderwerpen in *Licht* geheel buiten beschouwing. Dit geldt ook voor de noodzakelijkheid van het bestaan van God, wat bij Spinoza uitgebreid wordt aangetoond, omdat dit onlosmakelijk verbonden is met zijn deterministische inzet. Koerbagh besteedt in tegenstelling tot Spinoza bijvoorbeeld weer meer aandacht aan het hertalen van theologische begrippen en het afbreken van de triniteitsleer.

De substantieel leer van Koerbagh staat uitsluitend in dienst van de deconstructie van de triniteit. Deze leer contrasteert tevens op het punt dat Koerbagh vanuit zijn taalpurisme en etymologische houding het begrip “ipstantie” als “zelfstandigheid” definieert en “substantie” als “onderstandigheid”. Als uitvloeisel van de verschillende motivaties in de *Ethica* en *Licht* zou men de in paragraaf 4.2.1. genoemde metafysische gelijkenissen dus moeten nuanceren.

4.2.2. Methodologie

Een ander essentieel verschil is gelegen in de sfeer van de filosofische methodologie. Het betoog in *Licht* verloopt volgens een rationeel-etymologisch stappenplan, gevoed door taalpurisme. Hoewel *Licht* niet geometrisch is opgebouwd, hanteert Koerbagh deze methode wel op een tweetal plaatsen, waardoor dit op het eerste gezicht een belangrijke

overeenkomst tussen de beide denkers lijkt aan te duiden. Hij gebruikt de methode echter uitsluitend in twee specifieke contexten, namelijk in die van de goddelijke drie-eenheid en die van wonderen. Tevens heeft de geometrische methode in *Licht* meer een deconstruerende functie: het vormt het gereedschap waarmee Koerbagh de genoemde noties vernietigt. Hiermee verschaft hij zich zijn eigen invulling aan de geometrische methode en dit aspect vormt eerder een verschil dan een overeenkomst tussen de beide denkers.

De *Ethica* is een gestructureerd werk, wat samenhangt met haar meetkundige ordening, op grond waarvan het zich op deductieve wijze ontvouwt. Het werk van Koerbagh is daarentegen meer chaotisch en gefragmenteerd: onderwerpen als epistemologie, gelukzaligheid en het eeuwige leven worden niet behandeld in een afzonderlijk hoofdstuk, maar verspreid over het gehele werk. De temperamentvolle en daardoor ook chaotische stijl van Koerbagh staat in dienst van zijn hervormende doelstelling.

4.2.3. De waarheidsdomeinen van de theologie en de filosofie

Een ander belangrijk verschil is gelegen in de religieuze verbeelding en het intact houden van de mogelijkheid van een theologie die haar eigen waarheden kent, al zijn deze strijdig met de rede. De theologie gaat bij Koerbagh volledig op in de filosofie en de natuurkunde. Wat waar is in de theologie, moet ook waar zijn in de filosofie en natuurkunde, en vice versa. Voor zogenaamde bovennatuurlijke waarheden, of zoals Koerbagh het vanuit zijn etymologie noemt: “metafysica”, is in zijn denken geen enkele plaats. Spinoza lijkt aan de andere kant wel een bepaalde mate van godsdienstige verbeelding in zijn denken toe te laten, alhoewel dit niet zozeer aan de orde komt in de *Ethica*, maar in de *Tractatus Theologico-Politicus*. Een belangrijk thema in dit werk bestaat namelijk uit de gedachte dat de domeinen van de filosofie en de theologie strikt van elkaar gescheiden dienen te worden (Nadler, 2002, p. 239). Deze twee domeinen hebben hun eigen epistemologische objecten: zo ressorteren waarheden van metafysische aard onder de filosofie en valt de interpretatie van de Bijbel, waar Spinoza uitsluitend een moreel belang in ziet, onder de theologie. Spinoza aanvaardt dus wel dat theologische waarheden, die mogelijk boven of tegen de rede zijn, kunnen bestaan.¹⁰

4.2.4. De eeuwigheid van de ziel

Over de eeuwigheid van de ziel verschillen Spinoza en Koerbagh uiteindelijk ook van opvatting. Spinoza keert zich sterk tegen de cartesiaanse idee van een mogelijke scheiding van lichaam en ziel en het behoud van de verbeelding na de dood. Een wederopstanding uit de dood is bij Spinoza onmogelijk. Alleen een ideëel aspect van de mens blijft na de dood bestaan, namelijk de adequate ideeën *sub specie aeternitatis* die hij gedurende zijn leven heeft verkregen.

Koerbagh verruimt de aspecten van de mens die na de dood behouden blijven. Hij meent dat een wederopstanding van de mens uit de dood wel mogelijk is. De mens is volgens *Licht* namelijk in staat op te staan uit de dood en het eeuwige leven te hebben; de herrezen mens kent een bepaalde plaats, namelijk de hemel, waar hij het grootste geluk zal ervaren. Volgens Koerbagh is het, in tegenstelling tot Spinoza, ook nog alleen de mens die deze eeuwige aspecten kan bezitten; alle andere aardse wezens lijken in het denken van Koerbagh niet bezielde te zijn.

4.3 Welke conclusie kan uit deze vergelijking worden getrokken?

Uiteindelijk kan naar aanleiding van deze systematische vergelijking de conclusie worden getrokken dat de overeenkomsten tussen de beide denkers, wanneer zij aan een grondige analyse worden onderworpen, minder relevant zijn voor de uiteindelijke posities die zij in hun werk innemen, en bovendien bij nader inzien vaak aanleiding geven om toch van een verschil te spreken.

4.4 De relatie tussen Spinoza en Koerbagh

Toch kan niet ontkend worden dat *Licht* mede door Spinoza geïnspireerd is, gegeven de inhoudelijke verwantschap van hun werk in samenhang met hun gemeenschappelijke vriendenkring. Maar de aanwezigheid van overeenkomsten tussen ideeën impliceert nog geen afhankelijkheid. In welke mate is Koerbagh dan door Spinoza beïnvloed en heeft hij ideeën aan hem ontleend? Deze vraag is zeer moeilijk met zekerheid te beantwoorden. Koerbagh laat in *Licht* nauwelijks expliciet zien door welke denkers hij geïnspireerd is of tegen wie van hen hij ageert. In de zeldzame gevallen dat hij refereert naar andere denkers, citeert hij bijvoorbeeld

Plinius en Hobbes' *Leviathan*. Spinoza en Koerbagh verwijzen in ieder geval niet naar elkaar.

Daarnaast blijkt uit de recentelijk geïnventariseerde boekenverzamelingen van zowel Spinoza als Koerbagh dat zij geen boeken van elkaar bezaten (Van Bunge, 2001, p. 28). Dan is er nog het feit dat het werk van Koerbagh enkele jaren eerder is gepubliceerd dan het werk van Spinoza. *Bloemhof* is gepubliceerd in 1668; Koerbagh had het voornemen *Licht* in datzelfde jaar te publiceren. De *Tractatus Theologico-Politicus* is echter pas in 1670 gepubliceerd; de *Ethica* pas postuum in 1678. Sommige auteurs, waaronder Israel (Israel, 2001) en Nadler (Nadler, 2011), wijzen erop dat Koerbagh een vroeg manuscript van Spinoza's *Tractatus Theologico-Politicus* heeft bestudeerd of de *Korte Verhandeling*. Al lijkt dit inderdaad aannemelijk, toch moet men zich realiseren dat ook dit eerder berust op vermoedens en veronderstellingen dan op feitelijkheden. Voor 1670 hadden enkele goede vrienden van Spinoza namelijk wel toegang tot een conceptversie van het eerste deel van de *Ethica* en de *Korte Verhandeling*, maar niet de *Ethica* of de *Tractatus Theologico-Politicus* zelf (Van Bunge, 2011, p. 29). Bovendien is niet duidelijk of Koerbagh tot deze groep vrienden behoorde.

Wel staat vast dat Spinoza en Koerbagh elkaar kenden; waarschijnlijk waren zij ook bevriend (Leeuwenburgh, 2013, pp. 115-120). Dit blijkt onder meer uit de bekentenis die Koerbagh had afgelegd toen hij op grond van de publicatie van *Bloemhof* en de voorgenomen publicatie van *Licht* werd gearresteerd. In deze bekentenis stelde hij dat hij Spinoza weliswaar had bezocht, maar nooit met hem had gedelibereerd over filosofie en theologie. Ook zijn er veel indirecte bewijzen, zoals inmiddels verloren gegane briefwisselingen tussen de beide denkers (Ibid.). Pas nadat Koerbagh tijdens de verhoren met deze briefwisselingen werd geconfronteerd, gaf hij toe dat hij Spinoza kende en dat hij hem thuis had bezocht. Uit diezelfde verhoren blijkt duidelijk dat Koerbagh zijn vrienden en zijn broer Johannes in bescherming wilde nemen. Het lijkt mij in ieder geval zeer ongeloofwaardig dat Koerbagh niet met Spinoza over filosofische onderwerpen heeft gesproken, zeker gezien de duidelijke verwantschap van hun metafysica.

Een verdere complicatie is nog dat Koerbagh ook door andere filosofen en collegianten beïnvloed is. Israel stelt, in mijn optiek terecht, dat niet

alleen Spinoza, maar ook onder meer Van den Enden, Meijer, Hobbes, Van Berkel (1629 – 1686) en Bouwmeester het werk van Koerbagh in belangrijke mate hebben beïnvloed (Israel, 2001, p. 193). Op grond van de ideeën in *Licht* die contrasteren met het werk van Spinoza, lijkt Koerbagh met regelmaat meer geïnspireerd te zijn door Meijer, Hobbes en Descartes dan door Spinoza. Ook Meijer meende dat de theologie ondergeschikt was aan de filosofie; dat het de filosofen zijn die de Bijbel zouden moeten interpreteren en dat men de cartesiaanse analyse los zou moeten laten op de Heilige Schrift om het Woord van God te achterhalen (Leeuwenburgh, 2013, p. 157). Ook deelde Koerbagh met Meijer het taalpurisme, de idealen van de verlichting van de massa, de popularisering van wetenschap en het intellectuele belang van de Nederlandse taal (Leeuwenburgh, 2016, p. 6).

Klever stelt vast dat Meijer ook een belangrijke invloed heeft gehad op Spinoza (Klever, 1997, pp. 80-81). Hij noemt Spinoza een volgeling van Meijer wanneer het gaat om de grote waarde die aan het verstand wordt gehecht om de werkelijke aard der dingen te kennen. Ook vermoedt hij dat Meijer een belangrijke inspiratiebron vormde voor Spinoza's *Tractatus Theologico-Politicus*. Spinoza staat echter duidelijk niet in de traditie van Meijer, wanneer het gaat om de lexicografie, het massaverlichtingsideaal en zijn taalpurisme.

Het grote vertrouwen in de menselijke rede dat in het werk van Koerbagh centraal staat, is ook aanwezig in het werk van Spinoza. Dit had Koerbagh, die goed bekend was met het werk van Descartes, echter ook aan deze laatste kunnen ontleenen. Weliswaar menen zowel Spinoza als Koerbagh dat ethiek en kennis met elkaar samenvallen, maar het is moeilijk aan te geven of Spinoza op dit gebied Koerbagh heeft beïnvloed, of dat Koerbagh deze opvatting op een andere wijze of onder invloed van een andere filosofische stroming heeft gevormd, zoals de Stoa. Ook de Stoïcijnen meenden namelijk dat rationaliteit, kennis en geluk sterk samenhangen (Irwin, 1989, pp. 173-174).

Het blijft voornamelijk dus een lastige opgave het spinozisme, meijerianisme, cartesianisme en hobbesianisme in het werk van Koerbagh precies vast te stellen. Het wordt wel duidelijk dat Spinoza zeker niet de enige denker is die Koerbagh heeft beïnvloed. Dat Koerbagh een volgeling van Spinoza zou zijn en Spinoza dus zijn leermeester, lijkt mij ongefundeerd.

5 Conclusie

Op basis van de vergelijking is geconcludeerd dat de overeenkomsten tussen *Licht* en de *Ethica* minder relevant zijn voor de uiteindelijke filosofische posities, dat zij vaak aanleiding geven om toch van een verschil te spreken en dat er wezenlijke contrasten tussen de beide werken aanwezig zijn.

Uiteindelijk is het de metafysica van Spinoza, in het bijzonder zijn monistische substantieleer, waardoor Koerbagh geïnspireerd moet zijn geraakt: aan wie anders had hij dit idee immers kunnen ontleenen (Leeuwenburgh, pp. 158-159)? Maar dat hij in *Licht* geïnspireerd is geraakt door een metafysisch idee van Spinoza, maakt van Koerbagh nog geen spinozist. Koerbagh huldigt ook originele filosofische opvattingen, die hij niet aan Spinoza heeft ontleend. Hij is hierbij sterk geïnspireerd door Hobbes, Descartes, Meijer en andere collegianten.

De originaliteit van Koerbagh blijkt uit zijn taalpurisme. Ook staat hij in een lexicografische traditie, schreef hij in het Nederlands in plaats van het Latijn en verzette hij zich tegen de ontoegankelijkheid van de academische wereld. Hij zette zich daarmee ook in voor de autonomie en zelfstandigheid van de Nederlandse Republiek. Daarnaast was hij een rationeel-etymologische godsdienstcriticus: hij meende vanuit scherpe en zeer idiosyncratische taalkundige analyses dat de bevolking door geestelijken werd bedrogen en misleid. Dit terwijl Spinoza eerder een klassieke en ook elite filosoof is: het denken is bij hem meer een doel op zich (Ibid., p. 132). De opvattingen in *Licht* zijn eigenzinnig, bijvoorbeeld wanneer het gaat om de idee dat de Schrift voor verbetering vatbaar is. Koerbagh wilde godsdienstige conflicten opheffen door te pleiten voor een universeel geldende en op het verstand gestoelde godsdienst. Hij identificeerde zich met Jezus: iemand die hij beschouwde als een wijs mens die onwetendheid wegneemt, waarheid brengt en de mens aanzet tot zelfstandig denken.

Koerbagh zette zijn spinozistisch getinte metafysica niet op een ethische, maar eerder op een politieke, hervormende en verlichtende manier in. Hij was een intellectueel, die met zijn denken het politieke en religieuze landschap wilde veranderen, in het licht van welk doel hij het metafysische idee van Spinoza in de basis had ingezet, maar hier tegelijkertijd

naar eigen inzicht en opvattingen invulling aan heeft gegeven. Hij was een geloofscriticaster en verlichter die om filosofische en politieke redenen in het Nederlands schreef. Dit terwijl Spinoza meende dat verlichting slechts voor een zeldzame groep mensen was weggelegd. Zijn denken ging schuil achter een complex wiskundig opgebouwd systeem.

We moeten van Koerbagh daarom beslist geen verkapte of zelfs amateuristische spinozist maken en van zijn werk geen voorbode van de *Tractatus Theologico-Politicus* of de *Ethica*. Koerbagh was een maatschappelijk, politiek en religieus hervormer: een moedige denker bij wie het verlichtingsideaal in vrijwel ongelimiteerde vorm tot uitdrukking komt. Hij was een martelaar van het vrije woord.

Dankbetuigingen

Ik wil in de eerste plaats prof. dr. Wiep van Bunge bedanken voor het nomineren van mijn essay voor publicatie in het ESJP. Ook bedank ik de redactieleden hartelijk voor hun nuttige commentaar.

Stephan Rijsdijk (1991) studeerde Fiscaal Recht (LLM) en Filosofie van een bepaald wetenschapsgebied (MA) aan de Erasmus Universiteit Rotterdam. Hij is op dit moment werkzaam als fiscalist. Zijn academische interesses gaan uit naar de filosofie van de Verlichting en de ethiek van het belastingrecht.

Noten

1. In het vervolg van dit essay zal naar Adriaan Koerbagh worden gerefereerd wanneer uitsluitend de achternaam “Koerbagh” genoemd is. Wanneer op zijn broer Johannes wordt gedoeld, dan zal dit expliciet zijn aangegeven.
2. De volledige titel van dit werk luidt: “Een licht dat schijnt in duistere plaatsen. *Een verheldering van de voornaamste kwesties van theologie en godsdienst.*”
3. Hier sluit ik uiteraard niet mee uit dat een filosofische vergelijking tussen Spinoza’s *Tractatus Theologico-Politicus* en *Licht* of andere werken van Koerbagh een vruchtbaar onderwerp kan vormen voor toekomstig onderzoek naar de verhouding tussen de beide filosofen.
4. In komende verwijzingen zal uit deze Nederlandse vertaling van de *Ethica* worden geciteerd in de (beknopte) stijl die in secundaire literatuur met betrekking tot dit werk gebruikelijk is.
5. In de literatuur bestaan overigens meerdere discussies omtrent het begrip “attribuut”, bijvoorbeeld wanneer het gaat om subjectivistische of objectivistische interpretaties ervan.
6. Dit ideaal van massaverlichting wordt door Leeuwenburgh terecht als cruciale factor genoemd om te verklaren waarom Koerbagh zo streng is bestraft, ondanks het feit dat hij beschikte over invloedrijke vrienden en leefde in de relatief tolerante Republiek.
7. Voor geïnteresseerden in het taalpurisme van Koerbagh verwijs ik naar het volgende artikel dat nog in Brill gepubliceerd zal worden: Leeuwenburgh, B. *Koerbagh: Dutch as a weapon in politics*.
8. De gedachte dat Koerbagh niet alleen begrippen afbreekt, maar deze ook op een constructieve manier herinterpreteert, wordt in het volgende artikel nader geanalyseerd: Wielema, M. 2001. *The Two Faces of Adriaan Koerbagh*. In: *Geschiedenis van de Wijsbegeerte in Nederland* 12. Budel: Uitgeverij DAMON bv, p. 57-75.
9. Zoals eerder vermeld, zal ik mij alleen richten op een vergelijking van Spinoza’s *Ethica* met *Licht* van Koerbagh. Bij wijze van uitzondering zal in paragraaf 4.3.3. wel kort de *Tractatus Theologico-Politicus* in een vergelijking worden betrokken, namelijk wanneer het gaat om de verhouding tussen de theologie en de filosofie bij Spinoza en Koerbagh. Dit wil niet zeggen dat er niet ook bijvoorbeeld op grond van de genoemde *Tractatus Theologico-Politicus* of *Bloembhof* van Koerbagh andere belangrijke verschillen of overeenkomsten naar voren kunnen komen. Een voorbeeld van een dergelijke overeenkomst is bijvoorbeeld de volgende. Koerbagh meent dat de politieke macht voor zowel politieke als religieuze wetten niet toebehoort aan geestelijken, maar aan de overheid. Spinoza huldigt ook een dergelijke opvatting in de *Tractatus Theologico-Politicus*. Hij meent dat geestelijken geen politieke macht zouden moeten hebben, en dat de soeverein volledige zeggenschap zou moeten hebben over alle kwesties, zowel van religieuze als politieke aard. Andere overeenkomsten of verschillen zullen hoogstwaarschijnlijk betrekking hebben op Spinoza’s interpretatie van de *Schrift*.

10. Als nuancering dient wel genoemd te worden dat ook Spinoza weliswaar bijgeloof bekritiseert, omdat dit voortkomt uit de verbeelding. De ware religie wordt bij hem ook op zuiver rationele gronden geformuleerd. Maar in de *Tractatus Theologico-Politicus* wordt de verbeelding niet geheel afgebroken, omdat deze op een positieve manier politiek aangewend wordt.

Literatuurlijst

- Irwin, T. (1989) *Classical Thought*. New York: Oxford University Press.
- Israel, J.I. (2001) *Radical Enlightenment: Philosophy and the Making of Modernity 1650-1750*. Oxford: Oxford University Press.
- Klever, W.N.A. (1997) *Mannen rond Spinoza (1650-1700). Presentatie van een emanciperende generatie*. Hilversum: Uitgeverij Verloren, p. 80-81.
- Koerbagh, A. (1668a) *Een licht dat schijnt in duistere plaatsen*. Vert. Wielema, M. (2014) Nijmegen: Uitgeverij Vantilt.
- Koerbagh, A. (1668b) *Bloembhof*. DBNL. URL: http://www.dbnl.org/tekst/koer001bloe01_01/.
- Leeuwenburgh, B. (2013) *Het noodlot van een ketter: Adriaan Koerbagh 1633-1669*. Nijmegen: Uitgeverij Vantilt.
- Leeuwenburgh, B. (2016) *Koerbagh: Dutch as a weapon in politics*. (nog te publiceren in Brill).
- Nadler, S. (2002) *A Companion to Early Modern Philosophy*. Oxford: Blackwell Publishing.
- Nadler, S. (2006) *Spinoza's Ethics: An Introduction*. Cambridge: Cambridge University Press.
- Nadler, S. (2011) *A Book Forged in Hell. Spinoza's Scandalous Treatise and the Birth of the Secular Age*. Princeton: Princeton University Press.
- Parry, R. (2014) Ancient Ethical Theory. In: *The Stanford Encyclopedia of Philosophy*. Zalta, E.N. URL: <http://plato.stanford.edu/archives/fall2014/entries/ethics-ancient>.
- Spinoza, B. de (1677) *Ethica*. Vert. Krop, H. (2015). Amsterdam: Promotheus/Bert Bakker.
- Wielema, M. (2001) The Two Faces of Adriaan Koerbagh. In: *Geschiedenis van de Wijsbegeerte in Nederland 12*. Budel: Uitgeverij DAMON bv, p. 57-75.

Wielema, M. (2003) Adriaan Koerbagh: Biblical Criticism and Enlightenment. In: Van Bunge, L. *The Early Enlightenment in the Dutch Republic, 1650-1750*. Leiden/Boston: Brill.

Wielema, M. (2011) *Adriaan Koerbagh, A Light Shining in Dark Places, to Illuminate the Main Questions of Theology and Religion*. Leiden/Boston: Brill.



Compatibilism and the Basic Argument

Lennart Ackermans

1 Introduction

In his book *Freedom Evolves* (2003) and article (Taylor & Dennett, 2001), Dennett constructs a compatibilist theory of free will and responsibility. It is based on a conception of possibility that is broader than the libertarian conception, allowing it to be compatible with determinism. Dennett does not give much attention to the incompatibilist's Basic Argument (also known as the Consequence Argument). This is to the dissatisfaction of Fischer (2003, 2005), a semicompatibilist who believes that while we have the kind of possibilities that are required for moral responsibility, the Basic Argument shows that the libertarian free will is still worth having—and incompatible with determinism.

Dennett (2005) disagrees, and what follows is a quarrel about the burden of proof. Dennett maintains that there are too many versions of the Basic Argument and he cannot address them all. Fischer replies by giving two versions of the Basic Argument, to which Dennett does reply. I argue, however, that Dennett does not give convincing rebuttals. I will have a closer look at the arguments and suggest alternative ways in which compatibilists such as Dennett can make their defense stronger. Important shall be my conception of the 'epistemic world', which I claim is the relevant conception of the world when investigating the issue of free will.

In section 2, I summarize the disagreement between Dennett and Fischer. I then, in section 3, propose ways to address both versions of the Basic Argument given by Fischer, which Dennett could use to make his position stronger.

2 The debate

Taylor & Dennett (2001) defend a traditional version of compatibilism. They develop notions of 'possibility' and 'causality' that are compatible with determinism. Incompatibilism, the authors contend, subsists on a confused notion of these concepts. When determinism is true, the compatibilist notion of possibility allows one to say that 'I could have done otherwise', whereas this statement is necessarily false when using the incompatibilist conception of possibility (which requires that given the exact same past, something else could have happened). The compatibilist conception of possibility, we are led to believe, is what underlies what is actually meant by statements about possibility, and the incompatibilist notion has no merit in using.

2.1 Possible Worlds

To define their concepts of possibility and power, the authors make use of the framework of possible worlds. A possible world is a hypothetical universe. For the purpose of this essay, it is sufficient to look at a possible world as if it is a function of spacetime $w: \mathbb{R}^4 \rightarrow \{0,1\}$, which assign to each quadruple $(x,y,z,t) \in \mathbb{R}^4$ the value 1 if and only if at time t matter occupies location (x,y,z) .

The statement 'I could have cycled to the supermarket (instead of walking)' can be described as follows: there exists a possible world, which up to the point of going to the supermarket is sufficiently similar to our world, in which I cycled to the supermarket.

There is some vagueness created by the use of the phrase 'sufficiently similar'. Taylor and Dennett choose to leave this as it is, because the idea is clear enough. The precise interpretation of 'sufficiently similar' will depend

on what the intended meaning of the phrase ‘I could have cycled to the supermarket’ is. If the person saying it does not have a bike, he clearly imagines the existence of a possible world that is less similar than if he does have a bike, and him deciding to cycle would have only required to think about the matter (‘Will I cycle or will I walk?’) a second longer. However, when addressing the Basic Argument (in particular, Fischer’s version of it), we will need a more specific definition, which I give in section 3.2.

The Possible Worlds Characterization of Possibility allows one to say that ‘I could have done otherwise’ while determinism is true. Taylor and Dennett show that this concept of possibility allows us to explain and understand the world, whereas the incompatibilist narrow conception is useless for this purpose. They raise the question why free will requires the narrow conception, and why we should be interested in such a concept at all. The burden of proof, they claim, rests with the incompatibilists, who have to show why ‘real’ possibility demands the narrow conception (Taylor & Dennett, 2001, pp. 231-235).

2.2 Van Inwagen’s Basic Argument

In an appendix of the same paper, Taylor & Dennett attack a version of the Basic Argument as given by Van Inwagen (1975). Van Inwagen’s Basic Argument is a *reductio ad absurdum* starting from the premise that an agent has the ‘power to cause’ at least one event, concluding that the agent has the power to cause no event whatsoever if the world is deterministic. Taylor and Dennett (2001, p. 236) formulate Van Inwagen’s argument as follows.

First, some definitions. For any event a , let $\sigma(a)$ be the proposition asserting that a occurs. Let σ_0 be a comprehensive description of the universe’s state at some time in the remote past, and let λ be a statement of the laws of nature. Let A be a human agent.

Without defining precisely what it means to have the ‘power to cause’ an event, premise 1 and 2 are assumed to hold. We start with four premises:

1. For events a and b , if A has the power to cause a and $\sigma(a) \Rightarrow \sigma(b)$ obtains in every possible world, then A has the power to cause b .
2. For any event a , if A has the power to cause a , then A has the power to cause the event that a does not occur (in which case $\sigma(a)$ is false).
3. There exists at least one actually occurring event x which A has the power the cause.
4. Determinism is true. Hence, in every possible world, $\lambda \wedge \sigma_0 \Rightarrow \sigma(x)$ is true. Equivalently, $\neg\sigma(x) \Rightarrow \neg(\lambda \wedge \sigma_0)$.

From this the following conclusions follows:

5. A has the power to cause an event whose occurrence implies $\neg\sigma(x)$, and therefore implies the falsity of either λ or σ_0 , which is absurd.
6. Conclusion 5 must be rejected, and premise 1 and 2 are plausibly true. Hence, either determinism is false or A does not have the power to cause any of his actions.

However, Taylor and Dennett think that, when using a proper definition of ‘power to cause’ premise 1 is false. They maintain that causation should be understood as necessity. That is, a causes b means that a is a necessary requirement for b to happen ($b \Rightarrow a$ holds in all possible worlds close to actuality). ‘Power to cause’, then, should be rendered as follows in their possible world terminology:

A has the **power to cause a** if and only if for some event \bar{a} describing an action of A and a world w that is sufficiently similar to the actual world, $\sigma(a) \wedge \sigma(\bar{a})$ holds in w and $\sigma(a) \Rightarrow \sigma(\bar{a})$ in every world similar to w .

When such a definition is used, premise 1 is obviously false. A having the power to cause a means that $\sigma(a) \Rightarrow \sigma(\bar{a})$ in some set of worlds similar to w . It does not follow from $\sigma(a) \Rightarrow \sigma(\bar{a})$ and $\sigma(a) \Rightarrow \sigma(b)$ that $\sigma(b) \Rightarrow \sigma(\bar{a})$ — as we learn “in Logic 101,” Taylor and Dennett add.

2.3 Response by Fischer

In a 2005 issue of *Metaphilosophy*, Dennett and Fischer debate Taylor and Dennett's analysis of the Basic Argument. Previously, in a review of Dennett's book *Freedom Evolves*, Fischer argued that the Basic Argument gives us strong reason to deny the existence of free will in a certain sense if determinism is true (Fischer, 2003), complaining that Dennett has "declined to directly confront" this argument (Fischer, 2005, p. 427). In the debate in *Metaphilosophy*, Dennett challenges Fischer to look at his discussion of Van Inwagen's version of it, and that is what Fischer does.

Fischer starts by saying that it would be surprising if many philosophers have been making logical errors — "the sort one would learn about in Logic 101" — without anyone noticing. Fischer agrees that Taylor and Dennett's reading of Van Inwagen's argument contains a logical error, but contends that when causality is interpreted in terms of sufficiency instead of necessity, the argument becomes unproblematic. Fischer agrees that this weaker concept of causality does not capture some "ordinary, common-sense idea," but contends that the weaker concept of causality makes the argument valid (and presumably, Fischer thinks it is still convincing), and that therefore, Taylor and Dennett's worries do not affect the argument.

Fischer goes on to say that there are other versions of the Basic Argument that also do not face the problem suggested by Taylor and Dennett. He presents his own version, which he suggests is immune to Taylor and Dennett's criticism because it does not contain premises such as in line 1, and because it does not use phrases such as 'S can make it the case that some state of affairs obtains' (Fischer, 2005, p. 432). Fischer's version employs instead the *Principle of the Fixity of the Past and Laws*, which states:

An agent has it within his power to do *a* only if his doing *a* can be an extension of the *actual* past, holding the natural laws fixed. (Fischer, 2005, p. 432, my italics)

If this principle is true, the conclusion that free will does not exist if the world is deterministic follows readily from the following argument. Again, let σ_0 be some comprehensive description of the universe's state at some

time in the remote past, and let λ be a statement of the laws of nature. Let x be some action of *A*.

1. For an action *a*, *A* has it within his power to do *a* only if $\sigma_0 \wedge \lambda \wedge \sigma(a)$ is true in some possible world *w* (Principle of the Fixity of the Past and Laws).
2. Determinism is true. Hence, $\lambda \wedge \sigma_0 \Rightarrow \sigma(x)$ is true in every possible world. Equivalently, $\neg\sigma(x) \Rightarrow \neg(\lambda \wedge \sigma_0)$.
3. *A* did not have it within his power to not do *x* (as that would make the past or laws of nature false).

The Fixity Principle captures our intuition, according to Fischer, that our freedom is to "add to the given past, holding the natural laws fixed":

When we think of ourselves as being free to pursue a particular path into the future, we explicitly or implicitly think of that path as an extension of the actual past (holding the laws of nature fixed). We typically do not think of ourselves as having the power so to act that the past would have been different from the way it actually was, or so to act that the laws of nature would have been different. So, the picture is this: the various accessible pathways into the future branch off a fixed past (and contain the actual natural laws), and the reasons relevant to our practical reasoning are reasons that obtain only along the accessible pathways. A benefit that exists along some path is pertinent to my practical reasoning only if I can "get there from here." (Fischer, 2003, p. 635)

Denying this picture, according to Fischer, would be to substitute it for a "radically unintuitive picture," because it means we need to view the different pathways into the future as having different pasts. Since this is not how we actually view the world when we think of free will, the Basic Argument convinces us that free will as we commonly perceive cannot exist when determinism is true.

2.4 Dennett's response

Dennett is unimpressed. In his response to Fischer's defense of Van Inwagen's argument, he says:

[Fischer] agrees that given our notion of causation, the Basic Argument falls apart, but he proposes his own ‘theoretical’ notion of causation, acknowledges that it does not have even the imprimatur of ordinary language but does have the virtue of saving the Basic Argument for another day, or maybe another millennium. We wouldn’t want to say farewell to something as much fun as the Basic Argument in an appendix would we? Well, yes. (Dennett, 2005, p. 453)

Dennett (2005) argues that it is an impossible task to dispose of each and every version of the Basic Argument. Therefore, he chose to dispose of only the version that employs the concept of causality that he argues is correct. This is enough, he claims, until someone shows that he has left out something important. The burden of proof, says Dennett, rests with the incompatibilists.

He does reply to Fischer’s own version of the Basic Argument: he rejects the Fixity Principle, which “insists on taking the ‘actual past’ as ‘given’, down to the last electron, but it simply neglects the fact that this is an unmotivated insistence” (Dennett, 2005, p. 454).

Both authors agree that there is a definable concept of libertarian free will, based on the kind of power that the Fixity Principle talks about, which is incompatible with free will. However, Fischer takes “seriously the libertarian’s wish to have such freedom” (Fischer, 2005, p. 634). The difference with Dennett is that he does not care about the existence of such a libertarian free will at all. It seems that Dennett views it as similar to a free will arising from a ‘narrow conception’ of possibility, but it is unclear why exactly Dennett is so opposed to the Fixity Principle.

One may conclude from this that no side is right or wrong: both believe in their own theory of free will, both of which are internally consistent; proving one of them wrong may not be possible. Nevertheless, this is what Fischer seems to be trying: his version of the Basic Argument, he claims, proves Dennett wrong, *even if* he abides with his compatibilist conception of free will: the intuition underlying the argument is universal. The concept of free will seems to capture something we find *important*, and it is this value that is at stake. According to Fischer, it can only be found in libertarian free will. Hence, Dennett’s strategy to place

the burden of proof on the incompatibilist’s side does not work: Fischer’s argument already tries to disprove Dennett. To defend his position, Dennett needs to address it directly. This is what I will try to do next.

3 A defense of compatibilism

3.1 Van Inwagen’s Basic Argument

Fischer contends that Van Inwagen’s argument should be seen as using a sufficiency definition of ‘power to cause’, rather than a necessity definition. To see if this argument poses a threat to Dennett’s compatibilist theory, I will incorporate Fischer’s suggestion in Dennett’s ‘possible world’ version of ‘power to cause’:

Definition 1. *A has the power to cause a if and only if for some action b of A and a world w that is sufficiently similar to the actual world, $\sigma(b)$ holds in w and $\sigma(b) \Rightarrow \sigma(a)$ holds in every world similar to w .*

When this definition of power to cause (with causality as sufficiency) is used, Van Inwagen’s argument is indeed valid. Fischer agrees, however, that this notion of causality has some problems. For example, using this notion, and assuming that the sun is shining at 12:00 today, any action that I perform ‘causes’ the sun to shine at 12:00 today (except, perhaps, the action of using a cloud generator to stop the sun from shining). Despite this unintuitive concept of causality, Fischer maintains that the conclusion of the argument is still valid.

To see whether this is true, we need to observe exactly what the conclusion of the argument (line 5 in section 2.2) means when Fischer’s definition of ‘power to cause’ is used. Substituting, we get

5. For some action b of A and a world w close to actuality, $\sigma(b)$ holds in w and $\sigma(b)$ implies the falsity of either λ or σ_0 , which is absurd.

Now the question is, *is it really absurd*—and I see no reason why it is. Assuming that λ holds in every possible world, it should be concluded that there is some set of possible worlds in which $\sigma(b) \Rightarrow \neg\sigma_0$ holds, with at least one world (w) in which $\sigma(b) \wedge \neg\sigma_0$ holds. It is clear that this does not

mean that A has the power to change the past: we are talking about possible worlds that do not actually exist. In the actual world, we have $\sigma_0 \wedge \sigma(b)$ by assumption.

So it seems that, also with a sufficiency view of causality, Van Inwagen's argument will not be a problem for a compatibilist free will that is based on a Possible Worlds Characterization of Possibility. There may still be versions of the argument that have force (e.g. versions that do not use possible worlds), but Dennett could sidestep them by saying that they use the wrong conception of possibility.

3.2 Redefining the Fixity Principle

3.2.1 Introduction

Dennett chooses to ignore Fischer's version of the Basic Argument because he views it as "an instance of the conflation we were exposing," (Dennett, 2005, p. 454) by which he means conflation of causation and possibility. This is not very convincing, as he does not show in what way Fischer conflates the two. As long as Dennett does not show in what way Fischer's argument can be interpreted as an argument that conflates causality and possibility, *and* that faces similar problems as the Van Inwagen argument, the Fixity Principle remains standing.

I propose to address the Fixity Principle in a different way. It seems that the principle *does* capture how we think of 'having the ability to choose among different possibilities.' The mere fact that we universally think about our will in this way (even Dennett, I suspect) seems to be a reason that it is important to us whether the Fixity Principle is valid or not.

However, we can understand the Fixity Principle in a different way. Fischer understands the *past* in the Fixity Principle to refer to the (one and only) metaphysical reality. But there is no reason to suppose that this is what we mean when we refer to the past. In fact, I believe that we do not: when we talk about reality, we talk about models of reality—our limited knowledge of reality, which corresponds in some way to the metaphysical reality but is not the same thing. If the past in the Fixity Principle refers to our *understanding of* reality, the Fixity Principle is true and we have free will, even if metaphysical determinism is true. To show why this is so, I will give a formalization of 'epistemic worlds', which helps us understand

how our knowledge of reality relates to the actual reality without conflicting with it. I will then discuss a different understanding of the Fixity Principle, referring to the epistemic past, that I claim captures the same intuitions as Fischer's Fixity Principle, but in a better way.

3.3.2 The Epistemic Past

When we think about the past or the world, we do not have a full description of the world available to us. In possible world terminology, if W is the actual world, we do not know the value of $W(x,y,z,t)$ for all x,y,z and t . In fact, we do not know any value that W takes for certain. The question arises: what kind of thing is it that we *do* know about the world? I suggest that our knowledge of the world consists of a set of propositions describing the world as it is at particular points in time. Such as the proposition: 'On January 1 at 12:00, the sun is shining.'

Definition 2. The **set of knowledge about the world K_t of A** is the set of propositions about the world that an agent A would believe to be true at time t .

For simplicity, I will assume that K_t is the same for all people and that K_t is closed under logical deduction. In addition, I will assume that there are no errors: all $p \in K_t$ are true in the actual world, and the belief in them is justifiable.¹ Since beliefs do not necessarily come in the form of propositions, I use the phrase 'would belief' and assume that there is a one-to-one correspondence between K_t and A 's (justifiably true) beliefs.

Let Ω be the set of all possible worlds in which the laws of nature hold. Given determinism, there will be some world $W \in \Omega$ that is the actual world. But importantly, *we do not and will never know* the exact description of this actual world. However, our knowledge about the world K_t relates in some way to the actual world W , and I will try to formulate this relation as follows. Since our knowledge about the world is never complete, there will be multiple possible worlds $w \in \Omega$ in which all propositions $p \in K_t$ hold.

Definition 3. Call the set of possible worlds in which all $p \in K_t$ hold the **epistemic world E_{K_t}** .

By the assumption that all $p \in K_t$ are true, we have $W \in E_{K_t}$.

For example, suppose that $q \in K_t$ is a proposition stating that yesterday it was sunny. Suppose further that there is no $p \in K_t$ describing the weather as it was the day before yesterday (we forgot what weather was the day before yesterday). Then E_{K_t} contains only possible worlds in which yesterday it was sunny, but possible worlds in which it was sunny as well as rainy the day before yesterday.

I now turn to a situation in which someone's knowledge about the world is altered because he makes a choice. Let a be an action that may or may not be performed by an agent A . Suppose that slightly after some point in time t an agent A chooses to do a . This choice will add the proposition $\sigma(a)$ to the set of propositions about the world K_t . In addition, the inclusion of $\sigma(a)$ in K_t may imply that one has to add more propositions to one's knowledge set. For example, suppose that A 's choosing to do a (say, taking a job) could have only happened if some statement q (say, that A is raised by his parents in a certain way) about the past is true. That is, $\sigma(a) \Rightarrow q$, and A knows that $\sigma(a) \Rightarrow q$, but before choosing to do a , A does not know whether q is true. Once A has chosen to do a it is implied that q is true, and his epistemic world is changed to include q .

Definition 4. For an action a , the **deductive closure of K_t and a** , denoted $K_t(a)$, contains K_t , $\sigma(a)$ and the set of all inferences such as q in the example I gave.

Using these definitions we can formulate a condition for free will that is similar to Fischer's Fixity Principle, which I call the Epistemic Fixity Principle:

An agent has it within his power to do a after time t only if $E_{K_t(a)} \neq \emptyset$, and we have $E_{K_t(a)} \subset E_{K_t}$ (Epistemic Fixity Principle).

The first condition, $E_{K_t(a)} \neq \emptyset$, is equivalent to there being at least one possible world $w \in E_{K_t}$ in which $\sigma(a)$ is true; that is, it is possible to do a given A 's knowledge of the world. The second condition, $E_{K_t(a)} \subset E_{K_t}$, states that all possible worlds that after having done a are compatible with

our knowledge of the world, should have also been compatible with our knowledge of the world before we made the decision to do a . In a sense, then, the post-decision epistemic world should be a 'continuation' of the pre-decision epistemic world; it is an epistemic version of the fixity of the past. This is intuitively plausible: when we make a decision, we want our perception of the world after making a decision to be consistent with the world before we made the decision.

It follows trivially that, given our assumptions, all choices we make satisfy the Epistemic Fixity Principle. For suppose that A does a slightly after time t . From the truth of all statements in K_t it follows that there is at least one possible world $w \in E_{K_t}$ in which $\sigma(a)$ is true; hence, $E_{K_t(a)} \neq \emptyset$. For the second condition, let $w \in E_{K_t(a)}$. Then (at least) all propositions $p \in K_t$ are true in w ; hence, we have $w \in E_{K_t}$.

It is also clear that Fischer's Basic Argument does not work when the Fixity Principle is replaced by the Epistemic Fixity Principle. For suppose that A is choosing between a and not doing a and ends up choosing to do a after time t . If his epistemic world at time t contains possible worlds in which $\sigma(a)$ is true as well as possible worlds in which $\neg\sigma(a)$ is true, then the Epistemic Fixity Principle is satisfied for both choices, as shown above.

3.2.3 The intuition behind the Fixity Principles

Fischer's Fixity Principle and Basic Argument have force because our intuitions tell us that the Fixity Principle is plausible. The conclusion that free will is incompatible with determinism follows logically after accepting the principle. However, intuitions are vague by necessity: they can be interpreted in different — mutually exclusive — ways. The Epistemic Fixity Principle captures *the same* intuitions as Fischer's Fixity Principle. To see this, let's have another look at Fischer's motivation for his Fixity Principle.

When we think of ourselves as being free to pursue a particular path into the future, we explicitly or implicitly think of that path as an extension of the actual past (holding the laws of nature fixed). We typically do not think of ourselves as having the power so to act that the past would have been different from the way it actually was, or so to act that the laws of nature would have been different. So, the picture is this: the

various accessible pathways into the future branch off a fixed past (and contain the actual natural laws), and the reasons relevant to our practical reasoning are reasons that obtain only along the accessible pathways. A benefit that exists along some path is pertinent to my practical reasoning only if I can “get there from here” (Fischer, 2003, p. 635).

Each of these sentences can be reinterpreted with reference to the epistemic past instead of the actual past. We do not think of ourselves having the power so to act that what we held true about the world before will be false after making a choice; the various accessible pathways into the future branch of a fixed epistemic past: when we make a choice, the set of propositions about the world increases and the set of compatible possible worlds narrows correspondingly — but the relevant worlds are retained (‘fixed’). The epistemic ‘pathway’ is a path of epistemic worlds that becomes smaller (containing less possible worlds) when choices are made and larger when information is forgotten.

The straightforward reply from the libertarian would be that our intuitions about free will refer to the *actual* world, not to our knowledge of the world. But I see no reason why they should. Consider discoveries in science, such as special relativity and quantum mechanics, that have ‘proven’ that the intuitive way we think about the world is incorrect — *if* the world we think about were supposed to *be* the actual world. However, it is clear these discoveries are not a problem for us. As long as our intuitive understanding of the world does not *conflict* with a scientific understanding of the actual world (that there is some meaningful correspondence), there are no problems, and it does not follow that we should abandon our intuitive understanding of the world; rather, we should acknowledge that the actual world and its intuitive representation are different, and that intuitions are about the latter. The same holds for determinism. As long as the actual world W is an element of every epistemic world E_k , there is a meaningful correspondence with the epistemic world and the actual world: there is no conflict.

The notion of the actual world is relevant in a different way. In all of human history, the world has been what our senses told it to be. It did not take long to discover, however, that the phenomena that we observed behaved according to laws, and that knowledge of these laws is beneficial.

This, combined with the observation that the world seems to be the same for other humans (other humans perceive similar things, described by the same laws, with their senses), makes it beneficial to hypothesize the existence of an *actual* world consistent with these laws. This notion of the actual world is clearly useful, but why should its relevance extend to free will? It is rather our *model* of the actual world that seems to me to be relevant to free will. In fact, what *is* the actual world is never of concern, by definition: we cannot get any closer to reality than what is epistemically possible.

When we look at it from the epistemic perspective, it seems that our intuitions about free will make a lot more sense. Making a choice requires that we do not *know* what we will choose. If we would know what we will choose, there would be no reason to deliberate; even more, it could hardly be called a choice. Hence, it is an obvious requirement that multiple epistemic options are open to us—and this could be the reason that we have the intuitions that underlie incompatibilist ideas.

Another objection could be along the lines of the following. If we find out that the actual world is deterministic, our knowledge of the world K_i will contain the proposition that the world is deterministic. This seems to conflict with the epistemic world E_{K_i} not being deterministic and containing multiple possible worlds. But no such conflict exists. In the suggested scenario, E_{K_i} will only contain possible worlds that are deterministic, so it is likely that it contains less possible worlds than before we found out that the actual world is deterministic. However, as long as we do not have full knowledge about the world, multiple (deterministic) possible worlds will be elements of the epistemic world. Hence, the epistemic world is not deterministic.

Conclusion

Dennett’s strategy of denying the burden of proof did not seem to work, as his position was attacked directly by arguments from Fischer, so I tried to make his position stronger by addressing these arguments.

The concept of the Epistemic World can be seen as an extension of Dennett’s compatibilist concept of free will. Dennett explains his conception of possibility using the Possible Worlds Characterization of Possibility: something is *possible* if there exists a possible world describing it within

some set of possible worlds X . But this characterization of possibility leaves out from which set we are choosing (what counts as ' X '). By defining this in epistemic terms, we can build a more powerful fortification against incompatibilist arguments.

I have argued that the intuitions underlying libertarian free will, most notably the Fixity Principle, should concern us—that is, we should want them to be satisfied. However, these intuitions should be interpreted differently than libertarians do. They do not concern the 'actual' world but the world that we actually deal with: the epistemic world. If our conception of free will deals with the epistemic world, I have shown, Fischer's version of the Basic Argument does not work. Hence, free will can coexist with determinism of the actual world.

This is not a knockout argument against incompatibilists. They could maintain that our intuitions are in fact about the actual world, and it will be hard, if not impossible, to give a decisive argument for either position. Rather, my arguments put the compatibilist position back on par with the incompatibilist position. The epistemic free will gives compatibilists a defense of free will that is resistant against incompatibilist concerns.

Lennart Ackermans (1992) has a bachelor in Mathematics from Leiden University, with a minor in Philosophy and additional courses in Computer Science. Interested in the monetary system, he also obtained a minor in Macroeconomics at the University of Amsterdam. He is now enrolled in the Research Master in Philosophy of Economics at the Erasmus Institute for Philosophy and Economics (EIPE). In his spare time, he sings in a cappella groups.

Notes

1. These assumptions will need to be dropped when we want to analyze what it means for free will when different people have different beliefs, or when an agent has false beliefs. In that case, one might want to drop the term 'knowledge' in favor of 'beliefs', but presently, the use of the term 'knowledge' creates no misunderstanding.

References

- Dennett, D. C. (2003). *Freedom evolves*. New York: Viking.
- Dennett, D. C. (2005). Natural Freedom. *Metaphilosophy*, 36, 449-459.
- Fischer, J. M. (2003). Freedom Evolves. *Journal of Philosophy*, 100, 632-637.
- Fischer, J. M. (2005). Dennett on the Basic Argument. *Metaphilosophy*, 36, 427-435.
- Taylor, C., & Dennett, D. (2001). Who's Afraid of Determinism? Rethinking Causes and Possibilities. In R. Kane (Ed.), *The Oxford Handbook of Free Will*. Oxford University Press.
- Van Inwagen, P. (1975). The Incompatibility of Free Will and Determinism. *Philosophical Studies*, 27, 185-199.



The Case of a Fear of Flying

An enquiry into the rationality of recalcitrant emotions

Thijs Heijmeskamp

1. Introduction

Your palms are sweaty, your mouth is dry; you feel dizzy as you breathe heavily. The muscles tense and your chest hurts. Many people undergo these ordeals when they merely enter a plane. For some people taking an airplane is as normal as taking the bus, for others airplanes are a somewhat uncomfortable mode of transportation and for some flying is one of the most frightening things in the world. It does not matter how many statistics on the safety of flying a person who is afraid of flying sees, the fear does not subside easily. Even though a person believes that flying is perfectly safe, that person can still be nailed to the ground with fear. The fear of flying is an example of a *recalcitrant emotion*. A recalcitrant emotion is an emotion that we experience despite a judgement that seems to conflict with it. In the case of a fear of flying it is the judgement 'flying is perfectly clear' that conflicts with the emotion of 'fearing for one's safety'. The fear of flying and other phobias are not the only cases of recalcitrant emotions. Recalcitrant emotions occur in ordinary experience and need not be emotions of fear.

In the case of the fear of flying there is a clear tension between the judgement and the bout of fear. Prima facie, there seems to be a rational conflict between the judgement and the emotion; that is the emotion seems to contradict the judgement. In the case of a fear of flying this rational conflict seems to be clear. However, not all cases of recalcitrant emotions are problematic. Such in the case of a rollercoaster ride or Mark Twain's Huckleberry Finn who saves his friend Jim, there is no clear rational conflict. This raises the question how and if an emotion can rationally relate to an emotion. In which sense can a recalcitrant emotion involve rational conflict?

The aim of this paper is to present an account of the rationality of recalcitrant emotions. I will argue that a recalcitrant emotion is not irrational because of a conflict with a belief, but because of a practical conflict with a person's reasoned goals.

This paper is divided into two parts. In the first part, I argue that recalcitrant emotions have been for the most part judged unwarrantedly negative and that not all recalcitrant emotions can be judged as irrational. This is, in my view, most apparent in the *(neo)judgementalist* theories of emotions, which emphasises the evaluative aspect of emotions. The issue of recalcitrant emotions has been widely discussed by authors working within this tradition. In his paper *The Irrationality of Recalcitrant Emotions*, Michael Brady (2009) gave an extensive overview of why the judgementalist and neojudgementalist theories failed to adequately account for recalcitrant emotions and tried to formulate an alternative account within the *(neo)judgementalist* tradition. According to Brady, an emotion is an *inclination to assent to an evaluation*. I argue that Brady does not solve the problem of the irrationality of recalcitrant emotions. By examining the *(neo)judgementalists* and Brady an interesting problem comes to light. Namely, by focussing on phobias and other negative emotions, recalcitrant emotions are often unnecessarily portrayed in a negative light, while a recalcitrant emotion can be a very constructive experience.

In the second part, I provide an account of how recalcitrant emotions can be irrational. Although I do not identify my own position as *(neo)judgementalist*, I use Brady's theory of emotion as a springboard for an alternative account of recalcitrant emotions. I argue that Brady unnecessarily separates the emotions from the content of an evaluation. By applying the insights given by Brady in a framework that takes the

embodiment of emotions as forefront, I argue that the irrationality of an emotion is dependent on the action the emotion is oriented towards and the evaluative aspect of the emotion. I apply this account to the case of a fear of flying. I conclude that if an emotion hinders us to perform actions or goals that we deem worthy of pursuing, an emotion is irrational. The rationality of an emotion is largely derived from the action the emotion is oriented towards.¹

2. Judgementalism, Neojudgementalism and the Recalcitrant Emotion

One influential cognitivist theory of emotion is judgementalism, which states that emotions are a kind of evaluative judgement, that is assertive propositional attitudes. Although forms of judgementalism can be traced back to the stoics, its first contemporary formulation is usually attributed to Robert Solomon (1976). In this view, when a subject judges that he or she is in danger, then this warrants the feeling of fear. In other words, you cannot be afraid of X, if you do not believe X to be dangerous. Since an emotion contains propositional attitudes, an emotion can be stated in terms of propositions: I am afraid of the bear, because I judge the bear as dangerous. If emotions are (or embody) judgements, then this means that emotions have cognitive content and are subject to the same rational commitments as (other) beliefs.

2.1 Judgementalism and the problem of recalcitrant emotions.

Recalcitrant emotions pose a challenge for judgementalist theories, because recalcitrant emotions seem to suggest that judgements are neither necessary nor sufficient for the existence of an emotion. If emotions are judgements, then one should expect the fear to disappear once a person has judged flying to be safe. Just like you cannot genuinely hold two contradictory beliefs (believe both A and not A at the same time), you also cannot hold an emotion that contradicts a belief. A judgementalist is forced to ascribe two contradictory judgements to an agent and therefore the agent is irrational when holding a recalcitrant emotion. Judgementalists have a hard time explaining this conflict between judgements and emotions. In the case of a recalcitrant emotion a subject does not consciously assent to a judgement, so the judgementalist must hold that the judgement in question is held unconsciously

(Brady, 2009, p. 414). This is problematic because it implies that a subject is confused: he holds an incoherent evaluative profile. This would mean that the conflict between a judgement and an emotion is unintelligible. However, conflicts between emotions and judgements occur frequently and are quite intelligible. Although you know you are perfectly safe, the fear persists. This experience is common to most of us. Patricia Greenspan (1988) has argued that the judgementalists need a reason for why they violate the *principle of logical charity* (p. 18). The principle of logical charity states that we can assume that an agent behaves quite rationally in general. Although logical incoherency is possible, our ascription of beliefs to an agent ought to be governed by a principle of logical charity in such a way that “we need some special reason [...] for attributing to him an unacknowledged judgment in conflict with those he acknowledges” (Ibid., 1988, p. 18). Attributing an unconscious judgement to an agent is not the only solution in explaining the conflict between an emotion and an already consciously held judgement (this conflict might not necessarily even be a rational conflict, but could just be a psychological conflict); there are other possible (and more likely) explanations. The only reason judgementalists provide for ascribing an unconscious judgement to an agent in the case of a recalcitrant emotion, is that it follows from adherence to their theory. In order to explain recalcitrant emotions, the judgementalists need a better reason for assigning unconscious judgements to agents than that it follows from theory or risk undermining their own explanation of recalcitrant emotions.

The case of the fear of flying seems to be a nail in the coffin for judgementalist theories of emotion, since I can be well aware of the safety of flying, while still being afraid. This implies that the emotion in question is not or does not contain a judgement that is in contradiction with the belief that flying is safe. Therefore, judgements seem neither necessary nor sufficient for the existence of emotions.

2.2 Neojudgementalism

In response to the problem of recalcitrant emotions, judgementalist theories of emotions seem to have been largely abandoned in favour of neojudgementalist models of emotion. In reaction to the problem that (recalcitrant) emotions do not have propositional attitudes, neojudgementalists deny that emotions are judgements. A ‘simple’ neojudgementalist holds that emotions

do not involve evaluative judgements, but *evaluative perceptions or feelings or construals or thoughts* (Brady, 2009, p. 415). In this view, the subject of an emotional experience construes or thinks of an object in an evaluative way, which falls short of a fully-fledged judgement. Thus, there is no longer the question of contradictory beliefs. The fact that I see flying in a plane as dangerous, or construe flying as dangerous does not entail that I *believe* flying is dangerous. In order to believe that flying is dangerous, I need to *assent* to the appearance.

According to this neojudgementalist account of emotions, there is no rational conflict between a recalcitrant emotion and a judgement. An agent cannot be accused of irrationality when an agent is merely construing a situation as thus-and-so, while not believing the situation is thus-and-so. If emotions are more akin to perceptions, it is not clear how they can come into conflict with judgements. A person seeing a straight stick bending in water is not irrational. If emotions are like perceptions, then the mind of a phobic is not irrational, just like the mind of a person subjected to an optical illusion is not irrational. This goes against the intuition “(...) that there is something wrong, from the standpoint of rationality, when fear persists in the face of a subject’s judgement that she is in little to no danger” (Ibid., 2009, pp. 413–414). For a neojudgementalist there are two possible responses: either reject this intuition and maintain that emotions are not subject to rational requirements or try to meet this objection.

2.3 Emotions as inclinations

Brady aims to formulate a different neojudgementalist theory that can meet the objection that neojudgementalist theories fail to account for the intuition that recalcitrant emotions involve rationality. According to Brady, both the judgementalist and the other neojudgementalist answers fall short: “Whereas judgementalism imputes too much irrationality to someone experiencing recalcitrant emotion, neojudgementalism fails to impute enough” (2009, p. 416). Brady wants to hold on to the rational conflict between emotions and judgement, without having an overly strong conception of rational conflict that the judgementalists suffer from.

Brady intends to maintain a close link between judgements and emotions, but construing emotions as judgements imputes too much rationality

upon them. The neojudgementalist construe emotions as some sort of evaluation that falls short of a judgement. As a consequence, emotions cannot be construed as (ir)rational in this neojudgementalist view. What the neojudgementalist mean with emotion as a sort of evaluation is that when we experience an emotion the import of the situation impresses or trusts itself upon us (Brady, 2009, p. 420). The meaning of an emotion immediately becomes clear to us;² when we feel fear, we believe that we are threatened. So when we see a bear and genuinely experience fear, we believe that the bear is a threat to us. We do not have to infer from the feeling of fear and the perception of a bear, that the bear is fearsome. However, Brady adds to this that, to say that the import of a situation impresses upon a subject is roughly the same as saying that a subject is *inclined* to assent to a certain view of the situation. An emotion is an *inclination* to assent to an evaluation (Ibid., 2009, pp. 420–421). Emotions themselves do not contain rational commitments, but focus and hold one’s attention on a particular object or aspect of a situation in order to enhance our evaluative construal. An emotion focusses and holds our attention; therefore, an emotion requires many of the cognitive resources to incline a subject to endorse an evaluative construal. In the case of recalcitrant emotions this focus is a waste of energy, because an agent has already made a judgement in this situation and a focus of attention is no longer needed. That is why an emotion is irrational when it is recalcitrant (Ibid. 2009, pp. 426–429).

It is important to note that the inclination to assent to an evaluation is in itself arational. Thus, Brady successfully averts the criticism of judgementalist theories of having to account for two contradictory judgements. An inclination to assent is no more or less rational than to feel tired. It is not the emotion itself that is irrational because it contradicts judgement, but it is the waste of cognitive motivational energy that makes the recalcitrant emotion irrational. Brady formulates the irrationality of recalcitrant emotions in a second way: the recalcitrant emotion is irrational because it involves the violation of a certain epistemic norm. Namely, a recalcitrant emotion inclines a subject to accept and act on an evaluation the subject already judged as false. Thus, a recalcitrant emotion is not only epistemically irrational for what it inclines a subject to believe, but also for how it inclines a subject to believe despite being aware that there is no good reason to accept this construal (Ibid., 2009, p. 428). A third reason Brady gives for the irrationality of recal-

citant emotions is that people invent reasons to support their emotional claim (Ibid., 2009, pp. 428–429). However, this third reason explains little, because not all people invent reasons in the case of recalcitrant emotions and it does not explain the conflict between emotions and judgements.

2.4 The negative attitude towards recalcitrant emotions

For Brady, a recalcitrant emotion is epistemically irrational. However, this does not solve the problem of the recalcitrant emotion, but merely shifts the problem. Now Brady must explain what exactly constitutes a waste of cognitive energy. But more importantly, he must explain why one should hold the norm that wasting one's cognitive energy is irrational. A recalcitrant emotion can be a fruitful instrument for new experiences. Take for instance rollercoaster rides, through which I can experience fear while having the belief that I am safe. It is not irrational for me to focus on my fear when riding a rollercoaster if I derive pleasure from it, as long as that fear is no basis for forming a belief. In fact, many of our cultural practises depend on similar discrepancy between emotions and judgements. For instance, when an actor plays the role of a suffering character, the audience can empathise but at the same time know that the actor is not really suffering. Hardly anybody is inclined to say that this is a case of a recalcitrant emotion. It remains unclear what warrants the epistemic norm. Thus, Brady's account of emotions as inclinations to assent to an evaluation falls short in accounting for the irrationality of recalcitrant emotions.

However, Brady's account of the irrationality of recalcitrant emotions shows a bias that is also present in judgementalist and many neojudgementalist accounts: conflicts between emotions and judgements must be avoided and when there is a conflict between an emotion and a judgement it is the emotion that has to lose out. There is the tendency in Brady and (neo)judgementalist theories of emotions to discuss examples of recalcitrant emotions where the emotions are in the wrong. The judgementalists speak of recalcitrant emotions as contradicting justified beliefs. Brady speaks of recalcitrant emotions as a waste of cognitive motivational energy. By focusing on negative emotions, judgements are unjustifiably favoured over emotions. The standard of favouring judgements over emotions leads to an overly negative appraisal of recalcitrant emotions.

A well-used example of favouring a recalcitrant emotion over its conflicting judgement is the case of Huckleberry Finn, which was first given as an example by Jonathan Bennett (1974). Huck helps his friend Jim to run away from his slave owners, but he has a bout of conscience while helping his friend. Huck decided to turn his friend Jim in, because he believed he had stolen property from others.³ However, Huck's belief that he did something wrong conflicts with his sympathy and natural feelings for his friend. Instead of turning Jim in, he lies to the slave hunters in order to save his friend. Huck has many reasons for giving his friend up and none for stealing. Huck sees his compassion for Jim as a weak, ignorant and wicked felony. As Bennett (1974) puts it “he simply *fails* to do what he believes to be right” (emphasis in the original) (p. 4). In this case it is not the emotion that has it wrong. In this case adjusting our emotional disposition to our judgement would be wrong. Furthermore, it is difficult to judge Huck's reconsideration of his judgements to put them more in line with his emotions as irrational. In fact, the recalcitrance of his sympathy was even necessary, from the standpoint of rationality for Huck to reconsider his judgements (Döring, 2014, p. 129).

Examples like Huckleberry Finn show that there are no *prima facie* reasons to favour judgements over emotions and to judge each case of a recalcitrant emotion as irrational. The difference between a fear of flying and enjoying a rollercoaster ride seems to be of *practical* nature and *contextual*, not rational. Moreover, recalcitrant emotions become problematic when they are unwarranted. This implies that the (neo)judgementalists were wrong in putting the conflict between judgement and emotion in the forefront of understanding recalcitrant emotions. When it comes to cases like phobias such as a fear of flying, there is a clear conflict.⁴ However, this conflict need not necessarily be a rational conflict. So how are we to characterise the conflict between the emotion and judgement in the case of unwarranted recalcitrant emotions?

3. A Sketch of a Theory of Emotion

So far, the analysis of (neo)judgementalist positions has left the emotion far from the realm of rationality. The problem of recalcitrant emotions seems not to be a case of an irrational problem, but more a problem of practical nature. So before I continue to address the rationality of recalcitrant emotions, I need to give an alternative account of emotions. In order to do this,

I address two specific mistakes in Brady's account that in my view stand in the way of adequately addressing the rationality of recalcitrant emotions. In section 3.1, I argue that Brady sees the role of emotions in evaluations as too limited. Furthermore, in section 3.2 I argue that Brady splits the emotions in mental and bodily aspects, thereby separating arousal, appraisal and action. By addressing these two problems, it becomes clear that an emotion does not contradict a belief, but an emotion can still be rationally assessed. The judgementalists emphasised the evaluative part of emotions, but made the mistake of construing emotional evaluations as a form of judgements. Although judgementalists had no problem explaining why an emotion could be rationally assessed, in the case of recalcitrant emotions a judgementalist was forced to impute too much irrationality to a subject. Brady's answer to the problems of judgementalism (and neojudgementalism) was to separate the emotion from the content of the evaluation. However, Brady was unsuccessful in showing how (recalcitrant) emotion can be rational or irrational, because the emotion had no influence on the content of its accompanying evaluation. In order to speak of the (ir)rationality of an emotion, we need a theory of emotion that relates the emotion to the content of an evaluative construal without the strong rational commitments of the judgementalists. Although the scope of this paper does not allow me to go beyond a mere sketch of a theory of emotion, it should be sufficient to further the investigation into the irrationality of recalcitrant emotions.

3.1 Why do we feel?

In Brady's view the role of emotions is to allow us to respond to important matters in the right way. They do this by alerting us to significant events or objects in our environment, thereby facilitating cognitive processing and enabling us to act appropriately with respect to these objects or events (Brady, 2009, p. 422). It is important to note that for Brady the evaluative construal is not the emotion nor is it created by the emotion. The emotion does not detect the salient objects or events, but the emotion captures our attention and enhances the representation of emotional objects (Ibid., 2009, p. 423). An emotion also prepares us for a specific behavioural response; for instance, the subject is being primed for fight-or-flight behaviour (Ibid., 2009, p. 425). In this way, Brady introduces a 'motion' into his theory of

emotion. Although Brady sees emotions as primarily passive, only reacting to stimuli and evaluative construals, they do involve a 'movement' in the form of an inclination to accept an evaluation or perform an action. However, the emotion does not have any influence on the content of the evaluative representation that is formed, besides the amount of attention that is given to the formation of the evaluative representation.

There is, however, good reason to question the limited role Brady gives to emotions in forming the content of an evaluation. Consider the following example given by Sartre (1984, p. 195): imagine yourself walking alongside a country road. Suddenly you are feeling hungry. You encounter a tree with a bunch of grapes; you extend your hand in order to reach for the grapes, however, the grapes are beyond your reach. Consequently, you shrug your shoulders and move on. In this example, our emotions not merely perceive salient properties of the world, but act upon the world itself (Ibid., 1984, p. 195). The grapes first appear as 'desirable' and 'ready to be picked'; however, the grapes are beyond reach so the urge for the grapes cannot be realised. This leads to an unbearable tension between the urge and the inability to reach for them. This tension is resolved by bestowing the new quality of 'too green' unto the grapes, thereby construing the grapes as unwanted. In the words of Sartre: "Only I cannot confer this quality on the grapes chemically. I cannot act upon the bunch in ordinary ways. So I seize upon this sourness of the too green grapes by acting disgusted. I magically confer upon the grapes the quality I desire" (1984, p. 195). Without actually modifying the structure of the object in question another quality is conferred upon it. In the case of the grapes one makes them less desirable. What occurs is a transformation of the world by our emotions, so that when all ways are barred we can still act. This transformation is not necessarily a conscious process. According to Sartre, when we experience an emotion "it is the body, which [...] changes its relations with the world in order that the world may change its qualities" (1984, p. 195). For Sartre, emotions are self-deceptive, a coping mechanism for a world we can no longer live in, strategies to avoid action and responsibility. However, Sartre ultimately took the wrong lesson.⁵ Whether or not emotions are deceptive, it is emotions that show us possibilities for action. It is emotions that move us towards actions and redirect us towards alternative actions when we cannot act on our original urges.

Brady makes the same mistake as other neojudgementalists in seeing the emotions as passive in the sense of only responding to objects or events in the world and not as actively creating meaning. Again, the evaluative construal or judgement is favoured over the emotion in that the emotion is only there to aid the evaluation. Inclining us to actions is an activity in itself. Brady fails to see that introducing a vector in the emotions and thereby making the emotions more dynamic is at odds with seeing the emotion as fundamentally passive. He focusses too much on the fact that emotions consume cognitive attention, and too little on the unbreakable link between the emotion and the action. Seeing the grapes as green is not reacting to a new salient property or object, but is changing the relationship between the actor and the grapes by construing the grapes as undesirable.

3.2 Embodied appraisal

The second problem is that although Brady introduces an important notion of seeing emotions as inclinations, he also seems to split the emotion in mental and bodily aspects, each fulfilling different roles. Mentally, the emotions gather cognitive mechanisms for an evaluative appraisal. Bodily, they prepare an agent to act upon that appraisal (think of an increase in heart rate and breathing when one is afraid in order to facilitate running away). However, we do not seem to experience these two elements as separate and neatly distinct from each other (Colombetti, 2007, p. 83). A separation between arousal, appraisal and action is maintained by emotion theories such as judgementalism and Brady's neojudgementalism, which is not found in experience.⁶ Without an intervention of cognitive evaluation the arousal remains meaningless for a subject. This means that on a personal level⁷ there is no distinction to be found between bodily and mental aspects when it comes to emotions. According to Giovanna Colombetti the "emotional aroused body is rather that *through* which the subject evaluates her world" (emphasis in the original) (2007, p. 544).

Not only on a personal level does the distinction between cognition and emotion (as separate psychological faculties implemented in separate neural areas) not hold. A review of neuroscientific data by Pessoa (2008) shows that brain regions that are traditionally seen as involved with emotion are also involved with cognition. The amygdala, which is highly associated

with fear and is given a central role to fear responses and recognition, is also of crucial importance for cognitive functions such as attention, associated learning and decision making. This is part of the reason why Brady and neojudgementalism (or perceptual accounts) fall short in giving an adequate account of emotions. They rely on a dualistic view of the brain to which emotion and cognition are neurally distinct. However, cognition and emotion overlap and are distributed over the whole brain. All emotions are complex dynamical patterns of brain-body events and no longer separable from cognition (Colombetti, 2014, pp. 98-112).

What does this entail for emotions? Emotions can no longer be seen as inclinations to assent to evaluative construals, since an emotion is in itself already meaningful. In Sartre's example of the grapes, the grapes first appear as desirable and thus as something we can eat. However, when you are unable to reach those grapes, the grapes are construed as disgusting and as objects we can (or rather should) pass by. We move from the possibility of eating the grapes to the possibility of simply walking past the grape. Emotions do not just intervene in deliberative processes, but emotions shape the manner in which things appear to us and structure our reasoning as a consequence (Ratcliffe, 2005, p. 187). Emotions do not incline us towards evaluations but towards possibilities of action. Contrary to Brady, I see emotions as intrinsically evaluative and motivational. The action a subject feels drawn towards is usually related to the object of the subject's emotion. This can take the form of an urge to slap somebody's face to a tendency to express a certain emotion. Thus, Brady's notion of emotion should be reformulated: an emotion is an inclination to assent to an action. Emotions move us in the world. They contain a clear 'motor component' that directly relates to the motivational aspect of emotions. Emotions help us to make sense of the world, which is a world in which we act. Therefore, there is no longer any difference between the arousal part and the appraisal part of the emotion.

4. The Intuition of Rationality

In the previous section, I argued that an emotion is part of the evaluation of a subject. Secondly, I argued that an emotion pushes us to a certain action. However, these two points do not immediately clarify whether an emotion is apt for rational evaluation. In this section I argue that emo-

tions cannot contradict beliefs, but emotions can still be judged irrational because of a potential practical conflict with our reasoned goals.

A recalcitrant emotion seems to be a mere psychological conflict between an experiential state and a judgement. However, this seems to be at odds with the widespread intuition that a recalcitrant emotion that conflicts with a judgement is irrational. In the words of Brady (2009): “For we have an intuitive sense that there is something wrong, from the standpoint of rationality, when fear persists in the face of a subject’s judgement that she is in little or no danger” (pp. 413-414). How can we reconcile this intuition with the notion of a recalcitrant emotion as a psychological conflict and not a contradiction? I argued that emotions should be seen as an awareness of the possibilities of action. It is important to note that emotions do not contain propositional attitudes;⁸ hence emotions themselves cannot be in contradiction with a belief. This does not mean that we can do much against experiencing a recalcitrant emotion (warranted or unwarranted), there is definitely a conflict, but not in the sense that a subject is contradicting itself.

Sabine Döring (2014) argues that there is a non-inferential relation between emotions and other mental states, like judgements or beliefs (pp. 130-131). Inferential relations hold between the propositional contents of judgements. The *modus ponens* is an example of a deductive inference rule. Döring uses an example given by Peter Goldie to show why inference rules do not apply to emotions. If you feel that someone’s behaviour is irritating, you can feel amused by ‘how over the top’ his irritating behaviour is. However, your amusement cannot be deductively inferred from the premise that certain behaviours that are considered to be ‘way over the top’ are humorous and the premise that the behaviour in question is way over the top. The behaviour is immediately seen as funny. According to Döring (2014), this example shows that the evaluative content of emotions is in some way open to reasons, but those reasons are never compelling and that the contents of emotions are neither propositions, nor just particular objects (pp. 130-131). She gives the example that when we are afraid of a gorilla, this fear is generally not a fear of a sick or a dead gorilla. We fear gorillas only in situations where they present a danger to us. An emotion can persist both logically and psychologically in light of ‘better’ judgement and knowledge. The affective content of emotions is different in kind from the mental content of beliefs

and judgements. The content of beliefs and judgements is strongly susceptible to evidence, in the sense that a subject cannot at the same time consciously and explicitly believe *p* and not-*p* (Ibid., 2014, p. 132). According to Döring (2014), we intuitively sense that recalcitrant emotions are irrational, because they lead to a practical conflict by motivating the subject to act in ways that interfere with the *reasoned pursuit of goals* (p. 125). Conflict arises when an emotion persists in spite of the subject’s better judgement.

The conflict between an emotion and a judgement need not be resolved, as the example of the rollercoaster ride demonstrated, where the recalcitrant emotion does not lead to a practical conflict. However, in the case of Huckleberry Finn there is a conflict that needs to be resolved. Huck believed he acted wrongly in helping his friend Jim. Thus, he reasoned that he should give his friend Jim up to the slave hunters. That is (perceived by Huck and the society as) the moral thing to do. However, Huckleberry’s goal and accompanying actions of giving Jim up conflicted with his feelings of friendship. The compassion he felt for his friend Jim prohibited Huck from accomplishing his goal. In this case, Huck gave up his reasoned goal in favour of acting on his emotions. If an emotion leads to a problematic action in relation to our goals, we can adjust our emotions or our goals,⁹ with their accompanying judgements and beliefs. This is something we do all the time; we can choose not to trust somebody because we have a ‘nagging feeling in the back of our heads’, despite there being no reason to distrust somebody. We can also choose to ignore this feeling and consider someone trustworthy. The feeling of distrust is incompatible with the belief that someone is trustworthy; we have to choose one over the other. We can be right or wrong, but it is difficult to consider someone irrational in both cases, even if the person in question is mistaken. When an emotion comes into conflict with a judgement, it becomes problematic if the emotion prohibits us from accomplishing our goals.

Döring (2014) draws the conclusion that this means that emotions cannot be irrational, only the actions that follow from the emotion (p. 135). However, I believe Döring is too quick in drawing this conclusion. According to her, emotions are like perceptions and therefore incapable of contradicting a judgement (Ibid., 2014, p. 135). She is right to emphasise the link between actions and emotions, but she does not consider the emotion as

part of the action in question. Remember that emotions show us possibilities for action and that they are relational experiences in the sense that an emotion indicates how an agent is inclined to respond to some object in the world. Neither is there a strict separation between cognition and actions, meaning that emotions cannot be like perceptions as envisioned by Döring. Contrary to Döring's account, emotions are part of the *evaluative profile* of our reasoned goals. As I argued in the previous section, emotions contribute to what we find meaningful, and therefore to what our personal goals are and what actions are appropriate to achieve those goals. In other words, our emotions cannot be easily separated from the actions. Emotional content displays a characteristic phenomenology while at the same time presenting the world as being a certain way (Slaby, 2007, p. 432). Thus, emotional content can be opposed to beliefs or judgements, without themselves being like judgements or beliefs. This also implies that the content and quality of an emotion are not separate or even separable. When the content of an emotion changes, there will be a change in how we feel about the situation and vice versa. When the content of an unwarranted emotion changes, the felt quality also changes; fear changes into relief. The *intentionality* and *phenomenality* of emotions are essentially united in emotional experience (Ibid., 2007, pp. 431-432). In the next section, I examine the link between irrational emotions and our reasoned goals. I argue that we still need to make a distinction between cases where the practical conflict of emotions and our goals can be solved and cases where this conflict is unsolvable. Only in the later case is a recalcitrant emotion irrational.

5. The Irrationality of Action

Being afraid in a rollercoaster is an important element of the enjoyment of a rollercoaster ride. The fear does not prohibit a person from attaining his goals, namely fun, but leads to excitement, which is a fun experience. That fear can lead to enjoyment is without question, since there are many examples of people purposefully seeking out fear.¹⁰ In other words being afraid is part of the game. Here a difference is to be found in comparison with a fear of flying. In both cases, there is a conflict between a judgement and an emotion (being afraid of a thing and judging the same thing as safe), but in the case of a rollercoaster ride being fearful perfectly aligns with what we set out to do, while in the case of flying being fearful clashes with safely travelling

from A to B. Being afraid in an airplane is deemed inappropriate in normal conditions. In the example of the fear of flying, a person is motivated to resolve the conflict between the emotion and the judgement, because the emotion inclines the subject to behave in a way that is contrary to the person's actions. In the example of the rollercoaster, a person is not motivated to resolve the conflict between the emotion and the judgement, because this conflict leads to enjoyment.

Whether an emotion is *warranted* or not is dependent on the goals a person has. However, this does not necessarily make an emotion irrational. Huckleberry's emotion interfered with his goals and judgements, so he adjusted one of them. There are also cases in which we can assert some control over our emotions. Emotions do not necessarily lead us to certain actions, at least not in a strict causal sense. Emotions show us a world of opportunities, i.e. practical situations that are significant for us. In the case of a recalcitrant emotion the recalcitrant emotion relates the subject with the world in a way that conflicts with the subject's reasoned goals. I speak of reasoned goals, because the goals must be construed as attainable and reasonable. In a society where nobody flies, a fear of flying is not a problem because flying is not considered as a possibility. Typically, we tend to act¹¹ on our emotions, so it is not difficult to see how an emotion leads us to behaviour we do not wish to have.

There appear to be two cases of an *unwarranted recalcitrant emotions*. In both cases an emotion interferes with the reasoned pursuit of goals. In the first case, the subject relates to the world in a manner that is incompatible with the subject's goals. This conflict is often solvable; we either adjust our emotional response or at the least withdraw our confidence of its content, or we can adjust our goals. This can be a constructive and fruitful experience. An example of this case would be Huckleberry Finn saving his friend Jim.

The second case, of which the fear of flying would be an example, is when a recalcitrant emotion not only becomes unwarranted but also problematic. In this case, the conflict between the reasoned pursuit of a goal and the interfering emotion is unsolvable.¹² Either a person is unable to change or discard his emotions, or a person is unable to change his goals. Emotional responses are notoriously difficult to change, especially when something is construed as dangerous even if we suspect that the emotion is

deceiving. Emotions can continue to move us even after we have discarded their content. Even changing one's goals can be impossible. Our goals are a fundamental part of our understanding of ourselves as rational agents. Our goals pertain the beliefs, judgements, desires and normative requirements a person commits himself to or society demands of him.

An emotion can be judged as inappropriate, while the emotion in question still moves a person towards a certain action. In light of a person's goals an action can be judged as rational or irrational. If you believe that in order to learn you have to go to school, then the goal of learning reasonably implies that you go to school. However, if you want to learn but you stay in bed and do nothing, your actions are irrational. The same holds when somebody is afraid of flying. If you want to visit friends who live far away, the safest, most comfortable and fastest way to travel is by airplane. However, during the flight, you act afraid and even terrified. Or in worse cases, you do not even make it on the plane. Reasonably you should behave in a comfortable way or at the very least unafraid when flying. However, we act in a different way than what is reasonable in light of our goals, hence our action in the plane can be called irrational. There is a mismatch between our rational goals and our emotional actions and evaluation, meaning that the irrationality of our emotions is derived from the irrationality of its accompanying action and evaluation. A recalcitrant emotion is not irrational because it exists, but because it disrupts the actions leading to the rational agent's goal *for no good reason*. The emotions as an evaluation and an inclination towards a certain action can undermine our evaluative profile of ourselves as rational agents. Therefore, an emotion can only be considered irrational in a weak sense. The recalcitrant emotions in question are irrational because they do not fit in the agent's specific construal of himself as a rational agent.

This of course raises the question of what it means to be rational when experiencing emotions. A full discussion goes beyond the aim of this paper and warrants a whole discussion in itself. What rational goals are or what rational behaviour is, is not a matter of individual reasoning, but reflects the normative expectations of the social group. What is rational is what is recognised as meaningful and therefore rational behaviour. Since the agent's goals and accompanying emotions have to meet the normative expectations of the group, the rationality of emotions gains an ethical dimension.

6. Conclusion

In the (neo)judgementalist tradition on recalcitrant emotions, there seems to be misdiagnoses of the problem. It is not the conflict between a judgement and an emotion that makes a recalcitrant emotion problematic, let alone irrational, because some conflicts between judgements and emotions are productive and often rational.

An emotion is irrational because the actions that follow from the emotion in question are irrational in relation to the agent's rational goal. The subject fails to construe the world in possibilities for action that correspond with the subject's goals and as a result the recalcitrant emotion moves (or inclines) the subject to act against his reasoned pursuit of goals. There is a mismatch between the emotional actions and the reasoned goals of the agent; this means that there is a mismatch between how a person behaves and how a person thinks he should behave given his desires and goals. To be more exact: a recalcitrant emotion is not necessarily irrational if it comes into conflict with a judgement nor is it necessarily problematic. When a recalcitrant emotion moves a subject towards an action that conflicts with the reasoned pursuit of his goals and the subject is unable to either adjust his emotional response or his goals, the emotion can be seen as leading to an irrational action. To fully understand a recalcitrant emotion, one must not focus on the belief the emotions seems to contradict but one must look at the agent as a whole, his rational goals and the social context the agent lives in.

Acknowledgements

I would like to thank professor Heleen Pott for helpful comments and suggestions on earlier versions of this paper. This paper improved significantly as a result of the suggestions of the editors of the Erasmus Student Journal of Philosophy.

Thijs Heijmeskamp (1988) is currently writing his master thesis on an embodied, situated account of moral agency. Earlier he finished his bachelor's degree in Philosophy at Erasmus University in Rotterdam. His main research interests are in the area of philosophy of mind, metaethics, moral psychology and philosophy of cognition.

Notes

1. This paper is not meant to fully discuss all possible ways in which an emotion can be considered irrational, but only the specific cases of unwarranted recalcitrant emotions.
2. Of course, there are cases of conflicting emotions or cases in which we are not sure why we are feeling a particular emotion. However, the meaning of an emotion is always immediate.
3. moral turning in a runaway slave sounds to us, Huck was convinced that turning Jim in to the slave hunters was the moral thing to do. Thus, he viewed his own actions of helping Jim and not betraying him to the slave hunters as immoral.
4. Here, I mean there is at least a clear psychological conflict within a subject who has a phobia.
5. I do not wish to imply that emotions never deceive us. Sartre is right in the sense that emotions can deceive us and emotions do often deceive us. A fear of flying is a good example where it is possible to argue that an emotion deceives a person. However, emotions do not deceive us systematically, but are generally productive.
6. This argument is not necessarily a knockdown argument for a view of embodied emotions and appraisal and the aim of this paper limits room for a complete argumentation for a theory of emotion. However, my sketch here should be sufficient to support my larger thesis. For more on the integration between affect and appraisal see Colombetti (2007).
7. Between the personal and sub-personal level is derived from Daniel Dennett. In short, the personal level is where we have experiences. The sub-personal level is the underlying physical processes.
8. Also in Brady's account emotions lack propositional content as in many other neojudgmentalist accounts. (See section 2).
9. I speak of goals in a broad sense. Goals are not only desired results or outcomes, but they also imply actions. Thus, goals pertain behaviour, attitudes, judgements, beliefs and desires.
10. What can be questioned is whether the fear that is sought out is the same as the fear experienced in unwarranted cases.
11. Action here is to be understood in a broad sense to include irrational and arational expressive actions.
12. At least not immediately solvable. For instance, a phobia can eventually be overcome through therapy, although this is not always the case. However, in the moment a person undergoes a phobic experience, there is little a person can do about it.

References

- Bennett, Jonathan (1974) The Conscience of Huckleberry Finn. *Philosophy* 49, 123–34.
- Brady, Michael, (2009) The Irrationality of Recalcitrant Emotions. *Philosophical Studies*, 145, 413-430
- Colombetti, Giovanna (2007) Enactive appraisal. *Phenom Cogn Sci*, 6, 527-546
- Colombetti, Giovanna (2014) *The Feeling Body*. Cambridge, MIT Press
- Döring, Sabine S. (2014) Why Recalcitrant Emotions Are Not Irrational. In: *Emotion and Value*. Ed. Roeser, S. & Todd C., Oxford, Oxford University Press
- Greenspan, Patricia (1988) *Emotions and Reasons: An Inquiry into Emotional Justification*. New York: Routledge, Chapman and Hall
- Helm, B. (2001). *Emotional reason*. Cambridge, Cambridge University Press.
- Pessoa, Luiz (2008) On the relationship between emotion and cognition. *Nature Reviews Neuroscience* 9, 148-158
- Ratcliffe, Matthew (2005) William James on Emotion and Intentionality. *Philosophical Studies*, 13, 179-202
- Roberts, Robert C. (1988) What an emotion is: A Sketch. *The Philosophical Review*, Vol 97, No. 2, 183-209
- Solomon, Robert (1976) *The Passions*. New York: Doubleday.
- Sartre, Jean Paul (1984) *The Emotions: A Sketch of a Theory*. Transl. B Frechtman, New York, Philosophical Library
- Slaby, Jan (2007) Affective intentionality and the feeling body. *Phen. Cogn. Sci.* 7, 429-444

