

Opvattingen van naturalisme maken het verschil

Habermas en Dennett over natuur, wetenschap en vrije wil

Norbert van Ettinger

Vorm : Bachelorscriptie
Scriptiebegeleider : Joel Anderson
Periode : Mei-Augustus 2007
Subfaculteit : Wijsbegeerte
Instelling : Universiteit Utrecht

[bpapb02.doc / Linschoten 24-09-2007]

Inhoudsopgave

1. Inleiding	3
2. Centrale thema's in de filosofie van Habermas	4
2.1. Verband tussen taalspel, taalgebruik, rationaliteit en perspectief op de wereld	5
2.2. Theoretisch filosofische onderbouwing van de formele pragmatiek	7
2.2.1. Afbeeldende en communicatieve functie van de taal	8
2.2.2. Pragmatisme	9
2.2.3. Transcendentale analyse en detranscendentalisering van mogelijkheidsvoorwaarden ..	10
2.3. Het onderscheid hard en zwak naturalisme	11
2.4. Proefrit maken met Habermas' metatheoretische verbinding	13
3. Centrale thema's in de filosofie van Dennett	16
3.1. Dennett's interpretationisme	17
3.1.1. Intentionaliteit	17
3.1.2. De drie verschillende houdingen van het Interpretationisme	18
3.1.3. Realisme en interpretationisme	19
3.1.4. Interpretationisme in relatie tot reductionisme en eliminativisme	20
3.1.5. Commentaar op Dennett's interpretationisme	20
3.2. Dennett's visie op bewustzijn en zijn strijd tegen het Cartesianisme	21
3.3. Dennett's kijk op naturalisme	22
4. Zwak naturalisme versus antireductionistische fysicalisme	23
4.1. Sterke en zwakke punten van interpretationisme van Dennett en Habermas	28
5. De kwestie van de vrije wil	31
5.1. Habermas over de kwestie van de vrije wil	31
5.1.1. Enige relativering met betrekking tot het artikel van de neurowetenschappers	32
5.1.2. Habermas beoordeelt naturalistische benaderingen van vrije wil	33
5.1.3. Commentaar op de beoordeling van Habermas	35
5.1.4. Habermas' zwak naturalistische benaderingen van vrije wil	36
5.2. Dennett over de kwestie van de vrije wil	38
5.2.1. Dennett over determinisme en vrije wil	39
5.2.2. Dennett over de evolutie van de vrije wil	41
6. Habermas versus Dennett m.b.t. de kwestie van de vrije wil	42
6.1. Determinisme en vrije wil	42
6.2. Mentale veroorzaking	44
6.3. Habermas' kritiek op memetiek en Universeel Darwinisme.	46
6.4. Morele verantwoordelijkheid	47
6.5. Over de noodzaak van het participierend perspectief!	54
7. Pragmatische beoordeling van antireductionistisch fysicalisme en zwak naturalisme	59
7.1. Biedt de theorie een goede mogelijkheid voor een beschrijving van normativiteit?	59
7.2. Kan men met de theorie een kritisch standpunt vertolken?	61
7.3. Implicatie theorie adoptie voor omgang met wetenschappelijke onderzoeksresultaten?	63
8. Conclusie	65
9. Literatuuropgave en voetnoten	71

1 Inleiding

Hoe kunnen we het beeld van onszelf als vrije morele actoren nog hoog houden als de moderne wetenschap tegelijkertijd toont dat wij niet, zoals sommige verlichtingsfilosofen meenden, de hoog op de troon van de schepping gezeten wezens zijn, die met hun rationeel reflectieve vermogens en vrije wil zich als minigoden kunnen onttrekken aan de ketenen van causaal wetmatig verloopende processen in de natuur?

Habermas wijst in zijn antwoord op de pretentie van de wetenschap om haar objectiverende kennis te verabsoluteren tot de enige correcte afspiegeling van de allesomvattende wereld.

In al zijn filosofisch werk wordt telkenmale bezwaar gemaakt tegen deze verleende epistemische oppermacht. Er is iets mis met het vertrouwen in dit perspectief wat ten onrechte op het voetstuk wordt gezet als het allesomvattende 'gezichtspunt van God'. Dit objectieve perspectief kan niet zonder het participantperspectief op het communicatief handelen.

Het door het participantperspectief gevormde beeld van onszelf als vrije, morele actoren, wier handelen wordt bewogen door rationeel en bewust gekozen redenen, moet op metafysisch, metatheoretisch niveau zonder reductie worden verbonden met het objectiverende perspectief op de causaal nomologisch verloopende processen van de natuur van de natuurwetenschap. Kortom, Habermas pleit ten behoeve van het behoud van het beeld van vrije door redenen bewogen agents voor een zwak naturalisme waarin de allesomvattende natuur meer behelst dan de natuur van de wetenschap.

De naturalist en antireductionistische physicalist Daniel Dennett daarentegen betoogt dat het beeld van vrije morele agent uitstekend verantwoord kan worden zonder concessies te doen aan het primaat van het 'objectiverende' perspectief van de natuurwetenschap.

In deze scriptie analyseer ik beide benaderingen met betrekking tot natuur, wetenschap en vrije wil. Ik beargumenteer de these dat Dennett beter in staat lijkt een stevig fundament te geven voor de constructie van een verantwoordelijk moreel actorschap dat sterk genoeg is om overeind te blijven bij (neuro)wetenschappelijke ontdekkingen. Hij constateert mijns inziens terecht dat morele instituties afgestemd moeten zijn op onze natuurlijke instincten. Het gevaar dat die instituties, systemen en conventies niet overeind blijven onder druk van wetenschappelijke ontdekkingen is groter als we deze baseren op een metafysisch en mythisch mensbeeld.

Ik laat echter ook zien dat Dennett, in tegenstelling tot Habermas, de aspecten van communicatie met de concrete-ander in relatie tot kennisopbouw en handelen onderwaardeert. Dennett beschouwt de mens te veel als eenling, wiens handelen is gericht op het eigenbelang. Habermas legt een te eenzijdig accent op het verstand. Hierdoor lijkt hij niet te onderkennen dat de mens zich beter herkent in een beeld met meer plaats voor het gevoel, dan in het rationele beeld van een slechts door de rede voortbewogen handelen. Beiden geven in hun theorieën te weinig aandacht aan de fundamentele rol van empathie in de moraal.

Voor een goed begrip van de opvattingen van Habermas en Dennett aangaande naturalisme en vrije wil, kun je denk ik niet heen om enkele centrale thema's in hun filosofie. Van Habermas bespreek ik daarom [in sectie 2] eerst zijn formele pragmatiek, en het door hem aangebrachte onderscheid tussen een zgn. hard en zwak naturalisme. Van Dennett [in sectie 3] behandel ik het concept van intentionele systemen (interpretationisme) en zijn naturalistische theorie over het bewustzijn. In een onderlinge vergelijking [in sectie 4] zal ik Dennett's

naturalisme positioneren binnen het door Habermas aangebrachte onderscheid tussen hard- en zwak naturalisme. Dan richt ik me [in sectie 5] op hun opvattingen betreffende de kwestie van de vrije wil en morele verantwoordelijkheid, en vergelijk hun posities onderling [in sectie 6]. Tot slot besteed ik [in sectie 7] aandacht aan hoe beide filosofieën pragmatisch omgaan met de relatie tussen het manifeste en wetenschappelijke mensbeeld. Ik zal dan ook bekijken in hoeverre beider naturalistische posities, die van Habermas en die van Dennett, voldoende stevige basis bieden voor een mensbeeld dat het fundamentele concept van verantwoordelijke morele actor kan waarborgen in een door natuurwetenschap en technologie beheerst tijdperk.

2 Centrale thema's in de filosofie van Habermas

De filosofie van Habermas¹ is een voortzetting van de kritisch theoretische traditie van de 'Frankfurter Schule'. Adorno en Horkheimer, de grondleggers van deze school, die zich met sociaal georiënteerd maatschappijonderzoek bezighielden, zagen het proces van verlichting, met haar emancipatoire aspiraties van autonomie voor het individu, ontaarden in een alles doordringende instrumentele rationaliteit. Maar volgens de minder moedeloze Habermas zagen Adorno en Horkheimer niet, dat het verlichtingsproces ook nog een rijkere emancipatoire opvatting van communicatieve rationaliteit bevat. Deze vorm van rationaliteit komt communicatief handelend tot stand, als mensen als participanten van een discours zich richten op consensus in een ideale spreek situatie. Omdat daarmee de vervormende invloeden op de interpersoonlijke relatie worden overwonnen, bereiken de participanten alsnog hun autonomie.

In zijn 'Theorie des kommunikativen Handelns (1981)' presenteert Habermas deze zgn. formele² pragmatiek. In de opbouw van deze theorie breekt Habermas uit de traditie van een voornamelijk retrospectief gerichte discussie, door hierin niet alleen de filosofieën van Kant en Hegel maar ook elementen van Angelsaksische filosofieën op te nemen. Hij verbindt het pragmatisme van Peirce en Dewey, en de taal filosofie van Wittgenstein, Austin, Grice en Searl met een Kantiaanse, transcendentale analyse en gaat op zoek naar de mogelijkheden voor kennis.

Conclusie van zijn analyse is dat het louter op zichzelf beschouwd objectiverende perspectief van de instrumentele rationaliteit een incompleet en vervormd beeld van de wereld constitueert. Voor een betere cognitieve ontsluiting van de wereld is de aanvulling met het complementair participerend perspectief van de communicatieve rationaliteit vereist. Deze visie werkt Habermas in zijn communicatietheorie nader uit.

Uitgangspunt is dat de termen waarheid, juistheid (morele waarheid) en waarachtigheid geen onwrikbare objectieve kennis impliceren, maar door consensus in talige praktijken worden gedefinieerd. In de taal is volgens Habermas een element van onvervalste, ongeforceerde consensus aanwezig. Iedere taalhandeling bevat een aanspraak op de waarheid, juistheid en waarachtigheid. Degene tot wie de taalhandeling is gericht kan deze aanspraak bekritisieren en er een geldigheidsaanspraak tegenover stellen. Als dit op een open en belangeloze manier gebeurt, en geen van de sprekers op strategisch handelen overschakelt, zal deze performatieve vorm van communicatie een ideale spreek situatie vormen en tot instemming van alle betrokkenen leiden. Het communicatief handelen onderscheidt zich daarmee van het instrumenteel of strategisch handelen, omdat het op de realisatie van consensus en niet op privé-doeleinden is gericht.

In zijn theorie van de moderne maatschappij legt Habermas zijn vinger op de zere plek. Het ‘objectiverende’ perspectief, en de daaraan gepaarde instrumentele rationaliteit, dat haar object causaal nomologisch interpreteert, is functioneel en behulpzaam waar het object de fysische natuur (natuurwetenschap en technologie), technische of fysiologische systemen, (economie en biologie) dus mechanistische te verklaren en besturen processen betreft. Maar zijn er aspecten van de menselijke psyche bij betrokken, zoals in de typische menswetenschappen (psychologie, sociologie, antropologie) en in de communicatieve praktijken van de leefwereld, dan beschouwen we vanuit een louter ‘objectiverend’ perspectief de mens ten onrechte als ding, als mechanistisch te verklaren (eventueel te manipuleren) object³. In zijn maatschappijkritiek constateert Habermas dat de rationaliteit van deze technologische systemen te veel in de leefwereld, filosofie en menswetenschappen is doorgedrongen. Het verdringt daarin (het zicht op) het open en belangeloze communicatieve handelen. Habermas wil daarom ten behoeve van een volledig en correct maatschappijonderzoek een overgang van subject-object- naar een subject-subject perspectief voltrekken. Zo verkrijgt men een kritische theorie die analyseert en interpreteert.

“Een kritische theorie verzoent, berust op een beschrijving en een verklaring van de huidige toestand die kritisch is en tegelijkertijd ook duidelijk maakt dat het eigen kritische standpunt in de sociale werkelijkheid aanwezig is. De theorie verwoordt de morele ervaring van de mensen en verdedigt hun reële belangen....Zij laat zien dat in de maatschappij krachten aan het werk zijn die de feitelijke toestand overstijgen....Omdat alleen vanuit een performatief standpunt de rationaliteit van mensen kan verschijnen, kunnen we bij de studie van de maatschappij geen genoegen nemen met een objectieve beschrijving.” [Habermas⁴]

Dat voor Habermas dit performatieve, of participerende perspectief zo belangrijk is in een volledige beschrijving van de mens, wordt door hem verantwoord in zijn eigen uitwerking van een algemener verspreide opvatting van taal: de gebruikstheorie van betekenis die met het werk van Wittgenstein wordt geassocieerd. Volgens een bepaalde opvatting van deze theorie kunnen we pas begrijpen wat mensen zeggen, dus wat mensen beweegt, wat ze bedoelen, als we ze niet slechts als buitenstaander waarnemen, maar als we participeren in de context van het taalspel waarvan de spreker deel uitmaakt. Daar is niet alleen een ‘erklären’ maar ook een ‘verstehen’ voor nodig. Dit inleven in de concrete ander wordt vanuit het participerend perspectief performatief mogelijk gemaakt doordat we in een discours onze uitspraken ten overstaan van anderen d.m.v. argumentatie rechtvaardigen. Of Habermas gelijk heeft in zijn stelling, dat het participerende perspectief onontbeerlijk is voor een volledig en kritisch beeld van de mens in de samenleving, zullen we verder onderzoeken door wat dieper in te gaan op de filosofische onderbouwing van zijn theorie. Maar eerst zal ik een aantal begripsverbanden duidelijk maken.

2.1 Verband tussen taalspel, taalgebruik, rationaliteit en perspectief op de wereld

Habermas heeft zich als doel gesteld een kritische analyse te geven van de sociale interacties tussen rationele individuen in onze moderne maatschappij. De aard van de rationaliteit van deze interacties komt tot uitdrukking in de taalspelen. Volgens Habermas kunnen we een onderscheid maken tussen twee hoofdtypen: het taalspel van het objectiverende systeem, bijvoorbeeld de natuurwetenschap, en het taalspel van mensen die participeren in

interpersoonlijke relaties in het sociale domein. Elk taalspel heeft zijn eigen regels voor gebruik (interpretatie).

In het taalspel van de objectiverende natuurwetenschap ligt de nadruk op het afbeeldende gebruik van de taal. Ze wordt opgevat als een instrument om een mogelijke 'objectieve' stand van zaken (van objecten in causale interactie) in de werkelijkheid af te beelden; de zgn. 'objectieve' wereld. De taalregels zijn toegespitst op deze functie. Zo is de formele wiskundig georiënteerde taal van de fysica waardenneutraal, dwz dat het alleen kan worden gebruikt om uit te drukken hoe dingen zijn. Het hoeft voor haar doel (beschrijven, voorspellen, en manipuleren van objecten in de wereld) tenslotte niet uit te drukken hoe dingen *horen* te zijn. Consequentie is echter dat deze 'objectiverende taal' niet voor alle communicatie tussen mensen van vlees en bloed kan worden gebruikt. Daarvoor ontbreekt het de afbeeldende functie van taal aan vermogen om waardegeladen, (normatieve) (inter)subjectieve en communicatieve aspecten van het leven tot uitdrukking te brengen. Maar zolang men deze 'objectiverende' taal gebruikt voor haar geëigende doel, beschrijving en beheersing van de objecten in de wereld, is er niets aan de hand. De regels van de 'objectiverende' taal vormen de instrumentele rationaliteit. En omdat volgens Habermas de taal constitutief is voor ons wereldbeeld, is het bij dit taalgebruik behorende beeld een manipuleerbare wereld van dingen en causaal verlopende processen en levert als zodanig het perspectief van de objectieve wereld.

In het taalspel van participerende, communicerende actoren die zich richten op het door argumentatie bereiken van consensus, wordt de taal op een andere manier gebruikt. Deze taalgebruikregels, die de communicatieve rationaliteit vormen, maken het mogelijk om (inter)subjectieve aspecten van het leven en het menszijn uit te drukken en te interpreteren. De interpretatieregels omvatten regels voor gebruik van mentale termen, (termen van de volkpsychologie), regels voor hermeneutisch gebruik, d.w.z. regels voor het beschrijven van de historie van het leven, van actoren en gemeenschappen, existentiële aspecten, zingeving etc. De impliciet in de regels aanwezige wederzijdse verwachting is dat de door mensen gekozen handelwijze wordt gestuurd door rationeel geselecteerde redenen die de (inter)subjectieve waarden uitdrukken die het waard zijn bevonden om nagestreefd te worden. De overkoepelende regels van communicatieve rationaliteit impliceren dat wij wezens zijn die bewust en vrij redenen kiezen, tot de onze maken en ons handelen er door laten leiden. De regels van dit taalgebruik constitueren het participerende perspectief.

Habermas kritiek is nu dat de eenzijdige nadruk op het instrumentele gebruik van taal, dus taalgebruik waarin het accent op een niet-normatieve afbeelding van de werkelijkheid ligt, in onze moderne maatschappij veel te dominant is geworden. Het is het domein van de leefwereld binnengedrongen, de wereld die het terrein behoort te zijn van het communicatief, op consensus gerichte gebruik van de taal. D.w.z. het taalgebruik wat het vermogen heeft het (inter)subjectieve aspect van (samen)leven uit te drukken. En aangezien Habermas de regels van het taalgebruik gelijk stelt met rationaliteit, heet het dat de instrumentele rationaliteit domineert over de communicatieve rationaliteit. Omdat Habermas meent dat onze cognitieve toegang tot de wereld immer linguïstisch bemiddeld is, en dus de taal die we hanteren het beeld van de werkelijkheid constitueert, wordt onze kijk op de wereld te veel gedomineerd door het objectiverende perspectief.

Een natuurwetenschappelijke waarnemer, die meent dat zijn objectieve beeld van een fysische werkelijkheid volledig samenvalt met de ontologie en nomologie van alles wat er is

(de allesomvattende natuur), vergist zich volgens Habermas. De natuur omvat meer dan de natuur van de natuurwetenschap. Er zijn eigenschappen zoals die van de geest of de cultuur die niet kunnen worden afgeleid uit de eigenschappen van de elementaire bouwstenen van het objectiverende perspectief. Daarom is het zicht van de objectiverende waarnemer niet een volledig zicht op de werkelijkheid, maar vormt het slechts een segment ervan. Het vereist de aanvulling van het perspectief van de participant. Hoe kan dit filosofisch onderbouwd worden?

2.2 Theoretisch filosofische onderbouwing van de formele pragmatiek

Wat nu volgt is een reconstructie van Habermas' conclusie dat we de processen in de wereld vanuit een zwak naturalisme moeten beschrijven. Habermas verzet zich tegen de idee dat we de wereld, inclusief de mens en de menselijke interactie, uitputtend in fysicalistische termen kunnen beschrijven. Dit, wat hij noemt 'hard naturalistisch' beeld van de werkelijkheid, is gebaseerd op de onjuiste veronderstelling dat we vanuit het objectiverende perspectief -"the view from nowhere"- een onbemiddelde kijk op de wereld hebben. Habermas werpt tegen dat zoiets ten principale onmogelijk is. Geïnspireerd door Kant's dictum dat 'intuïties zonder concepten blind zijn', stelt Habermas dat elk door ons ingenomen perspectief linguïstisch door een conceptueel schema wordt bemiddeld. Dat ruwe, zintuiglijke data geen interpretatie zou behoeven is de 'Myth of the Given'. Waarnemen is interpreteren en dus classificeren. De wereld wordt waargenomen door de 'onverwijderbare bril' van een interpretatiestructuur; de regels van een taal. De taal is het ordeningsprincipe zonder welke we letterlijk en figuurlijk niets over de werkelijkheid te zeggen hebben.

De consequentie van deze theoriegeladenheid van de waarneming is dat, als we iets willen zegen over de status en aard van onze empirische kennis, we de aard van dit ordenende principe van de taal zouden moeten onderzoeken. Het probleem dat zich hier voordoet is dat zo'n onderzoek uiteraard ook linguïstisch geladen is. De constatering van deze vicieuze cirkel was voor Wittgenstein het moment om 'de ladder omver te gooien' en het theoretiseren hier verder te staken. Zo niet voor Habermas. Hij gaat verder. Hij ziet alleen maar een vicieus probleem als we een empirisch onderzoek zouden doen. Een gewoon empirisch onderzoek, met de methoden van de 'objectivistische' wetenschap, stuit op een cyclische deductie, omdat we in de aanname al datgene moeten veronderstellen wat we in de conclusie willen verklaren. Daarom blaast Habermas de transcendentale analyse van Kant nieuw leven in om op zoek te gaan naar de mogelijkhedenvoorwaarden en toepassingscondities van onze kennis. Maar Habermas voegt wel een extra conditie toe. Waar Kant's transcendent⁵ zich buiten ruimte en tijd bevindt, meent Habermas dat dat niet houdbaar meer is. Sterker nog, hij wil de transcendentale analyse met een evolutietheorie verenigen. In die evolutietheorie wordt uiteraard niet (zoals wel het geval is in het neo-Darwinisme) zonder meer de objectivistische houding aangenomen, maar deze wordt aangepast aan de eisen van een pragmatistisch transcendentale analyse. Het houdt rekening met het feit dat onze toegang tot de wereld nimmer linguïstisch onbemiddeld is, verdisconteert het gegeven dat kennis vooral performatieve kennis is, ziet deze kennis als een ontwikkeling vanuit leerprocessen waarin probleemoplossen centraal staat en is antireductionistisch; d.w.z. reduceert de beschrijvingen van de praktijken vanuit het participantperspectief niet tot een objectiverende neurofysiologische beschrijving. [De hieropvolgende paragrafen 2.3.1 t/m 2.3.3 construeren de theoretische onderbouwing van de formele pragmatiek in meer detail.]

2.2.1 Afbeeldende en communicatieve functie van de taal

In principe ziet Habermas de linguïstische wending, ingezet door Frege, als een in potentie positieve wending in de filosofie. Het is een mooie gelegenheid om af te rekenen met de eenzijdige nadruk van de filosofie van Locke, Kant en Hume. Met hun accent op het analyseren van sensaties, representaties en oordelen, typeerden ze de mens onvolledig als een van de medemens en het handelen geïsoleerd rationeel en kennend wezen⁶. De concentratie op de analyse van de taal zou dan hebben kunnen laten zien dat *“uiteindelijk ...taal net zo goed gebruikt wordt om te communiceren als om te representeren, en [dat] de linguïstische uiting zelf een vorm is van handelen, welke wordt gebruikt voor het produceren van interpersoonlijke relaties.[Habermas⁷]* Taal is niet alleen het vehikel van het denken maar toch ook vooral een instrument om te communiceren. Maar door na de linguïstische wending de relatie representatie-object te vervangen door de relatie propositie-feit bleef de hoofdstroming in de analytische filosofie (met boegbeelden als Russel, Carnap en Quine) vasthouden aan het primaat van bewerende proposities en haar representerende functies⁸. De enige relatie was er één tussen teken en iets in de werkelijkheid.

Maar de latere Wittgenstein toonde door een reflectie op het dagelijks taalgebruik dat taal voor veel meer wordt gebruikt dan voor alleen beschrijven en beweren. Zo heb je naast beweringen ook aanwijzingen, verplichtingen, uitingen en verklaringen. Er is dan ook geen starre betekenis te geven van woorden in onze taal. Betekenis wisselt van context tot context (taalspel). De relatie die de betekenis van een teken vastlegt, tussen teken en iets in de werkelijkheid, zou dus uitgebreid moeten worden met een relatie met de context. Bovendien toonde Wittgenstein in zijn strijd tegen het mentalisme aan dat het Augustiniaanse beeld van betekenis (betekenis opgevat als een mentale entiteit) incoherent is. Betekenis van een uitdrukking is niet zozeer een ‘know that’ als wel een ‘know how’, een regel voor gebruik. Betekenis heeft een regelgestuurde karakter. Men volgt een regel als men een uitdrukking interpreteert.

In zijn taalanalyse wees Wittgenstein ook op het feit dat het begrip betekenis op twee manieren kan worden opgevat. Enerzijds is er de semantische betekenis, als verwijzing. Het woord ‘brood’ verwijst naar een bepaald voorwerp. Dat is de manier waarop het woord ‘betekenis’ figureert in de representerende functie van taal, die taaldaden uitsluitend als beweringen (over een stand van zaken in de werkelijkheid) opvat. Maar er is ook een ‘betekenis’ als relevantie: dat wat iemand bedoelt als hij bijvoorbeeld zegt “Voor mij betekent Mozart heel veel, Wagner niets”. Het onderscheid tussen het begrijpen van de woorden van een gedicht en het begrijpen van het gedicht wordt duidelijk aan de hand van het verschil tussen de concepten ‘semantiek’ en ‘relevantie’. Meestal zal men bij de opbouw van een betekenisstheorie bijna exclusief aandacht besteden aan het semantische aspect, en zich niet de moeite getroosten verband te leggen met het concept ‘betekenis als relevantie’. Wittgenstein daarentegen was ervan overtuigd dat ook de semantische betekenis enkel ten volle begrepen kan worden vanuit betekenis als relevantie. Het essentiële van de menselijke betekenis is een subjectief aspect van geraakt of getroffen kunnen worden door iets waarvan men beseft dat het ook betekenisloos kan worden⁹.

De taalfilosofen, Austin, Grice en Searl werkten deze inzichten uit in een taaldadenanalyse. Dummett bracht een epistemische wending aan, door af te leiden dat de betekenis van

uitdrukkingen in het gebruik ligt dat we ervan maken bij het bereiken van overeenstemming door argumentatie.

Deze taalanalytische inspiratiebronnen vormden tezamen één van de peilers van de Habermas' communicatietheorie. Hij leidde hieruit af dat we de woorden van een spreker niet kunnen begrijpen zonder ons af te vragen of de taalhandeling die de spreker met die woorden stelt, gegrond of geldig is; een taalhandeling begrijpen is weten dat ze aanvaardbaar is.

Dat iemand iets heeft gezegd, is niet zomaar een feit uit zijn biografie (zoals het feit dat ie stottert of een stopwoord gebruikt). De spreker heeft de bedoeling een taalhandeling te stellen waarvoor hij gegronde redenen of argumenten meent te hebben. Of zoals Habermas het formuleert: *“hij spreekt zich uit over een wereld en beroept zich dus op iets in een wereld waarover ook de hoorder kan oordelen¹⁰.”* Habermas concludeert hieruit dat de representerende en communicatieve functie van taal elkaar wederzijds vooronderstellen. M.a.w. dat ze 'equiprimordiaal' zijn. Dus in het begrijpen van taaluitingen zijn het participantperspectief en het 'objectiverende' waarnemersperspectief van gelijk belang. Ze vooronderstellen elkaar, en tonen elk een eigen segment van de wereld

Participantperspectief moet begrepen worden als het aannemen van een bepaalde houding waarin de betekenis-als-relevantie van iemands uitdrukkingen naar voren kan komen. Deze ruimere opvatting van het begrip 'betekenis', wordt door Habermas verwoordt in de stelling dat, *“de betekenis[-als-relevantie] van de taalhandeling van een spreker pas duidelijk wordt als we weten hoe een spreker in een argumentatie¹¹ diens uitspraken rechtvaardigt¹²”* Door een taaldaad in een discours te beargumenteren worden de interpretatiekaders van participanten gesynchroniseerd. Dit scheidt de mogelijkheid ons in het particuliere perspectief van de concrete ander te verplaatsen.

2.2.2 Pragmatisme

Het verwerpen van de louter op de representerende functie georiënteerde opvatting van taal impliceert de verwerping van de idee dat waarheid synoniem is met correspondentie tussen afbeelding (propositie) en iets in de werkelijkheid. Habermas heeft dus een andere definitie van waarheid nodig. Inspiratie vindt hij bij de pragmatist Dewey in wiens darwinistische visie er geen streven naar van belangen ontblote waarheid zou bestaan. Al ons denken is een soort van probleemoplossen. We doen er volgens hem goed aan de beschrijving van onderzoek gebaseerd op de notie van werkelijkheid, dus de correspondentietheorie van waarheid, te verlaten. Evenals Kant verwerpt ook Dewey de zgn. toeschouwertheorie van kennis. Deze theorie van Locke beschouwt de geest als een tabula rasa, als een donkere kamer waar het licht van de ervaring instroomt, waarin het waarnemen zuivere receptie is, en de geest in de perceptie fundamenteel passief. Dewey brengt hier tegenin dat *“de ervaring...problemen creëert die moeten worden opgelost. ..De moderne mens heeft geleerd de natuur aan te passen en hoeft niet zijn passieve subject te zijn¹³”*. Kennis is het resultaat van handelen, en niet van een passief schouwen. Kennis ontstaat uit een performatieve omgang met de wereld, waarin we door intelligente omgang met een risicovolle omgeving door een ervaring van frustratie heen werken. Kennis is zo beschouwd dus praktische kennis: 'know-how' en niet 'know-that'. Het dualisme van theorie en praktijk werd daarom door Dewey ingeruild voor de 'experimental theory of knowledge'. Het soort van ervaring waaruit kennis ontstaat noemde Dewey 'inquiry'. Het doel van inquiry (aan de hand waarvan hij meende heel wat dichotomieën te kunnen elimineren¹⁴) is het verkrijgen van kennis die uitgedrukt wordt in

proposities die gegarandeerd verdedigbaar zijn, dwz. die door de gemeenschap van onderzoekers niet als problematisch maar als ‘geünificeerd’ geheel worden gevoeld. De ‘inquiry’ als kennis voortbrengend proces is een strikt empirische kennistheorie die anti-essentialistische is. Dwz. ze zoekt geen fundering in zoets als Hume’s impressies of Russell’s ‘sense-data’, waar weer dubieuze claims over onfeilbaarheid voor moeten worden uitgebracht. Habermas is duidelijk schatplichtig aan de het pragmatisme. Hij adopteert veel van de opvattingen van Dewey en verwerkt ze in zijn pragmatistisch model. “*De ‘spiegel van de natuur’ – de één-op-één representatie van de werkelijkheid – is het verkeerde model van kennis. De hantering van een twee-plaatsige relatie tussen beeld en hetgeen het afbeeldt, en de statische relatie tussen een propositie en een stand van zaken, verduistert de dynamiek van kennisaccumulatie als proces van probleemoplossen en rechtvaardiging. [Habermas¹⁵]*“

2.2.3 Transcendentale analyse en detranscendentalisering van mogelijkheidsvoorwaarden

Volgens Habermas biedt het waarnemerperspectief geen geprivilegieerde cognitieve toegang tot de wereld. Ze is niet direct maar altijd bemiddeld door een talige interpretatiestructuur. Het waarnemerperspectief vooronderstelt evenzogoed de talige, rechtvaardigende praktijken van het participantperspectief. Talige uitspraken moeten geïnterpreteerd worden. Die interpretatieregels liggen volgens Habermas vervat in socio-culturele normen van praktijken in de leefwereld, die alleen performatief, (intuïtief) gevat kunnen worden. Gegeven dat elk empirisch onderzoek zelf ook talig is, vraagt Habermas zich af wat hij kan doen om de mogelijkheidsvoorwaarden van onze kennis te onderzoeken, zonder in de aporie van circulariteit te stranden. De oplossing moet gevonden worden in de toepassing van Kant’s transcendentale analyse¹⁶. Maar een transcendentaal idealisme zoals dat van Kant is voor Habermas vandaag de dag geen optie meer. Dat zou elke wortel van realistische claim uit de natuurwetenschap verwijderen.

In ‘Truth and Justification’ geeft hij een filosofische onderbouwing van een aan het pragmatisme aangepaste transcendentale analyse, gericht op het onderzoeken van de mogelijkheidsvoorwaarden van kennis.

Deze de-transcendentalisatie van de transcendentale filosofie bestaat in een reconstructie van de Kantiaanse noodzakelijke (a-priori) voorwaarden onder welke iets het object van onze ervaring en cognitie kan zijn. Na een generalisatie door het afscheiden van mentalistische basisconcepten van zelfreflectie en het foundationalistische a-priori / a-posteriori begrippenpaar, verwijst ‘transcendentale analyse’ naar de zoektocht naar waarschijnlijke, universele, maar alleen de-facto onvermijdbare voorwaarden die moeten worden vervuld, willen fundamentele praktijken of prestaties verrijzen. De (transcendentale) mogelijkheidsvoorwaarden voor kennis moeten niet, zoals bij Kant, worden gevonden in begripsordenende vermogens in de geest, maar in de regels (normen) van het taalspel van het communicatief handelen. “*De intuïtieve kennis over hoe een regel te volgen en wat het is om een regel te schenden, is inherent normatief. Dit kan niet worden ingevangen door een notie van waarneming die is gelimiteerd door empirische regelmatigheden¹⁷.*” Het transcendentaal onderzoek richt zich nu op deze regels, leefwereldpraktijken, waarvan de kennis wordt gezien als een functie van leerprocessen, en die gelegen zijn in diepgewortelde structuren van de achtergrond van de leefwereld. Dit geeft aanleiding tot een architectuur van leefwereld en objectieve wereld, welke correspondeert met een methodologisch dualisme van verstaan en verklaren.

2.3 Het onderscheid hard en zwak naturalisme

We hebben nu genoeg informatie om de twee door Habermas geconstrueerde posities van het naturalisme te begrijpen. De door hem gegeven benaming 'hard naturalisme' is wat gewoonlijke onder naturalisme wordt verstaan. Het heeft de volgende karakteristieken.

Hard Naturalisme

- HN-1) Het model van kennis is gebaseerd op de 'spiegel van de natuur' – de één-op-één representatie van de werkelijkheid – en een correspondentietheorie van waarheid.
- HN-2) Wetenschap wordt beschouwd als het hoogste tribunaal voor waarheid. Het wordt niet onderworpen aan een transcendentaal onderzoek naar haar toepassings- en mogelijkheidsvoorwaarden daar het de aanname van een transcendente benadering verwerpt. Alle cognitie is uitiem te herleiden tot empirische processen.
- HN-3) Hanteert een reductionisme, d.w.z. haar verklaringsstrategie richt zich op het vervangen van de conceptuele analyse van praktijken van de leefwereld door een wetenschappelijk neurologische of biogenetische verklaring van de prestaties van het menselijke brein.

Habermas bezwaar op de karakteristiek, zoals genoemd onder (HN-1) is dat de representerende functie van de taal (de statische relatie tussen propositie en stand van zaken) niet voldoende rekenschap geeft van de met het handelen verbonden ervaringscontext, en van de discursieve rechtvaardiging.

Zouden we de wetenschap onderwerpen aan een transcendentaal onderzoek (HN-2), dan zou blijken, aldus Habermas, dat de wereld van verschijningen slechts een gedeeltelijk segment of een perspectivistisch verstoord beeld van de wereld-in-zichzelf representeert. Zo'n onderzoek zou laten zien dat er nood is aan een methodologisch dualisme wat bestaat uit (a) een interpretatieve reconstructie van onze leefwereld vanuit het perspectief van een participant op het communicatief handelen, en (b) een objectiverend perspectief op het verklaren van causale processen in de objectieve wereld. Maar in een hard naturalisme wordt de transcendente aanname verworpen. De Amerikaanse filosoof W.V.Quine¹⁸ is het paradigmatisch voorbeeld van een hard naturalist.

Het conceptueel reductionisme (HN-3) loop volgens Habermas vast op de tegenstrijdigheid tussen het zelfbegrip van competente sprekers en een contra-intuïtieve, revisionistische zelfbeschrijving. *“Subjecten die het vermogen hebben van spraak en handelen, en betrokken zijn in communicatieve praktijken kunnen niet anders dan zich oriënteren op hun denken en doen door middel van de normen, en worden bewogen door redenen. Zij kunnen zich niet herkennen in objectiverende beschrijvingen¹⁹.”*

Ook vormt de vertaling van intuïtieve kennis van normatieve praktijken van subjecten (met het vermogen van spraak en handelen) een probleem. Het is ten principale niet zonder rest te vertalen in het theoretische idioom van de nomologisch, empirische wetenschappen.

Het beeld dat Habermas in zijn zwak naturalisme schetst, is dat van een universum welke meer omvat (verwantwoordelijk voor bijvoorbeeld bepaalde eigenschappen van de geest en cultuur) dan de natuur van de natuurwetenschappen. Het erkent dat het waarnemerperspectief slechts een segment (of een perspectivistisch verstoord beeld) van de ‘welt-an-sich’ bestrijkt.

Zwak Naturalisme

- ZN-1) Waarheid wordt niet in een solo-activiteit verbonden aan correspondentie met een stand van zaken, maar wordt (net als juistheid en waarachtigheid) in een discours met aller instemming intersubjectief gevestigd.
- ZN-2) Wetenschap wordt als elk ander kennisproces onderworpen aan een (‘ge-de-transcendentaliseerd’) transcendentaal onderzoek naar haar toepassings- en mogelijkheidsvoorwaarden.
- ZN-3) Hanteert een anti-reductionisme. D.w.z. de continuering van leerprocessen op een hoger niveau worden niet reductionistisch verklaard in termen van wetenschappelijk neurologische of biogenetische verklaringen van de prestaties van het menselijke brein! [HBB-24]
- ZN-4) Het houdt vast aan een ontologisch monisme.

Ad ZN-1) Vanuit een pragmatistisch perspectief is werkelijkheid niet iets wat moet worden gekopieerd. We nemen er performatief notie van als de totaliteit van weerstanden die worden verwerkt en geanticipeerd. De werkelijkheid maakt zich bekend in de beperkingen en begrenzungen waaraan probleemoplossende activiteiten en leerprocessen worden onderworpen. Er wordt rekening gehouden met de dynamiek van kennisaccumulatie door probleemoplossen en rechtvaardiging.

- *“In de ruimtelijke dimensie is kennis het resultaat van het door een ervaring van frustratie heen werken, door intelligent om te gaan met een risicovolle omgeving.*
- *In de sociale dimensie is het ’t resultaat van het rechtvaardigen van jouw manier van probleem oplossen tegenover de bezwaren van andere participanten in een argumentatie.*
- *En in de temporele dimensie is het ’t resultaat van een leerproces gevoed door de revisie van de eigen fouten.” [Habermas²⁰]*

Ad ZN-2) De aanvullende voorwaarde die Habermas aan de transcendentale analyse stelt is dat (ondanks de detranscendentalisatie) een onderscheid wordt gehandhaafd tussen wat wereld (welt-an-sich) en wat inner-wereldlijk is.

“Deze opvatting is gebaseerd op één enkel meta-theoretische aanname: namelijk dat ‘onze’ leerprocessen (die mogelijk zijn binnen het framework van socioculturele vormen van leven), in zekere zin simpel de continuering zijn van voorafgaande ‘evolutionaire leerprocessen’, die op hun beurt aanleiding gaven tot onze vormen van leven. Want dan blijken de structuren, die de transcendentale voorwaarden van mogelijkheid van onze soort van leerprocessen vormen, het resultaat te zijn van minder complexe natuurlijke leerprocessen – die daarbij zelf een cognitieve inhoud verkrijgen.” [Habermas²¹]

De toename van ‘cognitieve’ vermogens van organismen in een eerder stadium van de evolutie moeten we dus begrijpen en verklaren vanuit het model van kennisaccumulatie door probleemoplossen, (en rechtvaardiging in argumenteren zodra de taal en cultuur begint te ontstaan ook).

Ad ZN-3) Een antireductionisme is *“tevreden gesteld met de achtergrondaanname dat de biologische begiftiging, en de culturele wijze van leven van homo-sapiens een ‘natuurlijke’ oorsprong hebben en in principe verklaard kunnen worden in termen van de evolutietheorie. Het is terughoudend in het adopteren van enige filosofische aanname t.a.v. de relatie tussen geest en lichaam, in de zin van eliminatief of reductief materialisme”*. [Habermas²²] “

En,

“Incorporeert noch onderwerpt het ‘interne perspectief’ van de leefwereld (maakt het ondergeschikt) aan het ‘externe perspectief’ van de objectieve wereld. In plaats daarvan houdt het deze theoretische perspectieven gescheiden, verbindt ze ze als het ware op een metatheoretisch niveau door een continuïteit tussen natuur en cultuur te veronderstellen.”[Habermas²³]

Zolang we de kwestie vormgeven in transcendentale termen, moeten we volgens Habermas scherp onderscheid maken tussen:

- hermeneutische benaderingen van een rationele reconstructie van de structuren van de leefwereld, welke we ondernemen vanuit het perspectief van de participanten, en de
- observatie-gebaseerde causale analyse van hoe deze structuren natuurlijk zijn geëvolueerd.

2.4 Proefrit maken met Habermas’ metatheoretische verbinding van perspectieven

Bepaalde aspecten van cultuur en geest zijn niet door de natuurwetenschap te verklaren. Het participantbeeld van een agent wiens handelen wordt bewogen door vrij gekozen redenen is volgens Habermas niet te reduceren tot het objectiverende beeld van de natuurwetenschappelijke waarnemer die alles in termen van causaal en nomologisch verlopende processen verklaard. Het door de natuurwetenschapper geschetste wereldbeeld mag niet tot allesomvattende natuur worden verabsoluteerd. De twee beelden van het participant-, en waarnemerperspectief moeten worden erkend als complementen die een zelfde status in epistemisch gezag vertegenwoordigen. Overeenkomstig de in ZN-3 gestelde eis moeten ze op niet-reductionistische wijze metatheoretisch worden ondergebracht in één geünificeerd beeld van de natuur, welke meer omvat dan de natuur van de natuurwetenschap.

Dit lijkt op het eerste oog misschien ongecompliceerd maar stuit op problemen bij nadere reflectie op de realisatie van zo’n metatheoretische constructie. Want hoe verloopt de interactie tussen enerzijds fenomenen van de geest en cultuur die volgens Habermas onder de allesomvattende natuur maar niet onder het domein van de natuurwetenschappen vallen, en fenomenen die daar wel onder vallen? In het zwak naturalisme wordt verondersteld dat er een ‘transcendentale wereld’ achter het participant-, en waarnemerperspectief ligt verscholen. Volgens deze visie is in die wereld een grond aanwezig voor eigenschappen van het participantperspectief (bijvoorbeeld die van het mentale), en voor eigenschappen van het waarnemerperspectief (bijvoorbeeld de nomologische causaliteit van neurofysiologische processen).

Habermas zal moeten aannemen dat de ‘transcendentale wereld’ op een of andere manier de correlaten van de interacties tussen fenomenen van het participantperspectief en die van het waarnemerperspectief bevat. De metatheorie moet de perspectieven en deze interacties verantwoorden door het geven van een verklarende beschrijving van de verbinding tussen hun bronnen in de transcendentale wereld. Dit is volgens Habermas niet zozeer een beschrijving van de transcendentale wereld, als wel van de mogelijksvoorwaarden van onze kennis. Habermas stelt dat deze metatheorie *...terughoudend [is] in het adopteren van enige filosofische aanname t.a.v. de relatie tussen geest en lichaam (bijvoorbeeld in de zin van eliminatief of reductief materialisme)*²⁴

Maar hoe kun je zo’n ‘terughoudende’ positie volhouden? In het filosoferen in dit domein kan je niet zonder een aanname betreffende de relatie geest-lichaam. Gebeurt het niet expliciet dan toch minstens impliciet. Dus zullen we toch op de aard van de relatie moeten reflecteren. En als je dan de consequenties van de door Habermas voorgestelde metatheorie gaat doordenken dan blijkt deze voor de natuurwetenschap allesbehalve terughoudend. De natuurwetenschap zal namelijk, als je de interactie tussen geest en lichaam wil verklaren, de aanname van causale geslotenheid (de eerste en tweede wet van de thermodynamica²⁵) voor haar domein op moeten geven, want dat vormt in het zwak naturalisme niet het allesomvattende universum. We zijn er zeker van dat geest en cultuur (waarvan de eigenschappen zich openbaren vanuit het participerende perspectief) in onderlinge interactie verkeren met fysische processen (die vanuit het objectiverende perspectief van de natuurwetenschap worden geopenbaard). Maar in het zwak naturalisme kan louter de natuurwetenschap deze kruislingse interacties niet meer verklaren. Het dogma van natuurwetenschap is echter de aanname van causale geslotenheid van het universum dus dat alle interacties in de wereld in principe fysisch causaal (nomologisch) verklaarbaar zijn. Dat dogma moet in een zwak naturalisme dus opgeofferd worden, tenzij het mentale als een epifenomeen (causaal impotent fenomeen) wordt beschouwd.

Het door Kim²⁶ ontwikkelde argument tegen eigenschapdualisme, wat toont dat de daarin gerepresenteerde mentale toestanden epifenomenen moeten zijn, geldt namelijk ook voor de door Habermas voorgestelde oplossing. Het eigenschapdualisme is de doctrine die er vanuit gaat dat ons brein twee verschillende klassen van onderling niet-reduceerbare eigenschappen bezit; mentale en fysische. Omdat Habermas’ zwak naturalisme uitgaat van een begrenzing van het domein van de natuurwetenschap en dat bepaalde eigenschappen van de geest en cultuur niet reduceerbaar zijn tot het fysische domein kunnen we zijn theorie ook als een eigenschapdualisme typeren.

Maar noch de verwerping van de causale geslotenheid van het universum noch het epifenomenalisme zijn aantrekkelijke oplossingen. We zullen daarom op zoek moeten naar een metatheorie die verklaart hoe de interactie tot stand komt tussen dat wat ten grondslag ligt aan ons beeld van geest en het beeld van het lichaam, maar daarbij niet toegeeft aan het epifenomenale van het mentale en de verwerping van causale geslotenheid. Aanvullend hierop moet de metatheorie verantwoorden dat deze interactie niet nomologisch van aard is, teneinde vanuit Habermas’ optiek de vrijheid van de geest te kunnen garanderen.

Dit wordt al met al een lastig op te lossen theoretisch probleem. Gezien het feit dat Habermas vasthoudt aan een monistische ontologie resteren vermoedelijk een fysicalisme (materialisme) of een metafysisch idealisme als opties. Een idealisme valt af (zo meent ook Habermas) omdat dit wetenschappelijke claims elke realistische grond ontnemt.

Laten we, in relatie tot het fysicalisme, eens wat nauwkeuriger kijken naar wat Habermas met zijn zwak naturalisme voor ogen staat. Hij meent dat de natuur meer moet zijn dan de natuur van de fysica. M.a.w. hij verwerpt reductionistisch fysicalisme. Maar wat houdt fysicalisme nu precies in? Om te beginnen moeten we een onderscheid tussen ontologisch en nomologisch fysicalisme aanbrengen. Ontologisch fysicalisme is de visie dat alles in de wereld is samengesteld uit door de natuurwetenschap erkende bouwstenen of basisentiteiten. Nomologisch fysicalisme is de visie dat bovendien alle eigenschappen verklaard kunnen worden aan de hand van de fundamentele wetten van de fysica. Petit formuleert dit als volgt;

1. Er is een empirische wereld bestaande uit microfysische deeltjes van de soort die de fysica veronderstelt.
2. Alles in de wereld is uitputtend zonder rest samengesteld uit microfysische deeltjes die zelf niet meer samengesteld zijn.
3. Microfysische deeltjes zijn onderworpen aan microfysische wetmatigheden met betrekking tot hun microfysische eigenschappen.
4. En het nomologisch fysicalisme voegt daar nog aan toe: alles in deze wereld is onderworpen aan microfysische wetmatigheden.

Een voorbeeld is het atomisme, dat verwant is aan nomologisch fysicalisme. Hierin staan de atomen op zichzelf. De wijze waarop atomen samenstellingen vormen kan volledig uit de atomen zelf verklaard worden. Stel dat een uit afzonderlijke atomen a,b en c samengesteld gecompliceerd molecuul een eigenschap $R(a,b,c)$ bezit, dan stelt het nomologisch fysicalisme de reductionistische eis dat deze macroscopische eigenschap R, logisch kan worden afgeleid (gededuceerd) uit de meest complete kennis van de eigenschappen van a,b en c.

Laat ons nu eens kijken naar een antireductionistische visie, het emergentisme.²⁷ Zou dit een kandidaat voor Habermas' zwak naturalisme kunnen zijn?

Het eigenschap-emergentisme (EE) accepteert ontologisch, maar verwerpt het nomologisch fysicalisme. Het emergentisme gaat er vanuit dat bij voldoende complexiteit in de samenstelling van elementaire fysische deeltjes nieuwe eigenschappen met eigen wetmatigheden emergeren die niet tot fysische wetmatigheden te reduceren zijn. Met andere woorden, een geheel samengesteld uit delen bezit eigenschappen die niet aangetroffen worden bij de delen en die ook niet tot de eigenschappen van de delen zijn te reduceren. Het beeld dat zo van de natuur ontstaat is gelaagd. De lagen zijn gerangschikt in termen van toenemende complexiteit waarbij elke laag met zijn eigen wetenschap correspondeert. Er bestaat verschil van mening over de rangschikking van de niveaus. Zo stelt Emmeche de volgende ordening van laag naar hoog niveau voor; fysisch, chemisch, biologisch, mentaal, en sociaal niveau. Salthe verdedigt een ordening waarbij het sociale niveau onder het mentale niveau valt.

Elk niveau heeft zijn eigen macroscopische wetten die niet tot de wetten op een lager niveau te herleiden zijn. De basisvraag in het debat tussen nomologisch fysicalisten en anti-reductionisten betreft de verhouding tussen macroscopische, en microscopische wetten. Een eliminativistische dus reductionistische positie ontkent dat er macroniveauwetten zijn die fundamenteel zijn. Alles is uiteindelijk te verklaren in termen van microscopische wetten.

Het eigenschapdualisme is een bijzonder geval van emergentisme. Het emergentisme is een ontologisch monisme en biedt ruimte aan emergerende eigenschappen die zichtbaar zijn vanuit het participantperspectief en niet vanuit het fysische perspectief. Maar de problemen

liggen weer in het tegenargument van Kim, die toont dat de macro-eigenschappen, en macrowetten niet anders dan in verband met epifenomenalisme moeten worden gebracht.

Een tweede argument tegen het bestaan van emergentisme treffen we in historische ontwikkelingen van de natuurwetenschap aan. Eigenschappen van fenomenen die aanvankelijk onvoldoende werden begrepen en daarom a.h.v. een emergentisme werden verklaard, bleken later een bevredigend reductionistische verklaring te kunnen worden gegeven²⁸.

Getuigd emergentisme niet van een wetenschappelijke luiheid, overgave of opportuniteit? Als je nog niet begrijpt hoe gecompliceerde structuren tot elementaire wetten kunnen worden gereduceerd, of als het veronderstellen van antireductionisme voor de eigen theorie opportuun is, dan kan gemakzucht of opportuniteit onterecht aanleiding zijn om emergerende eigenschappen te poneren. Emergentisme mag dan het ene probleem oplossen, maar je schiept er ook allerhande nieuwe problemen mee.

Een bijzonder probleem van emergentisme is het element van onvoorspelbaarheid. Zelfs de ideale theoreticus met complete kennis van de aggregaten en eigenschappen in een bepaald niveau is hulpeloos in het voorspellen van het plots opduiken van een emergente eigenschap vanuit een niveau daaronder. In het algemeen kan gesteld worden dat het sowieso moeilijk zal zijn aan de wetenschappelijke wereld een reductionisme op te dringen. De wetenschap wordt gedreven door het ideaal om zoveel mogelijk verschijnselen in geünificeerde theorieën onder te brengen. Zoals Herbert Feigl het uitdrukte: *“De wetenschap streeft naar het samenvatten van een maximum aan feiten en regelmatigheden in termen van een minimum aan theoretische concepten en veronderstellingen”*.

Emergentisme of niet, het ziet er naar uit dat het door Habermas voorgestelde zwak naturalisme ernstige implicaties voor de aannames en het verklarende vermogen van de natuurwetenschap zal hebben. Antireductionisme in combinatie met compartimenteren van de natuurwetenschap roept problemen voor haar op, en is tegennatuurlijk aan haar ideaal van unificatie en doorgronding van alle aspecten van de werkelijkheid. Er valt derhalve te verwachten dat het niet (eenvoudig) door de natuurwetenschap geaccepteerd zal worden.

Is er voor al de problemen die samenhangen met Habermas' introductie van het zwak naturalisme niet een andere oplossing te vinden?

Stel nu eens dat we uitgaan van een ontologisch, en nomologisch fysicalisme, dus van een louter door fysische entiteiten geconstitueerd universum. Waarin alle processen en gebeurtenissen corresponderen met hetgeen de fysica erover beweert. Maar waarin de verschillende perspectieven interpretaties zijn van die louter fysische werkelijkheid. Dan doen zich de hierboven besproken problemen van de causale interactie tussen het mentale en het fysische niet voor. Claimen we daarbij dan ook nog een antireductionisme, dan is in ieder geval aan de in ZN-3 weergegeven eis van Habermas voldaan. Dit geschetste scenario komt overeen met de visie van Dennett's fysicalistisch interpretationisme. Dit zullen we dan nu eens gaan bekijken.

3 Centrale thema's in de filosofie van Dennett

De afgelopen decennia heeft Daniel Dennett aan de filosofie van de geest, psychologie en moraal een invloedrijke bijdrage geleverd aan kwesties die in verband kunnen worden gebracht met notie van vrijheid, verantwoordelijkheid en persoonlijkheid. Net als Habermas tracht Dennett het beeld van de mens als vrije, morele actor te verzoenen met de visie dat de

gebeurtenissen in de wereld natuurwetenschappelijk kunnen worden verklaard. In tegenstelling tot Habermas wil Dennett deze verzoening bereiken vanuit een erkenning van het objectiverende perspectief (van de natuurwetenschap) als hoogste epistemische gezag.

In Dennett's filosofie staan twee opvattingen centraal: betreffende intentionaliteit als onderdeel van een interpretationisme en betreffende de aard van het bewustzijn.

Dennett's fysicistisch interpretationisme gaat er vanuit dat de mens in een principieel fysisch te verklaren wereld, omwille van verklaring en voorspelling van het gedrag van subjecten en objecten, over verschillende interpreterende perspectieven beschikt.

Dennett's theorie van bewustzijn als verklaring voor de werking van het brein behelst een verwerping van de metafoer van het cartesiaans theater ten gunste van de theorie van de 'meervoudige schetsen'. Het cartesiaanse beeld van iets gewaarworden als het binnentreden van een speciale kamer in ons brein, waar ons waarnemende en wilskrachtige ik zich bevindt is incoherent. In plaats daarvan moeten we bewustzijn opvatten als het tijdelijk dominant worden van één van de vele parallel aan elkaar verlopende gedistribueerde neurale processen in het brein. Beide thema's komen nu aan bod.

3.1 Dennett's interpretationisme

Een bijzondere kijk op onze epistemische relatie met de wereld is gezichtsbepalend voor Dennett's filosoferen. Ik heb het hier over zijn kijk op intentionaliteit, in relatie tot de propositionele attitudes zoals 'geloven dat', of 'verlangen dat'. Maar eerst kijken we even naar de betekenis van het begrip 'intentionaliteit'.

3.1.1 Intentionaliteit

Wat wordt eigenlijk bedoeld met de term 'intentionaliteit'? De term 'intentie' komt van het Latijns 'intendere' dat 'wijzen naar' betekent²⁹. De traditionele omschrijving van intentionaliteit luidt: "intentionaliteit is het kenmerk van het bewustzijn gericht te zijn." Bewustzijn is altijd bewustzijn van iets. Deze 'object-gerichtheid' van mentale verschijnselen werd door Brentano³⁰ gezien als het kenmerk waarmee het zich van het fysische onderscheid. Met zijn stelling dat intentionaliteit het onherleidbare en unieke kenmerk van de geest is, meende Brentano te kunnen rechtvaardigen dat de psychologie een onherleidbare zelfstandige wetenschap moest zijn. Elke poging om de geest in natuurwetenschappelijke termen te begrijpen zou bij voorbaat kansloos zijn. Er zou immers altijd één eigenschap van mentale fenomenen overblijven die zich niet in de termen van de biologie of neurofysiologie zou laten verklaren. De stelling van Brentano is vanaf het begin betwist. Zo kan men zich afvragen of alleen mentale verschijnselen intentioneel zijn. Tal van natuurverschijnselen lijken een soort van teleologie of doelgerichtheid te bezitten. Vooral in de biologie, of het nu gaat om de functie van een orgaan, om het handhaven van een toestand (homeostase), om de interactie met de omgeving of om het voortbestaan van de soort. Sinds de jaren 50 van de vorige eeuw hebben ook fysische artefacten zoals de computer vanwege hun gerichtheid op objecten de aandacht getrokken. Een kenmerk van intentionaliteit zoals Brentano dat bespreekt, is dat het in zijn formulering en beschrijving een beroep op introspectie doet. Maar de logisch positivisten en de behavioristen waren van mening dat we alleen kennis van publiekelijke waarneembaarheden kunnen hebben. Daarom mocht alleen het derde-persoonsperspectief dienen als verificatie van de zinvolheid van kennisuitspraken. Als men aannam dat uitspraken

bedoeld waren om naar innerlijke toestanden te verwijzen dan waren ze niet verifieerbaar, dus betekenisloos. Dus behavioristen verwierpen ofwel het hele mentale idioom, ofwel ze paraphraseerden mentale toeschrijvingen in termen van zogenaamde publiekelijk waarneembare responses op hypothetische stimuli. Vanwege moeilijk te repareren inconsistenties en de opkomst van de computertechnologie, moest het behaviorisme zijn positie als hoofdstroming afstaan aan de identiteitstheorie en het functionalisme. Deze theorieën vatten mentale toestanden van subjecten op als interne breintoestanden en onvervalste oorzaken. Maar zoals Dennett in zijn opvattingen betreffende intentionaliteit laat zien, is het behaviorisme nog steeds niet dood. Hij combineert het echter met aspecten van het functionalisme en pragmatisme. Het moet gezegd dat zijn interpretationisme subtieler is dan het behaviorisme van de eerste helft van de vorige eeuw.

3.1.2 De drie verschillende houdingen van het Interpretationisme

Als we het gedrag van complexe systemen (machine, computer, organisme, mens) willen voorspellen en verklaren dan staan ons volgens Dennett een fysisch, ontwerp, en een intentioneel perspectief ter beschikking.

In een benadering vanuit het **fysisch perspectief** bepaal je eerst de actuele toestand van het systeem, en de fysische aard van de inwerkingen. Vervolgens ga je met deze gegevens en kennis van de natuurwetten de uitkomst (eventueel tot op het microscopische niveau) voor een willekeurige input berekenen. *“Dit is de grootse en onpraktische strategie van Laplace voor het voorspellen van de gehele toekomst van alles in het universum. Ze is niet altijd praktisch beschikbaar, maar dat het altijd zal werken is een natuurwetenschappelijk dogma [Dennett³¹]”*. In principe kun je zo zeer betrouwbaar en nauwkeurig het systeemgedrag voorspellen. Deze nauwkeurigheid heeft echter wel een prijs. De hoeveelheid data die je tot je beschikking moet hebben, en de omvang van het rekenwerk, neemt met toenemende systeemcomplexiteit exponentieel toe. Dit maakt het al niet meer praktisch en zinvol om elke schaakcomputerzet op het fysische niveau door berekening te voorspellen. Daarbij maakt het opereren op het fysisch niveau blind voor wat het meest interessant is; waarom bijvoorbeeld de computer voor de rokade kiest. Om dat te zien moeten we naar een hoger abstractieniveau, het zgn. perspectief van het ontwerp.

Vanuit de bij dit **ontwerpperspectief** behorende houding bekommeren we ons niet om de innerlijke werking van het systeem. We voorspellen het gedrag louter op basis van de aanname dat het systeem een zeker ontwerp heeft, en zich overeenkomstig dit ontwerp zal gedragen. Computergebruikers hanteren vaak dit perspectief. Doorgaans hebben ze geen flauw benul welke fysische processen verantwoordelijk zijn voor de hoge betrouwbaarheid en voorspelbaarheid van het gedrag van hun computer. Wel hebben ze een goed idee van hetgeen de computer ontworpen is te doen, en voorspellen aan de hand daarvan het gedrag onder de aanname dat er zich geen storingen voordoen³². In het geval van het voorbeeld van de schaakcomputer kun je, als je daar de beschikking over hebt, naar de opzet van het schaakprogramma kijken en vind je vermoedelijk ook een procedure voor het bepalen van de opportuniteit van rokades. Met enige inspanning (maar toch beduidend minder dan op het fysische niveau) kun je a.h.v. de programmaroutines voorspellen wat de volgende computerzet zal zijn.

Maar wil je de grote lijnen ontwaren en beschik je niet over het programma-ontwerp, of is het systeem te complex om er vanuit de lagere abstractieniveaus effectief mee om te gaan, dan hanteer je het hoogste abstractieniveau: het **intentioneel perspectief**, waar je zicht op krijgt vanuit een door jou aangenomen intentionele houding; *“Hoe werkt het? Allereerst houd je het object waarvan je het gedrag wil voorspellen voor een rationele actor; dan probeer je, gegeven zijn plaats in de wereld, uit te zoeken wat die actor behoort te geloven en verlangen. [...]Vervolgens zal een beetje praktisch redeneren vanuit de gekozen verzameling van overtuigingen en verlangens jou in veel gevallen een beslissing opleveren over wat de actor behoort te doen. Dat is dan hetgeen wat jij voorspelt dat deze zal doen.[Dennett³³].”* Zoals we vanuit het ontwerp-perspectief het gedrag aan de hand van het veronderstelde ontwerp voorspellen, doen we dat ook vanuit het intentioneel perspectief, maar met dat verschil dat we nu overtuigingen, verlangens en andere intentionaliteiten aan het systeem toeschrijven. Overtuigingen die het systeem behoort te hebben gegeven haar perceptuele vermogens, epistemische behoeften en haar biografie. Verlangens die het systeem behoort te hebben gegeven haar biologische behoeften en de meest praktische middelen om ze te bevredigen. Dus (natuurlijk geëvolueerde) intentionele systemen verlangen overleving en voortplanting, en derhalve voedsel, veiligheid, gezondheid, seks, weelde, macht, invloed, etcetera. Zou je zo'n verzameling van overtuigingen, verlangens en bijbehorende gedragsdisposities in kaart brengen dan ontstaat er een normatieve theorie van rationaliteit, met daarin toegeschreven mentale 'toestanden' die de rol vervullen van calculerende devices (zeg theoretische postulaten). Zo'n theorie noemen we 'volkpsychologie'. De praktijk leert ons dat we t.b.v. gedragsvoorspelling intenties niet alleen aan mensen maar ook aan objecten toeschrijven. Vanuit het intentioneel perspectief beschrijven en voorspellen we het gedrag van de schaakcomputer in termen van zijn spelbegrip, bedoelingen en strategieën. Dit perspectief maakt ons grotendeels blind voor het programmaontwerp, en stekeblind voor de onderliggende fysica, maar is overigens uiterst nuttig. Het toont ons de logica van het systeem, de grote lijnen, de achterliggende wensen en ideeën. Het openbaart ons patronen in het systeemgedrag die ons anders volledig zouden ontgaan.

3.1.3 Realisme en interpretationisme

Het door Dennett geschetste beeld van de intentionele strategie lijkt te impliceren dat wanneer we in de volkpsychologie gebruikmaken van intentionele mentale termen ('geloven', 'verlangen' etc.) we daarmee niet naar innerlijke toestanden verwijzen. Dat is ook zo want volgens Dennett spelen de toegeschreven 'toestanden' feitelijk de rol van een geïdealiseerde fictie in een handeling-voorspellende en verklarende berekening³⁴.

Dennett ontmoette bij de introductie van zijn theorie veel kritiek. Het alleen maar toeschrijven van intentionaliteit omdat het een handige manier is om gedrag te voorspellen, doet de intuïtie dat overtuigingen en verlangens ook als subjectieve ervaringen bestaan, geweld aan. Zoals Dennett het stelt lijkt het dat wanneer we op een psychologische manier naar mensen kijken, we geen echte overtuigingen en verlangens aan ze toeschrijven. We doen alleen maar alsof ze die hebben. *“In Dennett's theorie bestaan de mentale toestanden slechts in de 'eye of the beholder'[Nozick³⁵]”*. In 'Real Patterns' heeft Dennett een antwoord op dit probleem geformuleerd. Hij ontkent daarin dat geloven of verlangen geen objectieve status zou hebben. Hij betoogt dat geloven, verlangen en andere intentionaliteiten hun

objectiviteit ontlenen aan (vanuit het intentionele perspectief herkenbare) gedragspatronen van intentionele systemen. Dus ook al is het niet gerechtvaardigd af te leiden dat deze patronen neurofysiologisch zijn gespiegeld, is er toch wel degelijk iets werkelijks en objectiefs aan zoiets als overtuigingen en verlangens.

3.1.4 Interpretationisme in relatie tot reductionisme en eliminativisme

Kunnen we volgens Dennett nu de beschrijvingen van de verschillende perspectieven tot elkaar reduceren? Het antwoord lijkt om meerdere redenen ontkennend te zijn. Allereerst verzet Dennett zich tegen de stelligheid waarmee identiteit-theoretici en functionalisten de visie verdedigen dat *“er voor elk toegeschreven geloof een functionele interne toestand van de machinerie bestaat, die net zo in functionele delen is te ontbinden als de geloofuitdrukking in haar delen (woorden en termen) kan worden ontbonden”*³⁶. Dennett meent dat zij die dit denken twee claims door elkaar halen. De eerste claim is dat beschrijvingen vanuit de intentionele houding over objectief bestaande gedragspatronen in de wereld gaan. De tweede claim stelt dat de gedragspatronen in de wereld worden geproduceerd door een ander werkelijk ruwweg isomorf patroon in de breinen van intelligente wezens. Die tweede claim volgt volgens Dennett niet uit de eerste claim, en is ook empirisch niet te rechtvaardigen. Kortom, volgens Dennett is de toeschrijving van het intentionele mentale te herleiden tot gedragspatronen, maar zijn de zgn. mentale ‘toestanden’ niet te reduceren tot isomorfe patronen in het brein.

Een tweede argument tegen reductionisme is dat de gedragspatronen alleen vanuit het intentionele perspectief zijn te herkennen. De verschillende perspectieven bedienen zich ook van een eigen taal met begrippen waarvan betekenissen exclusief verbonden zijn met die context. Het zou dus getuigen van een denkfout wanneer men het ene tot het andere perspectief tracht te herleiden³⁷.

En hoe zit het met Dennett’s positie t.a.v. eliminativisme van Churchland? Churchland kan weliswaar het vermogen van de intentionele houding als een strategie van voorspelling waarderen, maar hij denkt dat het gedoemd is om uiteindelijk te verdwijnen, omdat er langs de neurowetenschappelijke weg een patroon wordt ontdekt dat zo superieur is aan het ruisachtige patroon van de volkspychologie, dat iedereen bereid is de eerste voor het laatste te verruilen. Dennett’s antwoord is dat zoiets in theorie zou kunnen gebeuren. Maar Churchland moet dan wel de winst laten zien van het weinig attractieve uitruilen van de eenvoud waarmee wij mensen de volkspychologie toch betrekkelijk succesvol kunnen gebruiken, tegen immuniteit voor fouten van een zeer gecompliceerde en tijdrovende benadering. *“Een waarlijk ‘general-purpose’, robuust systeem van patroonbeschrijving, waardevoller dan de intentionele houding, is niet een onmogelijkheid, maar ik zou er niet op gokken!” [Dennett³⁸]*

3.1.5 Commentaar op Dennett’s interpretationisme

In het volgende hoofdstuk vergelijk ik in meer detail het antireductionisme van Dennett met dat van Habermas. Een eerste impressie laat zien dat Dennett’s interpretationisme voor verschillende perspectieven en de volkspychologie plaats en gezag inruimt, maar dat elk perspectief wel door een objectivering gekenmerkt wordt. We kunnen evenzogoed mentale toestanden en intentionaliteit aan objecten als aan subjecten toeschrijven omdat er in soort

geen wezenlijk verschil tussen bijvoorbeeld het rationele gedrag van een schaakcomputer en van een schaker van vlees en bloed bestaat.

In Dennett's bewustzijnstheorie komt heel duidelijk het objectiverende derde-persoonsperspectief als de enige cognitieve toegang tot de wereld tot uitdrukking. Hij tracht een weerlegging te geven van de idee van het Cartesiaanse ego als innerlijk geestesoog, dus als private, geprivilegieerde toegang tot het zelf.

3.2 Dennett's visie op bewustzijn en zijn strijd tegen het Cartesianisme

Descartes claimde dat het bewustzijn kon worden verklaard door een niet-materiële ziel, het Cartesiaanse ego, die een representatie van de wereld in de pijnappelklier gadeslaat. In deze notie speelde het Cartesiaanse ego de rol van een homunculus die over een sturende wilskracht beschikt. Het Cartesiaanse dualisme is inmiddels door wetenschappers en filosofen goeddeels ingeruild voor een materialistische visie waarin mentale gebeurtenissen op louter neurofysiologische breinprocessen terug te voeren zijn. Maar van het beeld van een Cartesiaans theater (een plek in het brein van waaruit alles gebeurt, waar ik zit) doen we maar moeilijk afstand. Deze visie op een te lokaliseren fysische homunculus als auteur van ons handelen en als betekenisgever van onze waarneming is echter incoherent³⁹, en levert in geval het als verklaringsmodel wordt gehanteerd, een verkeerde beoordeling van neurowetenschappelijke onderzoeksresultaten. Zo wordt herhaaldelijk, met het Cartesiaanse model in de hand, uit de resultaten van het roemruchte Libet experiment geconcludeerd dat het bewustzijn een epifenomeen moet zijn. Volgens één interpretatie van het Libet-experiment zou in het brein van proefpersonen een beslissing voor een handeling, voorafgaande aan een bewust besluit 'nu' te handelen, al in een onbewust cognitief proces zijn ingezet. Maar hoe zit het dan wel volgens Dennett?

Toen gedurende de evolutie aan de sociale interactie en de communicatie steeds grotere eisen werden gesteld, heeft zich een soort op de taal gebaseerde gebruikersinterface tussen sprekende mensen ontwikkeld. Met deze interface zijn mensen elkanders gedragscategorieën gaan ontdekken, en konden ze de ander dingen laten doen (door ze toe te spreken), met ze in wisselwerking treden, zonder dat ze iets van de werking van het brein hoefde te weten. De belangrijkste metafoor van deze co-ontwikkelde gebruikersillusies is het zelf, dat op een plek in het brein lijkt te zetelen, het cartesiaans theater, dat een beperkte metaforische kijk⁴⁰ geeft op wat er in ons brein omgaat. Een van de belangrijkste functies van de gebruiksiillusie die het brein van zichzelf heeft, en die Dennett het 'zelf als narratief zwaartepunt' (center of narrative gravity) noemt, is het zelf van het heden een middel te verschaffen om met het zelf op andere momenten te communiceren.

Dennett benadrukt dat het hier om een illusie gaat en niets over de werkelijke werking van het brein vertelt. De idee dat we door middel van een eerste-persoonsperspectief (een innerlijk geestesoog) een geprivilegieerde toegang tot ons innerlijk hebben is een illusie die ons door de metafoor van de cartesiaanse ego wordt verschaft. Dat ik meer van mijzelf weet dan ieder ander van mij, is niet omdat ik een bijzondere toegang tot mijn innerlijk heb, maar is alleen maar te danken aan het feit dat ik met mijzelf de meeste tijd doorbreng.

In 'Consciousness explained' heeft Dennett voorgesteld om ten behoeve van de verklaring van de werking van de hersenen en het fenomeen bewustzijn, het model van het Cartesiaanse theater te vervangen door het plausibelere "Multiple Drafts". Volgens dit MP-model bestaat in werkelijkheid het brein uit een enorme hoeveelheid aan simultaan optredende

gedistribueerde, competitieve, neurale, parallel verlopende processen (bijvoorbeeld gedachten); processen van interpretatie en bewerking van de zintuiglijke signalen, die schetsen (drafts) opleveren die voortdurend bloot staan aan revisie. Het proces wat tijdelijk de wedstrijd wint domineert de hersenactiviteit (en verwordt tot bewustzijnsinhoud).

Er is dus niet zoals in het Cartesiaanse theater sprake van een opzichter, één enkele homunculus (“author and meaner”) die regelt welke proceseenheid nu iets mag doen, maar van een grote groep eenvoudige homunculi die in onderlinge strijd verwickeld zijn om de macht over het brein te verkrijgen.

Om aan te tonen dat het bewustzijn wel een causale rol speelt geeft Dennett de volgende verklaring. Naast dat we tegen anderen praten, praten we als we vrijwillige handelingen uitvoeren ook tegen onszelf. Nu wordt een mentale inhoud niet bewust door een of andere speciale kamer in het brein binnen te komen, niet door omgevormd te worden in een of ander geprivilegieerd en mysterieus medium, maar door de wedstrijd om de dominantie van de gedragscontrole van andere mentale inhoud te winnen, teneinde op die manier langdurige effecten te bereiken.

“En aangezien we praters zijn, en omdat tegen onszelf praten een van ons meest invloedrijke activiteiten is, is het een van de meest effectieve manieren voor een mentale inhoud om invloed te krijgen, een positie in te nemen om de taalgebruikende delen van de controle te besturen. Dit alles moet in de arena van het brein gebeuren, in de ‘centrale verwerkingseenheid’; maar niet onder leiding van iets.” [Dennett⁴¹]

Doordat we dus tegen onszelf gaan praten, maken we dat tot de dominante hersenactiviteit, verwordt het tot inhoud van ons bewustzijn, en kunnen we ons als het ware zelf opdrachten geven. Het is een van de wijzen waarop we dus controle over onszelf hebben. Je kunt je afvragen of hier nu het bewustzijn mee is verklaard? Blijft er toch nog niet iets te verklaren over? Want wie of wat praat er dan hier tegen jezelf? Wie of wat geeft die opdracht om te praten en voert dat uit? Dat kan geen Cartesiaans ego zijn. Dat moet dus de rede zelf zijn.

3.3 Dennett’s kijk op naturalisme

Dennett’s benadering is naturalistisch en valt te typeren als antireductionistisch fysicalisme⁴². In zijn filosofie zijn kenmerken te herkennen van het behaviorisme, het functionalisme en het computationalisme. Het behaviorisme komt tot uitdrukking in de verwerping van de idee dat mentale termen verwijzingen zouden zijn naar, door introspectie geopenbaarde, innerlijke episodes. Mentale concepten zijn afgeleid uit louter publiekelijk waarneembare gedragspatronen. Dat Dennett meent dat beschrijvingen van mentale toestanden niet uitsluitend gereduceerd kunnen worden tot beschrijvingen van gedragspatronen, maar dat ze beschreven moeten worden in termen van causale relaties met andere mentale toestanden en stimuli en responses, maakt zijn opvatting functionalistisch.

Beschrijvingen van mentale toestanden kunnen ook niet gereduceerd worden tot neurofysiologische processen omdat er simpelweg geen één-op-één relatie tussen bestaat. Dus het mentale is weliswaar afhankelijk van, en wordt gevormd door het neuronale substraat, maar zonder dat er sprake is van een reductionisme zoals in het fysicalisme (materialisme) van de type-type identiteitstheorie. Zijn antireductionisme is zowel conceptueel als epistemologisch. Conceptueel omdat de begrippen voor de mentale toestanden niet kunnen

worden gereduceerd tot begrippen van een neurofysiologische theorie. Epistemologisch omdat de volkpsychologie een andere type van kennis is dan de kennis van neurofysiologische processen.

Dennett's computationalisme komt tot uitdrukking in zijn visie dat breinprocessen een symboolmanipulerend principe volgen, en dus evenzogoed in een ander geschikt medium (zoals silicium) zou kunnen worden uitgevoerd.

Hoewel Dennett erkent dat er verschillende manieren zijn om naar de wereld te kijken, is en blijft voor hem de allesomvattende natuur toch de natuur van de natuurwetenschap.

“Mijn fundamentele perspectief is naturalisme, de idee dat filosofisch onderzoek samengaat met, en niet superieur is of vooraf gaat aan, het natuurwetenschappelijk zoeken naar de waarheid.[...] De schat aan welverdiende wetenschappelijke ontdekkingen en theorieën moet worden verwelkomd als grondstof voor filosofisch theoretiseren, zodat goed onderlegde en constructieve kritiek van zowel de wetenschap als de filosofie mogelijk is.” [Dennett⁴³]

Dennett toont zich dus, in tegenstelling tot Habermas, tegenstander van een eerste filosofie, één die aan wetenschappelijk onderzoek voorafgaat. Filosofen hebben niet méér cognitieve toegang tot de werkelijkheid dan wetenschappers. De taak van filosofen is het expliciet maken van de impliciete normen en methoden van wetenschappelijke praktijken, en de vaak tegenstrijdige standpunten verhelderen en verenigen tot één enkele visie op het universum.

4 Zwak naturalisme versus antireductionistische fysicalisme

Een vergelijking tussen het zwak naturalisme van Habermas en het antireductionistisch fysicalisme van Dennett levert een aantal overeenkomsten maar ook een belangrijk verschil op. Overeenkomsten zijn; de erkenning dat de cognitieve toegang tot de wereld via een interpretatiestructuur verloopt, de toepassing van een interpretationisme, en de verwerping van reductionisme en Cartesianisme. Het verschil zit 'm in epistemische status die aan het objectiverend perspectief wordt toegekend.

Laten we beginnen met de overeenkomst in hun verwerping van het Cartesianisme, de klassieke filosofie van het bewustzijn die gedomineerd wordt door het onderscheid tussen het eerste-, en derde-persoonsperspectief. Dennett ontmaskert de illusie van een intern geestesoog en houdt alleen het objectiverende subject-object perspectief in stand. Voor Dennett maakt het niet uit of je een subject of object tegenover je hebt. Zo kun je evenzogoed aan een computer als aan een mens intentionele toestanden toeschrijven. Daarmee verstoffelijkt Dennett het subject tot object. Dat is natuurlijk ook in lijn met zijn verdediging van het fysicalisme. Zo is het menselijk brein en het bewustzijn bijzonder, maar toch niet zo bijzonder dat het niet uiteindelijk een naturalistische (antireductionistisch fysicalistische) verklaring heeft. Dus de mens is als al het andere in de natuur louter samengesteld uit de door de fysica beschreven bouwstenen. Al diens eigenschappen zijn in principe afleidbaar uit de fysische eigenschappen van de elementaire bouwstenen en de configuratieve relaties die er tussen bestaan. Daarmee is voor Dennett de allesomvattend natuur hetzelfde als de natuur van de natuurwetenschap. Het ontologisch, en nomologisch fysicalisme is voor Dennett een feit.

Ook Habermas verwerpt het cartesianisme, de klassieke opvatting van eerste- en derde persoonsperspectief, die hij resp. omschrijft als “de introspectie of zelfobservatie waarvan het object de eigen representatie is, en de observatie van een individu als hij zich vanuit een

objectiverende houding tot de objecten wendt.” Hij vervangt dit door een dualisme van tweede- en derde-persoonsperspectieven – door een subject-subject- en subject-object perspectief - dus door de interpretatieve prestaties van een participant in een ideale spreesituatie, en een waarnemerperceptie van een object.

Hoe ziet een onderlinge vergelijking tussen de perspectieven van de theorieën van Habermas en Dennett eruit? Het ‘objectiverende perspectief’ van Habermas, het waarnemerperspectief dat het beeld van objecten in nomologische causale relatie constitueert, komt goeddeels overeen met Dennett’s fysische perspectief. De beschrijvingen en verklaringen van dit perspectief zijn gesteld in de taal en theorie van de natuurkundige, de bioloog of neurowetenschapper.

Het perspectief waarin de mens als intentioneel, naar redenen handelend wezen verschijnt, is voor Habermas het participierend- voor Dennett het intentionele perspectief. De rationaliteit wordt zichtbaar vanuit Habermas’participantperspectief als die van het communicatief handelen, vanuit Dennett’s intentionele perspectief als volkpsychologie.

Zowel Habermas als Dennett verwerpen de idee dat beschrijvingen in mentale termen kunnen worden gereduceerd tot die van neurofysiologische processen. Maar ze doen dat elk op een andere grond.

Voor Habermas is het niet mogelijk om de normatieve socio-culturele praktijken, zonder residu in het jargon van het objectiverende waarnemerperspectief uit te drukken. De know-how van die praktijken kan alleen vanuit het participierend perspectief op het communicatief handelen performatief, intuïtief worden gevat. Breintoestanden zouden we niet begrijpen omdat ze ontdaan van hun interpretatieve context geen semantische inhoud hebben.

“We kunnen niet de cultuur en de gemeenschap als een context van communicatie vanuit het ‘perspectief van het brein’ begrijpen omdat er geen visie vanuit ‘nowhere’ is – er is geen observatie onafhankelijk van voorafgaande participatie in een of andere soort van communicatie.”[Habermas⁴⁴]

Een belangrijk verschil met Habermas is dat intentionaliteit voor Dennett wel iets is wat principieel tot een objectiverende perspectief is te herleiden. Het is op grond van de herkenning van de gedragspatronen van een subject of object dat de waarnemer vanuit diens ingenomen intentionele perspectief, er mentale toestanden aan toeschrijft. Het intentionele, en fysische perspectief van Dennett zijn in de betekenis die Habermas er aan geeft, beiden objectiverende waarnemerperspectieven. Het intentionele perspectief is uitdrukkelijk louter een bepaalde interpretatie van structuren in een fysische ontologie en nomologie. Voor Dennett, hoewel in principe niet onmogelijk, is het alleen door onze beperkte intelligentie en het gebrek aan toereikende gegevens dat een beschrijving vanuit het fysische perspectief niet altijd mogelijk is. Daarom is voor hem een verklaring op het laagste niveau, het fysisch perspectief, niet altijd de beste manier om (complexe) dingen te verklaren of te begrijpen.

Desalniettemin is een reductie van de beschrijvingen vanuit het intentionele perspectief (in termen van toegeschreven mentale toestanden) tot die van het fysische perspectief (in neurofysiologische termen) volgens Dennett niet mogelijk omdat de patronen die we in iemands gedrag herkennen niet exact in de neurofysiologische patronen in diens brein gespiegeld behoeven te zijn.

Het gegeven dat in Dennett’s optiek intentionaliteit ‘open to the eye’ en objectiverend is, heeft als consequentie dat we een uitputtende beschrijving van intentionaliteit vanuit het perspectief van een waarnemer kunnen geven zonder een beroep op zoiets als het

participantperspectief van Habermas te moeten doen. Voor Habermas is het participantperspectief noodzakelijk omdat alleen van daaruit het (inter)subjectieve aspect tot uitdrukking kan komen, en de betekenis voor mij, (de betekenis-als-relevantie) zich openbaart.

Het subjectieve aspect is in de filosofie van Dennett echter ontdaan van enig epistemologisch gezag. Voor hem berust het eerste-persoonsperspectief, opgevat als geestesoog, op een illusie, een construct van een talige interface. Alle uitspraken in onze taal berusten uiteindelijk op publiekelijke waarneembaarheden (hetzij direct hetzij met technische hulpmiddelen als brillen, telescopen, meetapparatuur), op abstracta of op ficties. Er worden in betekenisvolle uitspraken geen verwijzingen naar innerlijke processen verricht die nimmer publiekelijk waarneembaar zijn. Wittgenstein toonde reeds aan dat de idee van het spreken van een privé taal waarin naar innerlijke gebeurtenissen verwezen wordt geen hout snijdt. Taal en betekenis komt inderdaad intersubjectief tot stand. Maar volgens Dennett louter en alleen aan de hand van hetgeen publiekelijk waarneembaar is.

Maar waar Dennett het subjectieve perspectief behorende aan het interne geestesoog heeft geëlimineerd, heeft Habermas een verbuiging gemaakt door het onderscheid eerste-, derde-persoonsperspectief in te ruilen voor het tweede-, en derde-persoonsperspectief. Dennett heeft het belang van het subjectieve aspect uitgehold, waar Habermas het in het tweede-persoonsperspectief heeft ondergebracht. In zijn onderbouwing heet het dat het participantperspectief uitzicht op die concrete-ander geeft. Ik leer pas de betekenis van jouw taaldaden kennen, als ik weet welke argumenten jij voor de rechtvaardiging ervan geeft. Daarmee kan ik me inleven in jouw perspectief. Habermas onderbouwt dit niet direct met een verwijzing naar een door introspectie gekend aspect want het eerste-persoonsperspectief als intern geestesoog behoort aan de door hem verworpen klassieke Cartesiaanse filosofie van het bewustzijn. Maar de onderbouwing waarin er een subjectief aspect wordt uitgewisseld, kan in de communicatietheorie van Habermas wel getraceerd worden. Want als we actief participeren in een discours en argumenten uitwisselen, dan stemmen we de aan onszelf toebehorende interpretatiestructuur, die uniek en particulier voor ieder van ons is, gevormd door onze eigen unieke historie, op die van de concrete ander af. Zo komt dan de betekenis-als-relevantie-voor-mij alsnog in ieder geval in de consensus die we bereiken tot uitdrukking. Laten we eens kijken hoe dit er uitziet als we het vertalen naar de ideale spreek situatie van het performatieve communicatieve handelen. De betekenis(-als-relevantie) van concrete taalhandelingen wordt louter vanuit het participerende perspectief van de discours geopenbaard in een rechtvaardigende beargumentering. Volgens zijn communicatietheorie zijn bij elke uitspraak in de leefwereld drie aanspraken op geldigheid betrokken: een aanspraak op waarheid betreffende de objectieve wereld van de dingen en gebeurtenissen, een aanspraak op juistheid betreffende de normatieve wereld van erkende normen en een aanspraak op oprechtheid betreffende de subjectieve wereld van gevoelens en verlangens. Door een rechtvaardiging in argumenten te geven, komen volgens de opvatting van Habermas ook de subjectieve aspecten van de spreker tot uitdrukking⁴⁵.

Dennett's respons hierop zou als volgt kunnen zijn. Om te beginnen lijkt het erop dat in Habermas' theorie bij aanspraken op oprechtheid (betreffende de subjectieve wereld van gevoelens en verlangens) een kennisbron in de vorm van een innerlijk geestesoog wordt verondersteld. Maar dit vooronderstelt een soort van cartesiaanse ego, een ik-zelf die in een Cartesiaans theater een wereld gadeslaat. Maar het zelf is niet een bestaande entiteit zoals een oog wat naar binnen kijkt. Het zelf, een narratief zwaartepunt, is een handig abstract gegeven,

een gebruikersillusie die een plek geeft van waaruit verantwoordelijkheid wordt genomen. Dennett's terughoudendheid in een beroep op zo iets als de subjectieve ervaring in de rechtvaardiging van de betekenis van talige uitspraken is ook duidelijk herkenbaar in zijn benadering van het gebruik van termen voor mentale toestanden.

Vanuit een door ons aangenomen intentionele houding schrijven we aan onszelf intentionaliteit toe zoals we dat aan anderen toeschrijven, op grond van publiekelijk waarneembare gedragspatronen, en niet door met een innerlijk geestes oog ons binnenwerk te aanschouwen. We passen deze strategie toe omdat het ons helpt het gedrag van deze complexe entiteiten beter te begrijpen. Deze mentale termen zijn net zo reëel als een Newtoniaans zwaartepunt, handig om mee te rekenen, en gedragspatronen te voorspellen, maar verwijzen niet naar innerlijke toestanden of processen. Dus uitspraken aangaande een vermeende subjectieve wereld zijn volgens Dennett terug te voeren op uitspraken aangaande het gedrag van objecten in de objectieve wereld. Voor Dennett is kennis alleen gezaghebbend als deze tot de publiekelijk zintuiglijke waarneming is te herleiden. Aanspraken op geldigheid van uitspraken kunnen derhalve alleen refereren aan het subject-object perspectief.

De betekenis van de termen betreffende Habermas normatieve aanspraken op geldigheid van een taaldaad onderscheiden zich voor Dennett ook niet van het objectiverende perspectief, daar de conventies betreffende moraliteit te herleiden zijn tot waarneembaar gedrag, of als het om de ultieme bron van de moraal gaat, om publiekelijk zintuiglijk waarneembare gedragsdisposities behorende aan de natuurlijke drijfveren zoals passies en emoties. [Hierop kom ik in sectie 6.4 terug wanneer ik de moraaltheorie van Habermas en Dennett bespreek.]

Nu is het ook niet zo dat Dennett de beschrijving van een 'subjectief' perspectief, zoals bijvoorbeeld het fenomenaal bewustzijn, volledig negeert. In contrast met de fenomenologie (welke hij autofenomenologie noemt) gebruikt hij de naam 'heterofenomenologie' voor een expliciet wetenschappelijke derde-persoonsbenadering van de studie van het bewustzijn. Het bestaat uit het toepassen van de wetenschappelijke methode met een antropologische neiging, die de zelfrapportage van subjecten combineert met al het andere mogelijke bewijs om hun mentale toestanden vast stellen. Het doel is om te ontdekken hoe subjecten de wereld en zichzelf zien, zonder de accuratesse van de visie voor lief te nemen. In tegenstelling tot de fenomenologie is voor Dennett heterofenomenologie alleen gezaghebbend met betrekking tot hoe dingen voor subjecten lijken te zijn. Heterofenomenologie verlangt van ons om te luisteren naar het subject, diens rapportage serieus te nemen, maar tegelijkertijd te blijven kijken naar al de andere ons ter beschikking staande (derde persoon) gegevens, om soms te moeten concluderen dat het subject het fout heeft. De sleutelrol van de heterofenomenologie in Dennett's filosofie van het bewustzijn is dat het alles definieert wat over de 'mind' kan of moet worden gekend.

“De totale verzameling van details van de heterofenomenologie, plus al de data die we kunnen verzamelen van ermee samenvallende gebeurtenissen in het brein van de subjecten en in hun omgeving, omvat de totale dataverzameling voor een theorie van het menselijke bewustzijn. Het laat noch objectieve noch subjectieve fenomenen van het bewustzijn buiten beschouwing.” [Dennett⁴⁶].

Het objectiverende derde-persoonsperspectief, zo volgt uit de filosofie van Dennett, heeft het hoogste gezag. Al zou het linguïstisch bemiddeld zijn, het is toch onze enige toegang tot de wereld, inclusief onszelf. Natuurlijk zal Dennett niet kunnen toegeven dat er nog een eerste- of tweede- persoonsperspectief is, omdat dan de wereld vanuit het derde persoonsperspectief

niet zonder residu te beschrijven valt. Hij is al decennia lang bezig zijn opvattingen te verdedigen tegen horde critici die allen menen dat Dennett in zijn beschrijving iets cruciaals als bijvoorbeeld de Quale, 'what's it like for me', (de gevoelservaring en beleving) veronachtzaamd. Het lijkt me overigens best te verdedigen dat een wereld die volledig uit fysische entiteiten bestaat, aan de natuurwetten gehoorzaamt, een ervaringscomponent levert die met nog zoveel taal kan worden omschreven, maar nooit de essentie ervan volledig kan pakken, tenzij men volledig samenvalt met het wezen wat ervaart.

Enfin, wat zou Habermas op die kritiek van Dennett kunnen antwoorden? Habermas zou even afgezien van Dennett's veronachtzaming van het subjectieve aspect, kunnen tegenwerpen dat Dennett de verstrengeling van het waarnemerperspectief met het participantperspectief over het hoofd ziet. De toegang tot de wereld vanuit het perspectief van de waarnemer is verwrongen en toont slechts een segment van de werkelijkheid erachter. De opvatting dat alles in termen van causaliteit zou kunnen worden verklaard behoort slechts toe aan de instrumentele rationaliteit die het objectiverend beeld constitueert.

Dennett miskent dat andere segment, waar het participerende perspectief zicht op geeft, de communicatieve rationaliteit, die ons in staat stelt door middel van deliberatieve instemming de geldigheid van onze taaldaden te vestigen, en zo de mogelijksvoorwaarden levert voor de (taalbemiddelde) cognitieve toegang van het waarnemerperspectief. Circulariteit van redenering verhindert dat het objectieve perspectief zijn eigen rechtvaardiging vestigt. Dat moet plaatsvinden door middel van de communicatieve rationaliteit waarvan het inzicht niet op causale verklaringen maar op hermeneutisch verstaan is gebaseerd.

Habermas zal Dennett tegenwerpen dat betekenis van een objectiverende uitspraak niet louter tot stand komt door een correspondentie met een stand van zaken in de werkelijkheid, maar dat hier noodzakelijk de performatieve praktijk van een intersubjectieve vestiging van waarheid en betekenis-als-relevantie voor allen op is betrokken.

Tot zover deze verschillen tussen de visies van Habermas en Dennett. Laten we terugkeren naar de overeenstemmingen en kijken of die voldoende grond leveren om aspecten van Dennett's opvattingen in de filosofie van Habermas in te passen. Dennett's filosofie geeft een rechtvaardiging voor het epistemisch gezag wat aan de volkspychologie kan worden toegekend als redelijk betrouwbaar en handig te hanteren instrument om het gedrag van anderen (en jezelf) te verklaren en beheersen. Tegelijkertijd worden dan de impliciet in de volkspychologie opgenomen socio-culturele normen geaccepteerd dat we van elkaar verwachten dat we naar redenen handelen. Dit opent de weg naar de zienswijze dat rationeel en hermeneutisch verstaan een rol krijgt toebedeeld in het rechtvaardigen van het handelen. Het bezwaar dat als ons handelen ook verklaarbaar is vanuit het 'objectiverende' perspectief van de wetmatige causaliteit van de neurowetenschap, het semantisch geladen hermeneutische verstaan overtollig wordt, heeft geen grip op het antireductionisme van Dennett's opvatting betreffende de relatie tussen het intentionele en het fysische perspectief. Er is immers volgens Dennett geen één-op-één relatie tussen de patronen behorende bij de mentale toestanden van het intentionele perspectief, en de vanuit het fysisch perspectief zichtbare neurofysiologische patronen in het brein. Dus het antireductionisme en de erkenning van de volkspychologie en daarmee de veronderstelling dat mensen worden bewogen door redenen in de filosofie van Dennett, spoort met diezelfde aspecten van de filosofie van Habermas.

Het cruciale verschil tussen de opvattingen van Habermas en Dennett is en blijft het verschil in gezag dat wordt toegekend aan het perspectief van de waarnemer. Hoewel Dennett een

onderverdeling in drie perspectieven maakt, zijn het toch alle drie in Habermas'zin waarnemerperspectieven, niet performatief. Ook het intentionele perspectief is objectiverend van aard. De door Dennett aan het objectiverende perspectief toegekende prioriteit maakt daarom veel van zijn filosofische opvattingen voor Habermas onacceptabel. Omgekeerd zal het moeilijk voor Dennett te accepteren zijn dat het participantperspectief iets essentieels toevoegt wat ten principale niet vanuit het objectiverende perspectief is in te vangen.

Deze tussen de opvattingen van Habermas en Dennett lastig te overbruggen gebleken gapende kloof kan in historisch perspectief in de vorm van een verschil in visies van twee filosofische hoofdstromingen worden geplaatst. De ene stroming wordt gevormd door de idealisten en rationalisten langs de lijn die via Heidegger, Husserl, Freud, Marx, Hegel en Kant terugvoert naar Plato. Zij menen (vanuit een filosofische van binnen naar buiten benadering) dat de toegang tot de wereld immer aan ons verschijnt onder de aanname dat de constructie van ons denken er een belangrijke rol in speelt. In wat we kennen heeft het (inter)subjectieve aspect een aandeel wat niet mag worden veronachtzaamd. Kennis heeft betrekking op hoe de dingen door ons ervaren worden, op de verschijning, op datgene wat er voor ons mensen toe doet.

Dennett's filosofie is echter geworteld in de traditie van een empiristische lijn gevormd door filosofen als Quine, Ryle, Carnap, Russell, Hume, helemaal terug naar Aristoteles. Zij menen (vanuit een filosofische van buiten naar binnen benadering) dat daarbuiten een wereld min of meer klaar ligt om (desnoods conceptueel bemiddelt) vanuit het objectieve perspectief ontdekt te worden. Alleen datgene wat publiekelijk waarneembaar is, vormt een valide bron van kennis. Al het andere is speculatief. De natuurwetenschap die het onderzoek van de publiekelijk zintuiglijk waarneembare wereld tot haar domein heeft gemaakt, verdient derhalve het hoogste woord. Deze opvatting ligt aan de basis van het materialisme en fysicalisme. Storende (inter)subjectieve invloeden dienen te worden geëlimineerd. Kennis heeft betrekking op een correspondentie met de feiten in de wereld daarbuiten.

Dit is wat sterk uitvergroott en moet niet blind maken voor het gemeenschappelijke wat de filosofieën van Habermas en Dennett delen. Behalve dat ze beiden een interpretationisme hanteren treffen we niet alleen bij Habermas maar ook bij Dennett pragmatische elementen aan. Denk aan Dennett's pragmatische opvatting dat de volkpsychologie correct is omdat deze werkt. En belangrijk om daarbij te constateren is dat fysicalisme niet per se hoeft te impliceren dat de mens zich in het geheel niet zou kunnen herkennen in een objectivistisch geschetst perspectief. Dennett toont immers dat er een coherente filosofie mogelijk is waarin een ontologisch, en nomologisch fysicalisme wordt gepaard aan een antireductie van volkpsychologische, tot neurofysiologische beschrijvingen, en waarin het beeld van mensen die worden bewogen door redenen kan worden verdedigd.

4.1 Sterke en zwakke punten van interpretationisme van Dennett en Habermas

Habermas erkenning van het epistemisch gezag van het participerend perspectief wat ook rekening houdt met (inter)subjectieve aspecten, doet recht aan de intuïtie dat ze er toe doen. De idee van Dennett dat het perspectief van de waarnemer voldoende is om bijvoorbeeld "bewustzijn net zo diep en volledig te kunnen verklaren zoals natuurlijke fenomenen als metabolisme, reproductie, licht, zwaartekracht, etc⁴⁷", blijft ergens ook wringen. De idee dat een innerlijk geestesoog een illusie is lijkt contra-intuïtief. Natuurlijk moet men terughoudend zijn om de intuïtie klakkeloos als leidraad voor de waarheid (opgevat als

correspondentie met een stand van zaken in de werkelijkheid) te gebruiken. De geschiedenis toont van meerdere intuïties dat die uiteindelijk door de wetenschap als een onjuiste voorstelling van de werkelijke stand van zaken zijn ontmaskerd. Maar als je zoals Habermas de definitie van waarheid in termen van correspondentie loslaat, dan ligt het meer voor de hand om de fenomenale verschijning en intuïtie als een kennisbron te accepteren. Habermas' oplossing om het participant- en het waarnemerperspectief op een metatheoretisch niveau te verenigen met een ontologisch monisme compartimenteert echter de natuurwetenschap en levert daardoor [zoals we in sectie 2.4 hebben gezien] diverse problemen van incoherentie op. Voorts botst het postuleren van zo'n metafysische metatheorie met het wetenschappelijke ideaal van empirische weerlegbaarheid. De de-facto herintroductie van een eerste filosofie zal omstreden zijn omdat het gebaseerd is op de dubieuze, impliciete aanname dat de cognitieve toegang tot de werkelijkheid vanuit een rationalistisch, filosofisch benadering superieur is aan die van de empirische natuurwetenschappen. Het definiëren van zo'n metatheorie waarvan de toetsing (zich ontrekt aan het oog van het zintuiglijk waarneembare) alleen langs filosofische weg kan geschieden., kan bovendien de wetenschap mee gaan trekken in de twisten van de filosofie (waarvan het toneel immer wordt beheerst door een afwezigheid van consensus).

Het antireductionisme van Dennett levert in dat opzicht een beter te beheersen oplossing. Toegegeven, berust het ook op een aanname, doch het hanteert in ieder geval wel Ockhams' scheermes, "don't multiply entities beyond necessity". De suggestie van Habermas om een metatheorie te formuleren die beide perspectieven in één ontologie verenigt, suggereert dat beide perspectieven simultaan binnen één werkelijkheid (binnen één ontologie) hun oorsprong vinden. Dit resulteert dan in een hoop vragen en verwarring betreffende de causale verbinding, of wat voor een interactie dan ook, tussen mentale toestanden en neurofysiologische processen die sterk aan het eigenschapdualisme doen denken. De visie van Dennett lost een hoop van deze verwarring op. Vanuit zijn opvattingen is de hele vraag of er een causale relatie bestaat tussen mentale toestanden en neurofysiologische processen onjuist. Waarom? Hoewel de wereld een fysische ontologie heeft, zijn er verschillende houdingen die we kunnen aannemen van waaruit we de wereld interpreteren. Nemen we de fysische houding ten opzichte van gebeurtenissen in de wereld aan, dan hanteren we de fysica als theorie aan de hand waarvan we deze gebeurtenissen in termen van de causale natuurwetten verklaren. Zo wordt dan bijvoorbeeld een beschrijving van de causale relaties tussen neurofysiologische processen en stimuli responses gevormd. Vanuit de intentionele houding hanteren we de grove maar compacte wetten van de volkpsychologie aan de hand waarvan we causale relaties beschrijven tussen mentale toestanden en stimuli responses.

Maar omdat het expliciet twee verschillende elkaar uitsluitende perspectieven zijn op één en dezelfde ontologie, dus niet twee naast elkaar bestaande processen binnen één ontologie, is de vraag dus onzinnig of, en hoe het kan dat, er op- of neerwaartse causale verbindingen zijn.⁴⁸

De vraag getuigt van een 'category mistake'.⁴⁹

En het is ook niet zo dat ik delen van mijn beschrijvingen in termen van mentale toestanden zou kunnen vervangen door beschrijvingen van neurofysiologische toestanden. Er is immers volgens Dennett geen één-op-één relatie te vinden tussen de aan de mentale toestanden ten grondslag liggende gedragspatronen en neurofysiologische patronen. Daarom verwerpt Dennett ook juist het reductionistisch materialisme.

We ontlopen daarmee dus ook het beeld dat al het mentale opgesloten lijkt te liggen binnen dwingende causale keten, een beeld wat binnen het reductionistisch materialisme wordt

gevormd omdat we daarin beschrijvingen van mentale toestanden eenvoudig kunnen uitwisselen tegen beschrijvingen van neurofysiologische processen.

En het levert nog iets extra's op. Door te beweren dat er tussen mentale toestanden en neurofysiologische patronen geen één-op-één relatie bestaat, en door tevens te verdedigen dat de volkpsychologie een prima stuk gereedschap is om gedrag te voorspellen, verwordt de volkpsychologie tot een redelijk zelfstandig moeilijk te bedreigen bastion. ('It is here to stay'). De consequentie van deze antireductie is dat onze volkpsychologische opvattingen t.a.v. rationeel en vrij agency betrekkelijk immuun zijn voor speculaties in het neurowetenschappelijke domein. In het reductionistisch materialisme (wat uitgedragen lijkt te worden in het 'Manifest'⁵⁰, van de elf neurowetenschappers, maar meer daarover in sectie 5), zijn mentale toestanden één-op-één te reduceren tot hersenprocessen. We kunnen ons zo voorstellen dat indien het reductionistisch materialisme waar zou zijn de status van de volkpsychologie direct afhankelijk wordt van, en getroffen door, resultaten uit de neurowetenschap. Maar er is, zoals Dennett reeds opmerkte [sectie 3.1.4], onvoldoende rechtvaardiging te geven voor dit reductionisme.

Uit een opgemaakte tussenbalans vormt zich de tussentijdse conclusie dat de fysicalistische ontologie van Dennett de minst speculatieve visie oplevert. Tegelijkertijd weet Dennett het antireductionistische karakter, en daarmee de geloofwaardigheid van de volkpsychologie, en het door redenen bewogen handelen, plausibel te rechtvaardigen zonder verstrikt te raken in het poneren van mysterieuze relaties tussen het mentale en het lichamelijke.

Het ontbreekt Dennett's filosofie echter aan bepaalde aspecten die de filosofie van Habermas wel verdisconteert. Habermas' filosofie biedt meer plaats aan de communicatieve aspecten en de perspectieven van de concrete-ander die een rol spelen in het tot stand komen van kennis.

Habermas heeft meer oog voor (dat pretendeert zijn filosofie in ieder geval) de intersubjectieve vestiging van normen en waarden, de mogelijkheid van kritiek, hermeneutische benaderingen van de menswetenschappen, en erkenning van een (inter)subjectieve betekenis-als-relevantie voor mij of ons.

Het kan zijn dat dit alleen mogelijk is als we ophouden met het verbinden van waarheid aan correspondentie met een wereld buiten ons, en in plaats daarvan een waarheid gerelateerd aan een pragmatistische relevantie-voor-ons adopteren. Maar daar straks mee over.

Wat ik sowieso mis bij beide filosofen is aandacht voor de idee dat we niet louter denkende maar toch vooral ook nog voelende wezens zijn. Als één gedachte juist benauwt dan is het die wel waarin we vergeleken worden met louter cognitieve actoren, die als rationele automaten los van gevoel of beleven door het leven gaan. Habermas toont zich in zijn nadruk op het ideaal van de ratio meer een verlichtingsfilosoof dan Dennett, die in de fundamenteën van de moraal nog toegeeft aan ruimte voor de natuurlijke instincten en de emotie. Maar bij Dennett loop je weer tegen de uitholling van het subjectieve aspect aan.

In het algemeen zou ik willen zeggen dat een fysicalistische filosofie geaccepteerd zou kunnen worden als het daar bovenop nog een aantal aanvullende aspecten zou kunnen verdisconteren. Ze zou bijvoorbeeld een rechtvaardiging moet kunnen geven (door bijvoorbeeld verschillende perspectieven te hanteren op één en dezelfde werkelijkheid), die ruimte laat voor een uitzicht op agents die hun handelen laten leiden door reflectief gekozen redenen en daarin een vrijheid van de wil kunnen uitoefenen die het wensen waard is. We zullen dan nu daarom bekijken hoe Dennett en Habermas vanuit hun geschetste opvattingen omgaan met de concrete kwestie van de vrije wil.

5 De kwestie van de vrije wil

Vanuit het dagelijks leven kunnen we onszelf begrijpen als rationele, autonome morele agents die de vrijheid van de wil uitoefenen in het reflectief wegen van redenen, het weerstaan van verlangens, en het kiezen van onze handelswijze⁵¹.

Tegelijkertijd erkennen we bijna allen de epistemische autoriteit van de natuurwetenschappen die alle gebeurtenissen in de natuur aan de hand van deterministische natuurwetten verklaren.

Hoe zit het dan met onze vrijheid van de wil als we aannemen dat wij mensen een integraal onderdeel vormen van diezelfde natuur die de natuurwetenschappen aan de hand van deterministische wetten causaal wil verklaren? Als alles in de natuur gedetermineerd is, en wij zijn een onderdeel van de natuur, dan moeten wij dus ook in ons handelen gedetermineerd zijn. Hoe kan het dan, dat als we handelen, we het gevoel hebben dat we een waaier aan alternatieve mogelijkheden hebben? Hoe kan het dat als we terugdenken aan iets wat we gedaan hebben, we vaak het gevoel hebben dat we ceteris paribus (onder overigens gelijke omstandigheden), ook anders hadden kunnen handelen?

Het antwoord op deze vraagstelling, die reeds lange tijd wordt overdacht, verdeelt filosofen over twee kampen: incompatibilisten menen dat vrije wil en determinisme niet compatibel zijn. Compatibilisten denken dat, ondanks dat determinisme waar is, er nog steeds ruimte voor de vrije wil is.

De theorievorming betreffende vrije wil is significant omdat het implicaties voor het morele verantwoordelijkheidsbegrip heeft. Galen Strawson verwoordde het als volgt.

“De fundamentele motor van de vrije wil is de zorg om morele verantwoordelijkheid. Als niemand deze zorg zou hebben, is het twijfelachtig of het probleem van de vrije wil zo’n beroemd filosofisch probleem zou zijn. [Galen Strawson⁵²]”

Laten we nu kijken hoe Habermas en Dennett met deze zorg omgaan.

5.1 Habermas over de kwestie van de vrije wil

Recentelijk heeft ook Habermas zich met zijn essay “The Language game of Responsible Agency, and the problem of free wil.” in het debat gemengd. Voor hem heeft het concept vrije wil eveneens een sterke morele connotatie. Het vervult dan ook in zijn formele pragmatiek impliciet een belangrijke rol.

In een transcendentiaal-pragmatische reconstructie waarbij hij gebruikmaakt van de techniek van de ‘performatieve contradictie⁵³’, tracht hij aan te tonen dat het hebben van een vrije wil een noodzakelijke vooronderstelling is van deelnemers aan een coöperatieve zoektocht naar overeenstemming gebaseerd op een uitwisseling van argumenten. Op deze wijze brengt Habermas vrije wil in verband met vooronderstellingen die een rol spelen in het taalspel van verantwoordelijk ‘agency’. Na eenmaal vastgesteld te hebben dat de vooronderstelling van een vrije wil hiervoor essentieel is, gaat hij op zoek naar de mogelijkheid om deze overtuiging te verenigen met het geloof in de epistemische autoriteit van de natuurwetenschap en de eenheid van de natuur. Dat is globaal de door Habermas in het essay gevolgde strategie waarvan ik hier nu eerst een beknopte compilatie zal geven.

Habermas gebruikt een in 2004 verschenen manifest in “Gehirn und Geist” als inleiding voor zijn betoog. Hierin kondigen neurowetenschappers aan dat spoedig ons zelfbeeld opgeschud zal worden. Volgens Habermas heeft deze uitspraak de natuurwetenschap meegetrokken in een metafysische onderstroom, die in drie krachtige onderlinge strijdige intuïties kan worden weergegeven.

1. *Als agents zijn we overtuigd dat de geest causale vermogens heeft, en dat de geest niet-reduceerbaar onderscheiden is. We zijn er zeker van dat we vanuit eigen instemming handelen, en in de wereld dingen kunnen laten gebeuren.*
2. *Als kennende subjecten erkennen we de epistemische autoriteit van de natuurwetenschappen, die causale effectiviteit toeschrijven aan alle en alleen aan die gebeurtenissen in de wereld die op wetmatige wijze variëren.*
3. *Als wetenschappelijk verlichte personen die reflecteren over hun eigen positie in de natuur, zijn we ervan overtuigd dat het universum een eenheid is, en ons als een deel van de natuur omvat. [Habermas⁵⁴]*

De eerste intuïtie, die ook in verband gebracht kan worden met vrije wil, valt te herleiden tot de causale vermogens van de geest, de tweede tot de nomologische determinatie van alle gebeurtenissen in de wereld, en de derde tot de monistische constitutie van het universum.

Habermas geeft deze intuïties gezag door ze epistemische toestanden te noemen. Omdat hij meent dat de eerste epistemische toestand niet te reduceren is tot de tweede (en vice versa), spreekt hij over een epistemisch dualisme. Tot zover is er nog niets aan de hand ware het niet dat de derde intuïtie, die een ontologisch monisme impliceert, de vereniging van de eerste twee epistemische toestanden problematisch maakt. Want hoe kunnen we de ogenschijnlijk onderling elkaar uitsluitende beschrijvingen van onszelf onderbrengen in één geünificeerd gedacht universum? Habermas wil deze vraag bevredigend beantwoorden. Maar hij mengt zich niet in het debat voordat hij zich eerst op de hoogte heeft gesteld van de posities die door anderen op het terrein reeds zijn ingenomen. Hij kijkt eerst bij naturalistisch geïnspireerde filosofen of die de kwestie reeds bevredigend hebben opgelost.

5.1.1 Enige relativering met betrekking tot het artikel van de neurowetenschappers

In het artikel van de elf neurowetenschappers ‘Das Manifest’ wordt inderdaad melding gemaakt van een op handen zijnde herziening van ons zelfbeeld; *“Was unser Bild von ins selbblt betrifft, stehen uns in sehr absehbarer Zeit beträchtliche erschütterungen ins Haus⁵⁵”*. Maar ten eerste, zoals ook Timothy Schroeder in een commentaar⁵⁶ op het artikel van Habermas opmerkt, kan de ‘schokkende aanpassing van ons zelfbeeld’ vooral van toepassing zijn bedoeld voor hen die nog menen dat we een volledig vrije onveroorzaakte zelf hebben, en als mini-Goden, als eerste oorzaken ons handelen kiezen. Maar deze ex nihilo veroorzaking vind onder filosofen vrijwel geen aanhang meer.

Enige relativering t.a.v. de voorspelling is ook op zijn plaats. In het artikel wordt gesproken over de resultaten van het hersenonderzoek, op drie kennislagen. In de onderste kennislaag, die de werking van afzonderlijke neuronen betreft, is in de laatste decennia een enorme vooruitgang geboekt. Dat geldt tevens voor de bovenste laag, de kennis van de verschillende

functionele hersenmodules, zoals amygdala, limbisch systeem en cortex. Maar van het middengebied, de beschrijvingen van het gebeuren binnen neurale netwerken van honderdduizenden cellen, weet men nog zeer weinig.

“Über die mittlere Ebene – also das Geschehen innerhalb kleinere und Größere Zellverbände, das letztlich den Prozessen auf der obersten Ebene zi Grunde liegt – wissen wir noch erschreckend wenig⁵⁷.”

Frank Rösler toont de betrekkelijkheid van de mogelijkheid om te voorspellen of iemand in een zekere situatie een moord zal gaan plegen. *“Om met zekerheid te voorspellen zou men de toestand van 10^{15} synapsverbindingen moeten kennen⁵⁸.”* Gesteld dat men een manier zou kunnen vinden om dat te doen, wat zeer onwaarschijnlijk lijkt, dan nog moet men ook de codering van de informatieoverdracht ofwel communicatie tussen de neuronen onderling en tussen de netwerken kennen. Dat vraagt om een soort van psycho-semantiek. En hoe die semantiek tot stand komt, ongetwijfeld met verdiscontering van de gehele context (de omgeving) is nog in het geheel niet bekend. De hersenen zijn daarbij zeer plastisch, dwz veranderen voortdurend onder invloed van omgevingsfactoren. Als we ons dan ook nog realiseren dat een token-identiteitstheorie veel waarschijnlijker lijkt dan een type-identiteitstheorie, (geen twee breinen realiseren het mentale op dezelfde neurofysiologische wijze), dan kan men nagaan wat voor een enorme problemen en een uitdaging hier in het verschiet liggen. Dus aannemen dat ons gedrag uiteindelijk een neurofysiologische basis heeft is één ding. Maar voorspellen van gedrag is nog weer een heel ander verhaal. Al met al is daarmee het Manifest nog een tamelijk relativerend en mild stuk.

5.1.2 Habermas beoordeelt naturalistische benaderingen van vrije wil

Het geünificeerde universum, dus het ontologisch monisme van hard naturalistische filosofen, wordt uitgedrukt in de verdediging van de naturalistische premissen, dat alles in het universum door natuurwetten wordt gedetermineerd.

Habermas brengt een onderscheid aan in het soort van oplossingen voor het vrije wil probleem. Ofschoon compatibilisten aannemen dat we geen (ceteris paribus) onvervalste open keuzemogelijkheden hebben, is dit geen belemmering voor een verzoening met de aanname van vrije wil. Hetzij, omdat ze menen dat de verzoening mogelijk is omdat de aanwezigheid van alternatieve mogelijkheden voor de toeschrijving van verantwoordelijkheid irrelevant is. Hetzij omdat ze net zoals Dennett denken dat het in de praktijk niet uitmaakt dat we de wereld als materialistisch en causaal gesloten begrijpen daar de agent toch slechts een beperkte epistemische horizon heeft. De toekomst lijkt open omdat we immers niet kunnen weten welke mogelijke wereld in de oneindige verzameling, de onze is. Habermas verwierpt beide oplossingen omdat ze de scheiding tussen het waarnemer- en participantperspectief vertroebelen. De waarnemer weet iets, wat de participant zgn. niet zou weten, namelijk, dat de wereld gedetermineerd is, en dus geen alternatieve mogelijkheden toestaat. Maar dit is volgens Habermas een ongeldige redenering omdat de participant en de waarnemer in één en dezelfde persoon verenigd kunnen zijn. En wat iemand in de hoedanigheid van waarnemer weet kan hij niet simpelweg vergeten als hij in de rol van participant zit. Het resultaat is dat de participant uiteindelijk ook gaat denken dat hij geen alternatieve mogelijkheden heeft. *“wat de een weet kan de ander simpelweg niet ontkennen⁵⁹”.*

De andere categorie filosofen voor wie het hebben van een vrije wil compatibel is met de naturalistische premisse, tracht voor de mentale toestanden een geschikte plaats te reserveren binnen het, in fysische termen geconceptualiseerde, naturalistische geraamte. Habermas neemt hier als voorbeeld het zgn. "anomalous monism"⁶⁰ van Donald Davidson.

Ook deze oplossing wordt door Habermas verworpen. In zoverre het doel van Davidson's argument was om ruimte te reserveren voor mentale activiteiten binnen de in fysicalistische termen verklaarde wereld, heeft het 't mentale nu overtollig gemaakt. Indachtig de kritiek van Jaegon Kim, die onder de aanname van causale geslotenheid van het universum argumenteerde dat mentale toestanden in het eigenschapdualisme van Davidson epifenomenen zijn, "heeft Davidson's poging om de geest voor reductie te behoeden, nu onbedoeld tot een eliminatie ervan geleid".

Zo komt Habermas uiteindelijk tot een verwerping van alle gangbare naturalistische oplossingen voor het probleem van harmonisatie van epistemisch dualisme en ontologisch monisme. En de gemeenschappelijke grond van verwerping ligt in het feit dat bij al deze naturalistische opvattingen, in geval van een conflict met het participantperspectief, het objectiverende (waarnemer) perspectief als de finale arbiter beslist over de ontologische en nomologische eigenschappen van het allesomvattende universum. Maar dat is onjuist daar dit volgens Habermas gebaseerd is op de methodologische fictie van een exclusieve 'view from nowhere'. Een fictie die geen rekening houdt met het feit dat de waarnemer geen directe maar bemiddelde cognitieve toegang tot de wereld heeft, en hem daardoor slechts een bepaald (gestoord) segment van die werkelijkheid toont. Het complementaire segment van het participerende perspectief mag niet ontbreken, of meer dan dat, verdient een zelfde status als het objectieve perspectief. Het participerend perspectief brengt immers performatief middels het intersubjectief tot stand komen van de bepaling van de betekenis(-als-relevantie) van onze taaldaden de mogelijkheidsvoorwaarden voor de objectiverende toegang tot de wereld tot stand.

De zojuist besproken naturalisten lossen het conflict tussen participant, en waarnemerperspectief, dwz tussen de intuïtie van vrije wil en determinisme op, door het perspectief van de waarnemer het laatste woord over de kwestie te geven. Daartoe reduceren of elimineren ze de beschrijvingen vanuit het participantperspectief. Dat doen ze bijvoorbeeld door de conceptualisering van het dagelijks gebruik van het begrip 'vrije wil' (hetgeen we er onder plachten te verstaan) zodanig af te zwakken dat het in een deterministisch perspectief past.

In het antwoord op de vraag of vrije wil compatibel is met determinisme, speelt de conceptualisering van vrije wil een belangrijke rol. Habermas lijkt vast te houden aan een sterke definitie van vrije wil. Het antwoord op de vraag wat Habermas dan onder vrije wil verstaat leidt onverkort naar zijn theorie van communicatief handelen. Het openbaart zich wanneer iemand om een verklaring voor zijn doen en laten wordt gevraagd. In het taalspel van verantwoordelijk agency wordt als vooronderstelling aangenomen dat participanten naar zelfgekozen redenen handelen, en dat ze desgevraagd hun handelingen door het geven van beweegredenen kunnen verantwoorden. Daarin wordt volgens Habermas impliciet aangenomen dat de participanten *ceteris paribus*⁶¹ ook anders zouden hebben kunnen handelen.

"De kracht van goede redenen hangen af van de resonantie die ze aan het licht brengen binnen een subjectieve geest, terwijl de ruimte van redenen het milieu levert voor

geldigheidsclaims welke personen vis-a-vis met elkaar naar voren brengen waarbij ze een Ja- of Neen positie innemen. De ‘agent’ kan niet op het zelfde moment zijn eigen optreden waarnemen als een gebeurtenis die oorzakelijk tot stand is gebracht. De normatieve structuur van het deliberatieve proces, gecombineerd met de feilbaarheid van haar uitkomst, maakt het voor de participanten conceptueel noodzakelijk, om in hun praktisch redeneren de beschikbaarheid van alternatieve mogelijkheden te veronderstellen. [Habermas⁶²]”

De achtergrondbetekenis van vrijheid die personen performatief vergezelt omvat in Habermas opvatting twee belangrijke aspecten:

- We verwachten als rationele agents van elkaar dat we ons handelen laten leiden door reflectief gekozen redenen die de onze zijn.
- Als agents gaan we er vanuit dat we in het kiezen van onze redenen (ceteris paribus) alternatieve mogelijkheden hebben.

Het eerste aspect (bekend onder de naam ‘reason responsiveness’) omvat het besef van een verwachting dat agents deliberatief, reflectief en rationeel, tot een redelijke afweging van redenen voor hun handelen komen, en dat ze zich vervolgens de beslissende redenen ook toe-eigenen. Maar het is juist de tweede eigenschap die Habermas met vrije wil associeert, die met determinisme conflicteert. We zullen weldra zien dat in Dennett’s compatibilisme deze eis niet wordt gesteld.

Habermas meent dat de inhoud van de aanname van vrije wil zichzelf alleen openbaart aan participanten die (vis-à-vis) een tweede persoon, als toehoorder een performatieve houding aannemen. Het blijft ontoegankelijk vanuit het gezichtspunt van de waarnemer, d.w.z. vanuit het perspectief van de niet betrokken derde persoon (bijvoorbeeld een hersenonderzoeker). Het laatst genoemde perspectief constitueert louter nomologisch causale verklaringen, terwijl participanten de vrije wil als vooraanname tot de ‘wereld’ rekent (d.w.z. deze vooraanname als element van de verzameling van noodzakelijke vooronderstellingen in de ideale spreek situatie niet in twijfel trekken), omdat zonder vrije wil het gehele taalspel van het communicatief handelen onmogelijk zou zijn. *“Zonder zijn twee peilers – de aannamen van de causale effectiviteit van de geest en de verbinding tussen reflectie en vrijheid – stort het taalspel van verantwoordelijk ‘agency’ in elkaar”*[Habermas⁶³] En waarom is volgens Habermas mentale effectiviteit onmogelijk in een nomologisch gedetermineerde wereld? Omdat *“als de complexe veroorzaking van menselijk gedrag direct moet worden gevonden in het activeringsproces van het brein, ..het vreemdsoortig [is] om te denken dat personen de toestand kunnen veranderen door intentioneel te interveniëren. Dan is er niet zoiets als neerwaartse veroorzaking van geest naar brein”*

5.1.3 Commentaar op de beoordeling van Habermas

Habermas’opmerking aangaande ‘neerwaartse veroorzaking van geest naar brein’ en zijn conclusie dat mentale effectiviteit in een nomologisch gedetermineerde wereld onmogelijk is houdt verband met zijn impliciet aanwezige aanname van een vorm van eigenschapdualisme. Hij lijkt te veronderstellen dat eigenschappen van de cultuur en geest waarvan de beschrijvingen niet zijn te reduceren tot beschrijvingen van fysische eigenschappen, simultaan met de fysische eigenschappen door een monistische ontologie moeten worden

geinstantieerd. Dat is althans de consequentie van zijn voorstel om het participierend en objectiverende perspectief zonder onderlinge reductie samen te brengen in één metatheorie. Het is waar dat eigenschapdualisme inderdaad in verband wordt gebracht met epifenomenalisme van het mentale, maar de genoemde citaten worden door Habermas ingezet tegen het reductionisme van de neurowetenschappers van het 'Manifest'. Daarvoor is het een incorrect tegenargument. Reductionistisch fysicalisme kan niet epifenomenalisme van het mentale in de schoenen worden geschoven. Want reductionisme, dwz een type-identiteitstheorie, houdt in dat onze beschrijvingen in mentale termen kunnen worden vervangen door neurofysiologische beschrijvingen. En daar neurofysiologische toestanden causaal effectief zijn, moeten mentale toestanden dat dan ook zijn.

Een andere verwarring betreffende de neurofysiologische realisatie van het mentale is aanwezig in de bewering: "*Redenen staan in semantische relaties tot andere redenen. Ze zijn fundamenteel onderworpen aan kritiek vanuit tegenovergestelde redenen, terwijl causaal verklaarbare mentale toestanden of episodes elkaar niet kunnen tegenspreken.*" [Habermas]⁶⁴ Hier kunnen we het volgende bezwaar tegen maken. Zoals redenen in spraak of schriftsymbolen van onze omgangstaal worden uitgedrukt, en we daarvoor een semantiek kunnen ontwikkelen die verklaart hoe de intentionele betekenis van deze symbolen tot stand komt, zo kan evenzogoed een psycho-semantiek denkbaar zijn, die verklaart hoe neurofysiologische patronen in het brein een intentionele betekenis kunnen hebben. Er is voor wat betreft de verklaring voor het verkrijgen van betekenis geen belemmering in het verschil in fysische aard van de drager die het symbool vormt; of het nu klanken, tekens op papier, of neurale patronen zijn.

5.1.4 Habermas' zwak naturalistische benaderingen van vrije wil

In zijn essay 'The Language Game of Responsible Agency' bespreekt Habermas een op Sellars' visie gebaseerde ontwikkeling van onze cognitieve toegang tot de wereld welke laat zien dat de twee epistemische perspectieven van waarnemer en participant kunnen worden herleid tot 'wereld perspectieven' die simultaan hun oorsprong hebben in onze vorm van linguïstische communicatie. Habermas voert een aantal argumenten aan voor de stelling dat het dubbele perspectief (participerende, en, objectieve) en niet een reductionisme, tot het correcte zelfbegrip van de mensheid behoort.

- Het ontbreken van een geünificeerde terminologie voor zowel mentale operaties en breintoestanden.
- Het ontbreken van een naadloos aansluitende ordening van wetenschappelijke theorieën.

Gegeven nu het feit dat Habermas het participerende perspectief een zelfde gezag toekent als het waarnemend perspectief, en bovendien vasthoudt aan een ontologisch monisme, ziet hij zich gesteld tegenover een lastig op te lossen probleem. Een naturalistisch aanname die de natuur van de natuurwetenschappen met het geünificeerde universum identificeert komt dan namelijk niet meer in aanmerking. Zo'n aanname zou je tenslotte conflicterende overtuigingen opleveren. Enerzijds geloven dat je als mens het vermogen hebt om (ceteris paribus) evenzogoed zus als zo te kunnen handelen, botst met het geloof dat je als mens in zijn geheel, onderdeel uitmaakt van een geünificeerd universum die in theorie geheel verklaard kan worden a.h.v. de causale determinerende wetten van de natuurwetenschappen.

De oplossing voor het vastlopen van naturalistische opties tegen de starre weerstand van een onvermijdbaar perspectivistisch dualisme, moet volgens Habermas, gevonden worden in de adoptie van zijn zwak naturalisme.

We moeten gaan zoeken voorbij de nomologische conceptie van natuur. Zoals we hebben gezien, omdat we geen mogelijkheid hebben om over en achter de bemiddelende mogelijkhedenvoorwaarden van kennis te komen, is de enige oplossing een aangepaste transcendentale analyse teneinde te reflecteren op de pragmatische voorwaarden voor onze cognitieve toegang tot de wereld. Aangepast omdat een detranscendentalisering wordt uitgevoerd die bestaat in het plaatsen van de transcendentale mogelijkhedenvoorwaarden van kennis in de linguïstische socio-culturele vormen van de leefwereld.

“Onze taak is een verzoening tussen Kant en Darwin”, zo zegt Habermas. Hij bedoelt daarmee dat we het onderzoek naar de transcendentale voorwaarden voor kennis moeten combineren met een evolutietheorie.

Niet een evolutietheorie in de snit van de nomologische natuurwetenschappen, uitgebreid tot voorbij de grenzen van de geneselectie. Dus niet louter een verklaring van de ontwikkeling van de socioculturele vormen van de leefwereld door een getransponeerd nomologisch model gebaseerd op mutatie, selectie en overerving. Hiermee wijst Habermas de toepassing van het universeel darwinisme en de memetiek van Richard Dawkins af.

We moeten volgens Habermas niet uit het oog verliezen dat in de ontwikkeling van de mogelijkhedenvoorwaarden voor kennis, de kwaliteit van iets wat het waard maakt om te weten, niet alleen afhangt van de wereld maar ook van onze positie in de wereld. Dus we moeten ons niet laten verleiden tot een visie waarin al onze kennis wordt verondersteld een afspiegeling van de natuur te zijn. Bekeken vanuit het perspectief van de evolutie is waarheid niet noodzakelijk synoniem met de correspondentie van een teken met iets in de werkelijkheid. Een betere benadering van de reconstructie van de ontwikkeling van onze cognitieve toegang tot de wereld begint met de benadering van Dewey die stelt dat kennisaccumulatie in nauw verband staat met actief handelen en probleemoplossen. Waarheid wordt i.p.v. met correspondentie met succes in het oplossen van problemen, en met de selectie in de strijd om het bestaan in verband gebracht: “*waar is datgene wat door de praktische gevolgen wordt bewaarheid*⁶⁵.” Waarheid is wat tot stand komt door het testen en verdedigen van onze uitspraken ten overstaan van anderen, en een instemming van allen in een open deliberatief proces van allen voor allen. Feiten zeggen natuurlijk iets over de wereld maar wij zijn degenen waarvoor ze iets zeggen.

Bij de reconstructie van de evolutie van onze mogelijkhedenvoorwaarden van kennis horen de twee methodieken van het verklaren en het verstehen, elk met hun eigen logica, die in verband kunnen worden gebracht met de perspectieven van subject-object en subject-subject. In het kennisbegrip van Habermas is kennis niet een afbeelding van een buiten de mens gelegen wereld. Dat betekent dat noch het objectiverende, noch het participerende perspectief als beeld van de werkelijkheid kan dienen. De mens construeert eerder zijn werkelijkheid in de vorm van deze twee perspectieven. En in die constructie evolueert ze, leert ze, door problemen op te lossen in de ruimtelijke dimensie door intelligent om te gaan met een risicovolle omgeving, en in de sociale dimensie door het argumenterend rechtvaardigen van de eigen manieren van probleem oplossen tegenover die van anderen. Beide leerprocessen van kennistoename zijn bepalend voor de overlevingskansen van onze soort.

Het zwak naturalisme van Habermas verbiedt dat het ene leerproces wordt geassimileerd in het andere. Beide perspectieven moeten als het ware worden verbonden op een metatheoretisch niveau, door een continuïteit tussen natuur en cultuur te veronderstellen. Dat wil zeggen dat we in ons historisch onderzoek naast een constructie vanuit het objectiverende perspectief in termen van causaliteit, evenzoveel aandacht besteden aan de hermeneutische onderzoeksmethoden⁶⁶ vanuit het perspectief van participanten. Anders gezegd: we moeten een instrumentele verklaring combineren met een communicatief rationeel verstaan. Wat zo ontstaat is een anti-reductionistische gedetranscendentaliseerde natuurlijke geschiedenis die de natuur bevrijdt uit het korset van de fysicalistische ontologie. Zeg maar, een ontologie waarin de natuur meer dan alleen de natuur van de natuurwetenschappen omvat. De plaats van de natuurwetenschap met haar nomologisch causaliteitsbegrip moet, overeenkomstig de interventionistische opvatting van causaliteit van Kant en Peirce, worden teruggewezen naar dat deel van de natuur wat kan worden geobjectiveerd en onder de technische beheersing valt.

Resumerend zien we dat de reconstructie van de evolutie van de vrije wil bij Habermas vooral betrekking heeft op een bezinning op de mogelijkhedenvoorwaarden van kennis.

Net als Kant beweegt Habermas weg van een verklaring van de wereld in zichzelf om zo de epistemologie als eerste filosofie weer bovenaan de agenda te plaatsen. Voortbordurend op Sellars' onderscheid tussen manifeste en wetenschappelijke kennis, rekent Habermas de intuïtie dat wij vrije morele door redenen bewogen agents zijn, tot onze manifeste kennis, die complementair is aan de objectiverende kennis van de wetenschap. Beide bijbehorende perspectieven, die elk bijdragen aan onze cognitieve toegang tot de wereld en we dus even series moeten nemen, moeten we zonder onderlinge reductie metatheoretisch verbinden. De consequentie hiervan is dat het onderzoek naar de mogelijkhedenvoorwaarden van onze cognitieve toegang, naast verklaren evenzo gericht moet zijn op het verstaan van het verloop in de evolutionaire ontwikkeling van deze cognitieve toegangspaden tot de wereld. Habermas' onderzoek van de evolutie van de mogelijkhedenvoorwaarden van dit epistemologisch dualisme maakt dus niet louter gebruik van de instrumentele ofwel empiristische onderzoeksmodellen, zoals speltheorieën of universeel darwinistische modellen. Die leveren slechts verklaringen in termen van waarde-neutrale beschrijvingen van blinde causale processen. De normatieve aspecten worden pas zichtbaar als naast deze empiristische modellen ook hermeneutische methoden worden toegepast. Dat openbaart niet alleen de ontwikkeling van de instrumentele maar ook van de communicatieve rationaliteit.

5.2 Dennett over de kwestie van de vrije wil

Niet een reductionisme maar het dubbele perspectief vormt volgens Habermas het correcte zelfbegrip van de mensheid. Maar waar Habermas het antireductionisme realiseert in een metafysica bestaande uit een metatheoretische verbinding van epistemisch dualisme met een monistische visie op een natuur die meer omsluit dan de natuur van de wetenschap, levert Dennett's antireductionisme een fysicalistisch alternatief. In contrast met de visie van Habermas beweert Dennett dat vrije wil en een deterministisch universum wel compatibel zijn. Sterker nog, voor Dennett is determinisme de noodzakelijke voorwaarde voor een vrije wil 'die het ook waard is om te hebben'⁶⁷. Dennett postuleert geen transcendente wereld of andere metafysica. Zulke oplossingen omschrijft hij als 'skyhooks'. Dat is mysterieuze hulp uit de hemel die verder niets verklaart. Hij houdt het bij 'Cranes' hijskranen, fysische verklaringen, dwz.

verklaringen vanuit een fysicalistisch naturalisme. Vanuit dat kader probeert hij in zijn boek 'Freedom evolves' de evolutie van de vrije wil te reconstrueren door een verhaal van 'zo zou het kunnen zijn gegaan' te vertellen. Het verhaalt over de ontwikkeling van de eerste primitieve vormen van vermijdbaarheid, in zijn ogen een synoniem voor vrijheid, van eencellige tot aan de bewuste, rationele en vrije mens van vandaag.

5.2.1 Dennett over determinisme en vrije wil

Om het probleem van de veronderstelde incompatibiliteit van vrije wil en determinisme eens goed tegen het licht te houden, reflecteert Dennett eerst op wat een gedetermineerd universum wel en niet impliceert. Uit de definitie van determinisme die luidt dat er op elk moment precies één fysisch mogelijke toekomst is,⁶⁸ wordt volgens Dennett maar al te vaak de onjuiste conclusie getrokken dat determinisme onvermijdelijkheid impliceert. De verbeelding van Laplace twee eeuwen geleden betreffende de calculeerbaarheid van de toekomst, heeft bijgedragen tot een verkeerd begrip van de consequentie van determinisme voor de mens. Laplace schreef: *"Een oneindig intellect (een Laplaciaanse demon⁶⁹) kan, als het over de toestandbeschrijving van een zeker moment van een deterministisch universum⁷⁰ beschikt, de toekomst (niet het verleden) berekenen"*.

Maar hoe gedetailleerd moet je zijn met je toestandbeschrijving? Moeten we elk subatomair deeltje beschrijven? Om de gedachten te bepalen en de hele kwestie te verhelderen introduceert hij twee instrumenten (vereenvoudigde modellen van de wereld). Allereerst laat hij aan de hand van Quine's Democritische werelden (gedigitaliseerde driedimensionale universa) en de toepassing van modale logica⁷¹ zien, dat er reeds bij de beschouwing van een klein volume ontzaggelijk veel mogelijke deterministische werelden⁷² zijn. Tevens verbeeldt Dennett met rommelige werelden de constatering van Hume, dat causaliteit niet een logisch begrip is maar een uit de empirie afgeleid concept, waardoor er geen contradictie schuilt in het denkbeeldige geval dat van de ene op de andere dag de natuurwetten (overgangsregels) zouden veranderen. Dus strikt genomen kunnen we niet eens zeker weten of we in een gedetermineerde of een "rommelige" wereld leven. De verbeelding van zo'n klein speelgoed universum laat zien dat we het vermogen de toekomst van ons universum te voorspellen, met scepsis moeten benaderen. Onze epistemische horizon is simpelweg te beperkt om te weten in welke mogelijke wereld we ons bevinden. *"Iedere eindige informatiegebruiker heeft een epistemische horizon; hij weet minder dan alles over de wereld die hij bewoont en deze onvermijdelijke onwetendheid garandeert dat hij een subjectieve open toekomst heeft. Spanning is een noodzakelijke levensvoorwaarde voor al dergelijke actoren"*⁷³

Het tweede instrument is de deterministische Conway's Leven-Wereld (een speelgoedmodel uit zijn eerdere essay "Real Patterns"). Deterministisch omdat het moet gehoorzamen aan twee overgangsregels (natuurwetten) op "fysisch niveau". Op een meer panoramisch niveau er vanuit een ontwerphouding naar kijken, waarbij we een andere beschrijvende taal gebruiken, kunnen we in de patronen creaturen herkennen en zien "leven" en als gevolg van "milieuomstandigheden" zien ontstaan en vergaan. Programmeurs, hij noemt ze hackergoden, kunnen, door geschikte begintoestanden in te voeren, creaturen gaan ontwerpen die onafhankelijker worden van de omgeving, aan vrijheid winnen door bijvoorbeeld een stapje opzij te doen bij een aanstromend projectiel en zo door een strategie van vermijden zo lang mogelijk in "leven" proberen te blijven. En zie: daar is vrijheid ontstaan!

Dennett tracht de parallellen in 'Freedom evolves' met onze wereld nog verder door te trekken door te zien of we aan de "deus ex machina" kunnen komen. Eerst zijn het nog de hackergoden die het "onderzoek" naar meer ontwikkelde creaties verrichten maar kan er niet een evolutie mogelijk zijn in de Leven-wereld? De botsingen kunnen zorgen voor mutaties! Als zelfreproducerende creaties kunnen de door Conway voor de leven-wereld ontwikkelde Turingmachines dienen. En selectie is er ook omdat een zelfreproducerende creatie die beter tegen botsingen bestand is meer nakomelingen zal krijgen.

De analogie die hij wil aanbrengen is duidelijk: we hebben nu het model van een wereld die gedetermineerd is maar waar zich toch zoiets als vrijheid in kan ontwikkelen. Vermijding noemt Dennett dat. En van de Leven-wereld kan, door er op een bepaalde manier naar te kijken, worden aangetoond dat het veel parallellen met onze wereld vertoont. Natuurwetten op het grondniveau. Maar bekeken vanuit het perspectief van hogere abstractie, structuren en patronen, ja, als je wilt een vorm van 'intentioneel leven' met een graad van vrijheid in de vorm van een vermogen gevaar af te wenden. Gevaar wat zonder dit vermogen fataal voor het voortbestaan van dat leven zou zijn geweest.

Dennett concludeert dat determinisme geen onvermijdelijkheid impliceert. Als men zegt dat determinisme alles onvermijdelijk maakt, dan begaat men de vergissing het woord 'onvermijdelijk' in de fysische context te gebruiken.

"Als 'onvermijdelijk' niet slechts een manier voor 'bepaald' is, wat voegt het dan nog toe? Onvermijdelijke uitkomst? Onvermijdelijk voor wie? Voor de wereld als geheel? Dat slaat nergens op aangezien de wereld geen actor is die er belang bij heeft iets te vermijden." [Dennett⁷⁴]

Het woord hoort niet thuis in de fysische context maar op het ontwerp, of intentioneel niveau. Vermijdbaarheid is een vorm van vrijheid; het betekent dat je iets aan bedreiging kan doen dankzij een mechanisme of strategie van vermijding zonder welke je onvermijdelijk een slachtoffer zou zijn geworden. Het is vermijdelijk vanuit de optiek van de actor. Zo'n strategie heeft ook alleen maar zin in een gedetermineerde wereld. Vermijdelijkheid is compatibel met determinisme, onvermijdelijkheid met indeterminisme. Iets is onontkoombaar wanneer je er niets aan kan doen, zoals een onbepaalde bliksemstraal je doodt. Dat is onvrijheid.

"De denkgewoonte die ik hier voorstel moet herkend worden voor wat ze is, een antropocentrisch (of tenminste actorcentrische) projectie. [...] De begrippen onvermijdelijkheid en vermijdelijkheid, horen eigenlijk thuis op het ontwerpniveau, niet op het fysische niveau." [Dennett⁷⁵]

Hoe zit het verder met onze alledaagse ervaringen die aan kwesties van de vrije wil raken? Ons gewone denken over mogelijkheid, noodzakelijkheid en veroorzaking lijkt strijdig te zijn met determinisme. De aanname dat het gehele universum gedetermineerd is, lijkt ons van onze mogelijkheden te beroven en schijnt ons lot in de totale reeks causale ketens die tot ver in het verleden reiken te bezegelen. Maar dat is een illusie volgens Dennett. Het determinisme impliceert niet dat, wat we ook doen (a) we het niet anders hadden kunnen doen, (b) dat elke gebeurtenis een oorzaak heeft, en (c) dat onze karakters onveranderlijk zijn.

- a) We hadden iets anders kunnen doen: Neem het geval waarin Austin na het missen van een bal zegt ‘ik had de bal kunnen putten’. Natuurlijk, in de enge zin van ‘mogelijkheid’ waarin alle omstandigheden hetzelfde worden gehouden, en Austin weer put, zal hij weer missen. Maar zo denken we niet over ‘mogelijkheid’. In plaats daarvan ‘veranderen’ we in onze gedachten de wereld dan een klein beetje, kiezen een mogelijke wereld die iets afwijkt, en kijken dan of we de bal hadden kunnen putten.
- b) Niet elke gebeurtenis heeft een oorzaak: Stel dat S0 een zin is die tot in het kleinste detail de beschrijving van het universum op t0 specificeert, en S1 op t1, dan dicteert determinisme (S0 → S1) dat S0 weliswaar toereikend maar niet noodzakelijk is voor S1. Aangezien causaliteit in het algemeen noodzakelijkheid veronderstelt, heeft de waarheid van determinisme weinig of helemaal niets te maken met de geldigheid van onze causale uitspraken. In feite is determinisme zelfs perfect compatibel met de opvatting dat sommige gebeurtenissen helemaal geen oorzaak hebben [bezien vanuit onze beperkte horizon]. Tossen is zo’n voorbeeld van een gebeurtenis die een resultaat oplevert (bijvoorbeeld kruis) die eigenlijk geen oorzaak heeft. Het resultaat van een toss is ongetwijfeld het resultaat van de som van totale krachten die op de munt inwerken, maar doordat het gevoelig is voor zoveel variabelen is er gewoon geen mogelijke eindige lijst voorwaarden als oorzaak te selecteren.
- c) Onze karakters zijn niet onveranderlijk: Het is een vergissing om te denken dat echte mogelijkheden onder determinisme verdwijnen. De implicatie van een ‘vaste persoonlijke toekomst’ is iets heel anders dan de implicatie van een ‘vaste persoonlijke aard’. Het kan heel goed zijn dat iemands gedetermineerde persoonlijke toekomst gezegend is met een veranderlijke aard die in hoge mate ontvankelijk is voor de ‘activiteit van het zelf’. Onze karakters zijn juist niet bepaald omdat de evolutie ons de mogelijkheid gegeven heeft om onze aard te veranderen naar de eisen die de omgeving er aan stelt

5.2.2 Dennett over de evolutie van de vrije wil

Dennett’s opzet in ‘Freedom evolves’ is een naturalistische reconstructie van de evolutie van de vrije wil. Daarmee tracht hij deze geschiedenis te verklaren vanuit een door mutatie, selectie, en reproductieve overerving getypeerd mechanistisch model.

Aldus reconstrueert hij een natuurwetenschappelijk, biologisch verhaal over de ontwikkeling van de groei van de vrijheid van de eerste eencellige naar die van de vogel. Hij gebruikt de tot de verbeelding sprekende vrijheid van de vogel om naar de vrijheid van de mens te springen. Met de mens ontstond de cultuur. En zo werden de verticale (van ouder op kind) paden van sociale informatieoverdracht aangevuld met de horizontale paden van memen⁷⁶-verspreiding, verbreed tot culturele informatie-super-hoofdweg. Via deze verticale en horizontale wegen leerden wij de taal en werden van mensen tot personen. En zo past Dennett ook de verklarende mechanismen van mutatie, selectie, en erving van het universeel darwinisme⁷⁷ op cultuuroverdracht toe. Hij schuwt daarin niet het regelmatig gebruik van speltheoretische inzichten. Maar tegen het eind van de route zet hij de bril van Richard Dawkins af.

Aangekomen bij ons in het heden, zoekt hij naar een concept van verantwoordelijkheid en moreel actorschap wat stand kan houden onder de druk van nieuwe wetenschappelijke onderzoeksresultaten. Zijn verklaringen blijven daarbij het naturalisme getrouw, dwz vertrouwen in de (natuur)wetenschap als hoogste autoriteit in de zoektocht naar waarheid. Vertrouwen in het objectiverende perspectief, en de inname het intentionele perspectief waar processen gecompliceerd worden. Maar het gaat over gedragspatronen, blijft ‘open to the

eye', en is objectiverend van aard. Aldus reflecteert hij als een psychisch ingenieur op het fundamentele concept van verantwoordelijk morele actor. Hij meent daarbij dat het waarborgen van onze vrijheid is gebaat bij het demythologiseren van het zelf. Gebruik de wetenschappelijke studies over onze natuurlijke aard om de morele instituties daarop aan te passen. Dit geeft in ieder geval een naturalistisch kader waarin de traditionele problemen van de moraal ergens op slaan. Bouw geen verantwoordelijkheidsbegrip op een metafysica omdat dat niet sterk genoeg is om de nieuwe bevindingen van de wetenschap te weerstaan. In een vergelijking met Habermas' moraaltheorie zal dit in meer detail worden besproken.

6 Habermas versus Dennett m.b.t. de kwestie van de vrije wil

Habermas zwak naturalistische voorstel roept een reeks van vragen op. Hoe verloopt daarin de interactie tussen dat wat voorgesteld wordt gelegen te liggen achter het participant-, en waarnemerperspectief? Onze intuïtie vertelt in ieder geval dat er iets van een interactie tussen mentale en fysische fenomenen moet zijn. Veel van de problemen waar we in het eigenschapdualisme (bijvoorbeeld een emergentisme) mee te maken hadden keren hier terug. Denk aan het probleem met de verklaring van mentale veroorzaking waarvoor epifenomenalisme dreigt als we de causale geslotenheid, gedicteerd door de wetten van de thermodynamica willen blijven eerbiedigen.

Verder mag de metatheoretische verbinding tussen beschrijvingen vanuit het perspectief van de participant en die van de waarnemer niet nomologisch zijn, daar dan toch weer het determinisme met Habermas' sterke definitie van vrije wil (gedefinieerd als *ceteris paribus* alternatieve mogelijkheden) zal conflicteren. Maar als deze verbinding niet nomologisch is, hoe kan toeval dan bijdragen aan een vrije wil?

Voorts is de consequentie van het postulaat van het transcendentale (het achter de perspectieven gelegen niet direct door ons kenbare) dat de metatheorie empirisch onweerlegbaar is. Zo'n metatheoretisch oplossing wordt dan vatbaar voor de bezwaren van empiristen tegen rationalisme en idealisme, dat als je iets postuleert wat achter de zintuiglijke ervaring verborgen blijft, je dan de sluizen opent voor een waterval aan speculatieve theorieën die alleen nog maar op basis van het criterium van coherentie kunnen worden weerlegd.

Kortom, het zwak naturalistisch voorstel van Habermas is zeker niet onproblematisch. Misschien moeten we het scheermes van Ockham ter hand nemen, trouw zweren aan zijn *maxime*, ('don't multiply entities beyond necessity') en trachten een eenvoudiger model te vinden om hetzelfde of meer te verklaren. Daarom wil ik nu de zwak naturalistische benadering van de vrije-wil-kwestie uitzetten tegen het fysicalisme van Dennett.

6.1 Habermas versus Dennett m.b.t. determinisme en vrije wil

Omdat de kwestie veel haken en ogen kent om een coherente verklaring van de wereld vanuit Habermas metafysische voorstel te geven, wil ik eens kijken of we uitgaande van een antireductionistisch fysicalisme, zoals dat van Dennett, niet in staat zijn om aan twee vooronderstellingen te voldoen die Habermas essentieel acht voor het taalspel van verantwoordelijk agency. Deze vooronderstellingen zijn zoals we gezien hebben:

- We verwachten als rationele agents van elkaar dat we ons handelen laten leiden door reflectief gekozen redenen die de onzen zijn.
- Als agents gaan we er vanuit dat we in het kiezen van onze redenen (*ceteris paribus*) alternatieve mogelijkheden hebben.

Dus teneinde het communicatief handelen in een theorie te kunnen rechtvaardigen moet deze de visie rechtvaardigen dat participanten naar redenen handelen en *‘in hun praktische redeneren de beschikbaarheid van alternatieve mogelijkheden.[...] veronderstellen’* We vooronderstellen dus het vermogen van rationaliteit zoals het tot stand komt in een vrije keuze voor de beste reden, waardoor we ons handelen laten leiden. Kan Dennett’s interpretationisme het vermogen rationeel te handelen naar zelfgekozen redenen rechtvaardigen? Ja, Dennett’s aanname [Zie sectie 3.1.2] is dat we vanuit het intentionele perspectief *‘het object waarvan we het gedrag willen voorspellen voor een rationele agent’*⁷⁸ houden. De idee dat we ons handelen laten leiden door redenen is een volkspychologische aanname. We voegen daar Dennett’s opvattingen aan toe dat de volkspychologie, waarin alle mentale concepten als calculerende ‘devices’ optreden, een redelijk betrouwbaar instrument is om te verklaren en voorspellen. Aangenomen dat Dennett ook een bevredigende verklaring voor mentale veroorzaking kan geven (ik kom hier in 6.2 op terug), dan lijkt het dat Dennett’s positie in ieder geval een rechtvaardiging geeft voor de opvatting dat we rationele agents zijn die hun handelen laten leiden door redenen.

De tweede veronderstelling van agents in het taalspel van verantwoordelijke agency is dat ze alternatieve mogelijkheden hebben zodat ze kunnen kiezen voor de beste handelingsturende reden. Agents moeten dus bij het reflectief wegen van redenen kunnen geloven dat er ook onvervalste opties voor hen zijn; dat er verschillende paden voor ze in de toekomst open liggen. Kan Dennett dit vanuit zijn filosofische positie rechtvaardigen? Ja, door een beroep op de beperkte epistemische horizon van agents te doen. Het klopt dat in een gedetermineerde wereld elke mogelijke wereld maar één toekomstverloop heeft. Maar a.g.v. de eindige intelligentie van de agents en de beperkte tot hun beschikking staande informatie, kunnen ze nimmer vaststellen welke van de oneindige verzameling van mogelijke werelden, hun actuele wereld is. Maar hier is Habermas niet van onder de indruk. Zijn respons luidt:

“Om aldus de contradictie tussen een naturalistische wereldvisie en het zelfbegrip van agents te verdrijven, is het niet voldoende om vanuit een ‘bird’s-eye’ gezichtspunt van de wetenschapper op te merken dat zulke alternatieven en gradaties van vrijheid alleen vanuit het perspectief van de participanten kan bestaan. Want ondanks het eindige karakter van de gesitueerde rationaliteit welke hen in staat stelt de toekomst alleen binnen enge grenzen te voorspellen, zijn die agents identiek aan de personen die de wereld als het ware van buiten of van boven bekijken. Wat de een weet, kan de ander simpelweg niet veronachtzamen. [Habermas⁷⁹]”

Habermas voert dit vervolgens terug op de wetenschappelijke aanname dat het objectiverende perspectief van de natuurwetenschappen prioriteit op het participerende perspectief moet hebben. De veronderstelling juist van Dennett dat we als actor een beperkte epistemische horizon hebben, maakt het tegenstrijdig om vervolgens aan te nemen dat het objectiverende perspectief wel in staat is tot een soort van ‘gods-eye view’. Want consequent doorgetrokken zouden we ook voor het objectiverende perspectief moeten concluderen dat het geen directe toegang tot de wereld geeft, en bovendien, aldus Habermas zodanig verstrengeld is met het participerende perspectief dat het niet gerechtvaardigd is het participerende perspectief ondergeschikt te maken. *“We moeten erkennen dat kennisdaden deel uitmaken van een*

context die pas wordt geconstrueerd in de levenspraxis, in het spreken en handelen van nooddrufelige wezens.”⁸⁰

Dit verlegt dus de hele discussie naar de meer fundamentele vraag naar de geldigheid van de filosofische onderbouwing van de stelling dat het participerende perspectief in onze toegang tot de wereld zodanig essentieel is, dat het niet door het waarnemer perspectief overruled mag worden. [Daar kom ik in sectie 6.5 op terug].

Mijn kritiek op Habermas is dat hij in zijn essay nauwelijks in gaat op Dennett's formulering van vrijheid, en Dennett's positie ook niet expliciet weerlegt. Wat Dennett biedt met o.a. zijn speelgoedmodellen is dat (a) er ook een begrip van vrijheid is te geven, welke beslist niet onaantrekkelijk is om te hebben, namelijk vrijheid in de zin van het hebben van het rationele vermogen om de dingen naar je hand te zetten, om het verloop van de toekomst te beïnvloeden, om gevaar te vermijden, en dat (b) het beeld van een actor die vanuit deze vrijheid zijn handelen kan laten leiden door vrij gekozen redenen niet incompatibel hoeft te zijn met een ontologisch en nomologisch fysicalisme. Dat we daarbovenop uiteindelijk niet eens zeker weten in welke mogelijke wereld we zitten zodat we genoeglijk de verbinding tussen determinisme van het fysische perspectief en voorspelling en onvermijdelijkheid vanuit het actorcentrische perspectief kunnen verbreken, is Habermas nog niet voldoende. Ik meen dat Habermas een onredelijk sterke definitie van vrijheid veronderstelt, namelijk dat onze overtuiging van vrije wil in de zin van het hebben van alternatieve mogelijkheden een zekerheid 'beyond any doubt' lijkt te moeten zijn. Maar is dat eigenlijk wel consequent gezien Habermas' opvatting dat de werkelijke wereld feitelijk altijd achter de theoriegeladenheid van de waarneming schuil gaat. Hoe kun je dan kennis verheven boven elke twijfel eisen?

6.2 Habermas versus Dennett m.b.t. mentale veroorzaking

Volgens Habermas zijn redenen niet tot causaal nomologische oorzaken te reduceren. Ze determineren het handelen van subjecten niet op dezelfde starre noodzakelijke wijze als natuurwetten dat voor de oorzaak-gevolg relaties van objecten doen. Habermas is van mening dat (1) als handelingen direct op nomologisch gedetermineerde gebeurtenissen terug te voeren zouden zijn, dan (2) de semantisch gecodeerde propositionele attitudes geheel zouden worden gepasseerd, zodat (3) mentale veroorzaking een fictie zou zijn. Het zou dan onlogisch zijn om te denken dat participanten (door hun performatief handelen en communiceren) mentaal in de wereld zouden kunnen interveniëren. Hoe moet immers een vrije morele agent, die vanuit het participantperspectief overtuigd is van het feit dat deze zich in zijn handelen door zelfgekozen redenen laat leiden, zich herkennen in het objectiverende mensbeeld van de nomologische causaliteit van neurofysiologische processen? Het vanuit het participantperspectief ervaren mentale proces verwordt tot een epifenomeen wanneer we reductie tot een objectiverende verklaring van gedrag in termen van onwrikbare noodzakelijk verloopende neurofysiologische procesketens van oorzaak (stimulus) en gevolg (response) veronderstellen. Volgens Habermas is mentale veroorzaking alleen te begrijpen als we beide perspectieven verenigen in één metatheorie zonder daarbij het objectiverende perspectief te laten domineren, dus zonder een nomologisch, en ontologisch fysicalisme te veronderstellen.

Dennett zal in een respons ook het reductionisme afwijzen. Maar in tegenstelling tot Habermas zal hij argumenteren dat daarvoor het nomologisch en ontologisch fysicalisme niet

verlaten hoeft te worden. Allereerst (ad 1) zal Dennett zijn antireductionisme verdedigen door te stellen dat de volkpsychologie een grofmazige theorie is. Hierin fungeren redenen in termen van mentale toestanden als hypothetische calculerende postulaten die behulpzaam zijn in het voorspellen van gedrag. Je kunt niet de intentionele beschrijvingen, in termen van door redenen bewogen actoren, één-op-één naar de nomologische neurofysiologische processen vertalen. Dennett wijst de type identiteitstheorie af.

Maar Habermas' conclusie, dat in een fysicalistisch wereldbeeld de semantische gecodeerde propositionele attitudes worden gepasseerd waardoor mentale veroorzaking een fictie is, zal Dennett bestrijden. Habermas lijkt er vanuit te gaan (ad 2) dat propositionele attitudes, omdat ze een zekere gerichtheid bezitten, niet in relatie met een fysische correlaat te brengen zijn.

Maar die gerichtheid kun je al veronderstellen bij een microprocessorgestuurde kamerthermostaat. Omdat er zich een causale link gevormd heeft tussen temperatuursensoren in de ruimte en het programma van de processor zou je je kunnen voorstellen dat de elektronica een soort, zij het minimale, interne representatie van de omgeving om haar heen heeft. Dit kan reeds als een vorm van intentionaliteit worden beschouwd⁸¹.

Veronderstellen dat neurofysiologische patronen geen semantische inhoud kunnen hebben is ook een voorbarige conclusie. Er wordt voor wat betreft de verklaring voor het verkrijgen van betekenis, of het nu de semantiek van de omgangstaal of psycho-semantiek betreft, geen principiële belemmering opgeworpen door het verschil in de fysische aard van de drager die het symbool vormt; of het nu klanken, tekens op papier, of neurale patronen zijn.

In een respons op het epifenomenale van het mentale (ad 3) is het mogelijk om aan te tonen dat Habermas zich in de problemen werkt doordat hij telkens het beeld van een soort eigenschapdualisme voor zich heeft. Omdat hij aanneemt dat het domein van het mentale en het fysische simultaan (als een conjunctie) zijn verbonden in één metatheorie, ziet hij zich in dezelfde ontologie geconfronteerd met twee ketens van interactie, namelijk in dat wat aan de oorsprong van enerzijds het mentale, en anderzijds het fysische domein ligt.

Maar dit, wat Ryle omschreef als een 'category mistake', vormt een belemmering voor een coherente natuurwetenschappelijke verklaring voor de relatie geest-lichaam⁸².

Dit probleem ontmoet je niet als je zoals Dennett aanneemt dat je met twee verschillende interpretaties van één en dezelfde werkelijkheid (de ontologie van het antireductionistisch fysicalisme) hebt te maken. In zo'n benadering wordt er geen van elkander te onderscheiden ontologie van het lichamelijke en het geestelijke verondersteld. Je kunt daarin immers niet tegelijkertijd het ene en het andere interpretatiekader innemen⁸³. Uitgaande van Dennett's interpretationisme kunnen we enerzijds vanuit het fysische perspectief een beschrijvingen van alle fysische (waaronder neurofysiologische) processen en hun causale relaties geven. Anderzijds kunnen we het intentionele perspectief innemen van waaruit we een beschrijving geven in de taal van de volkpsychologie. En ook vanuit dit perspectief kunnen we causaliteit herkennen⁸⁴. Door aldus uit te gaan van twee verschillende perspectieven die niet op hetzelfde tijdstip kunnen worden ingenomen, ontloop je het probleem van de zgn. schending van de wetten van de thermodynamica, en daarmee het epifenomenalisme van het mentale. Door ook nog de één-op-één relatie tussen mentale toestanden en neurofysiologische toestanden te verwerpen, kun je de dwingende beschrijvingen van de causaal nomologisch verlopende processen van het fysische perspectief, niet zonder meer overzetten op, of vergelijken met de vrijere volkpsychologische beschrijvingen vanuit het intentionele perspectief.

Een tweede verdediging van mentale veroorzaking komt uit Dennett's bewustzijnstheorie. In antwoord op bezwaren als zou het bewustzijn een epifenomeen zijn, stelt hij dat we door middel van praten tegen onszelf de ervoor verantwoordelijke neurale processen de hersenprocessen laten domineren en zo de leiding over ons handelen geven.

6.3 Habermas' kritiek op memetiek en Universeel Darwinisme.

Habermas bekritiseert de universeel darwinistische aanpak van Dennett. De memetiek doet volgens Habermas geen recht aan het participierend perspectief op het communicatief handelen. Vanuit deze instrumenteel rationele atomistische benadering, die voornamelijk gebaseerd is op het principe van macht, en maximaliseren van individueel profijt, blijven de meer holistische aspecten van normativiteit consensus en communicatieve rationaliteit onzichtbaar.

“Het is precies vanuit het door Dennett ingenomen standpunt van de evolutietheorie, (en ontwikkeld tot buiten de grenzen van de evolutionaire biologie) dat dit ‘sciëntisme’ ophoudt zelfevident te zijn. De eindigheid van een geëvolueerde en ingebedde geest gaat in tegen het, ondergeschikt maken van de rol van het participantperspectief (van waaruit de objectieve wereld ons in onze dagelijks omgang confronteert met oncontroleerbare contingenties), aan een transcendentiaal standpunt buiten de wereld ...Het is juist de idee van een transcendente, finale, ultieme beschrijving van het universum wat inconsistent is, ..op basis waarvan .. ons gevoel van vrijheid een fictie wordt. Zodra men ziet dat de capaciteit voor het participeren in rechtvaardigende praktijken essentieel wordt voor alle kennis – en voor elke intelligibele conceptie van kennis – stort deze constructie in elkaar” [Habermas⁸⁵]

Kenmerkend voor dit soort empiristische modellen is dat ze geen normatieve uitspraken doen. Ze geven slechts een beschrijvende analyse vanuit het objectiverende perspectief, die moet aantonen hoe de ontwikkeling tot stand is gekomen. Het centrale uitgangspunt is een telkens door louter instrumentele rationaliteit aangedreven machtsstrijd van de actoren die uitzijn op het nastreven van eigenbelang⁸⁶. Typerend is het veelvuldig gebruikmaken van speltheorieën waarin actoren worden opgevat als losse rivaliserende atomen, calculerende wezens die op rationele wijze uit zijn op het maximaliseren van hun eigen profijt. Habermas meent dat het onterecht is om de ontwikkeling van geest en cultuur te typeren als een blinde machtsstrijd waarin eigenbelangen worden nagestreefd. Dit schetst een te eenzijdig beeld. De veronderstelling dat mensen slechts calculerende subjecten zijn die hun eigen nut willen maximaliseren, miskent de mogelijkheid van het tot ontwikkeling komen van een ander meer omvattende rationaliteit. Habermas' theorie van het communicatief handelen wil juist duidelijk maken dat mensen wel degelijk hun particuliere belangen en visies kunnen overstijgen. Door redelijk discussies te voeren zijn mensen samen met anderen in staat rekening te houden met de individuele belangen van iedereen. De mens heeft zich ontwikkeld in gemeenschappen en handelt daarbij niet louter uit eigenbelang.

Een bijkomend bezwaar tegen de memetiek is dat het de autonomie van het intentionele individu tot een zowat passieve (over)drager, een gastheer van informatiepakketjes reduceert. Als grote afwezige schittert in dit verhaal het perspectief van de intentioneel optredende participant in de taaldaad rechtvaardigende sociale context van verdediging door argumenteren. Maar zoals Dennett wel zegt, er moet nog steeds gedacht worden.

“Memen zijn afhankelijk van het menselijk brein als hun tehuis; de menselijke nieren of longen zijn geen geschikte alternatieven, want memen zijn afhankelijk van de denkkraft van hun gastheren...Als memen gereedschappen zijn voor het denken, moeten ze worden gebruikt om hun fenotypische resultaten te laten zien. U moet dus nog steeds denken....Het is waar dat een goed darwinistische denkmodel er niet net zo uitziet als de traditionele modellen. We moeten echt het slechte oude cartesiaanse model van een centrale, niet mechanische *res cogitans*, [*denkend ding / één enkele opzichter; de homunculus*] dat het serieuze geestelijke werk doet, vervangen....al het denkwerk moet worden verdeeld over minder fantastische actoren [*vele concurrerende eenvoudige homunculi zonder opzichter*] [Dennett⁸⁷]”

Dennett is vooral begaan met de zorg of de ‘fantasie’ van onze filosofische opvatting wel de toets van kritische reflectie op de mogelijkheid van een fysische realisatie kan doorstaan. Ook hier geldt dus weer zijn devies ‘geen hemelhaken maar hijskranen’. Ik kan me wel aansluiten bij de idee dat we in onze filosofische bespiegelingen contact moeten blijven houden met de (natuur)wetenschap. Dat neemt niet weg dat ik tegelijkertijd ook Habermas’ kritiek op het sociaal-atomistische karakter van de memetiek en de speltheorie deel. Ik mis daarin de rol van gevoelens van zorg voor de medemens (empathie), en andere hermeneutisch aspecten die ontegenzeggelijk een rol spelen in het menselijk gedrag binnen een samenleving. Maar ik betwijfel Habermas conclusie dat het primaat van het objectiverende denken noodzakelijk tot een blindheid voor normatieve kwesties zou moeten leiden. Daar kom ik in de komende secties, waarin normativiteit en moraal centraal staat, op terug.

6.4 Habermas versus Dennett m.b.t. ontstaan van morele verantwoordelijkheid

De opvattingen van Habermas en Dennett betreffende moraal en verantwoordelijkheid, wortelen in twee verschillende moraalfilosofische tradities gevestigd door resp. Immanuel Kant en David Hume.

Kant begint met de constatering dat de mens ontegenzeggelijke een moreel wezen is. Ofschoon we vaak verleid worden ons handelen door sentimenten en verlangens te laten leiden, kunnen we in ons morele denken en handelen zulke motieven weerstaan. Kant argumenteerde dat het niet het verstand (Habermas’ instrumentele rationaliteit) kan zijn wat met het morele denken moet worden geïdentificeerd. Dat constitueert immers met haar categorieën de objectiverende wereld van de causaal nomologische natuur, en biedt volgens hem daarom niet de vrijheid die voor het moreel handelen is vereist. Het de praktische rede (Habermas communicatieve rationaliteit) die ons handelen als morele agents bepaalt, en de verleiding door de emoties (om te handelen in eigenbelang) moet weerstaan. Omdat slechts een rationeel principe en niet een verlangen of emotie de goede handeling kan bepalen, moet in het morele handelen de rede de meester over de emotie zijn. Dit rationele principe is het categorisch imperatief: ‘handel zodanig dat je zou willen dat het maxime van jouw handelen een universele wet is’, een principe wat een redelijk individu zal kiezen als het zichzelf de morele wet stelt, en wat door haar universaliteit onpartijdigheid garandeert.

Heeft bij Kant de rede de potentie om ons te doen laten handelen, zo niet bij Hume. De ratio is voor hem inert. Elke impuls tot handelen moet uit de emotie komen. In een moreel oordeel werken de rede en sentiment samen. De rede kan weliswaar instrumenteel ons oordeel sturen, door ons bijvoorbeeld informatie te geven, maar verlegt daarmee slechts de drijfveer van de ene naar de andere emotie. Een moreel oordeel is een oordeel dat gebaseerd is op de emotie van goed- of afkeuring afkomstig uit het empathisch mechanisme van de sympathie. Daarmee

is voor Hume de moraal niet zoals bij Kant in de ratio doch in een natuurlijke drift (de emotie) geworteld.

Deze verschillen in opvatting vormen aanzienlijke consequenties voor de cognitieve toegang tot de ultieme bron van de moraal. Door moraal in de emotie te funderen ontkent Hume dat de ultieme grond in de rede (anders dan de instrumentele rede) is te vinden. Daarmee plaatst Hume de bron van de moraal buiten het bereik van demonstratie, kennis en waarheid. Morele oordelen kunnen niet exclusief appelleren aan feiten en aan rationele demonstratie van (in)correctheid.

Dat daarmee Hume's moraaltheorie volledig subjectivistisch zou zijn, en elke grond van universaliteit zou missen, is een onjuiste conclusie. Ten eerste beargumenteert Hume namelijk dat sommige van de sentimenten waarin de moraal is gefundeerd, sentimenten zijn die door alle mensen worden gedeeld. Ten tweede kunnen we volgens Hume, ondanks dat de rede een aan de emotie ondergeschikte rol krijgt toebedeeld, "goede redenen" voor het handelen hebben. Goede redenen kunnen worden begrepen in termen van geschikte motieven en goed bewijs in de vorm van relevante empirische feiten.

Een van Humes' premissen is dat de regels voor hetgeen we behoren te doen in de morele instituties zijn uitgekristalliseerd. De regels van deze instituties zijn een zaak van consensus in het sociale domein. Dus afgezien van de psychologische gronden van de moraal, vormt de vaststelling van de morele belangen door de gemeenschap, de enige basis voor moraliteit. De morele regels zijn dan niet louter de formulering van wat een individu voelt. Het zijn gefixeerde punten in een culturele matrix van richtlijnen en besturingen, die volgen uit de gemeenschappelijke menselijke natuur.

Nu we de traditionele bronnen van de meta-ethische opvattingen van Habermas en Dennett hebben gezien, gaan we kijken hoe ze deze hebben aangepast en uitgewerkt.

De in de jaren 80 van de voorbij eeuw door Habermas uitgewerkte moraalfilosofie heeft de naam discoursethiek gekregen. Het gaat om een bewerking van zijn theorie voor het communicatief handelen die op de morele context is toegespitst. Ons handelen moet worden geleid door morele normen die het resultaat zijn van een redelijke morele discussie tussen alle betrokkenen. De overeenstemming die in de discussie wordt bereikt, garandeert daarbij de geldigheid van de normen in kwestie. Dit noemt Habermas het discoursprincipe waarin wordt aangegeven dat handelingsnormen het voorwerp van consensus moeten zijn: "*(D) alleen die normen kunnen aanspraak maken op geldigheid, die in praktische discourses op de instemming van alle betrokkenen kunnen rekenen.*" Hoe die overeenstemming wordt bereikt wordt aangegeven door het aanvullende moraalprincipe: "*(U) Een norm is slechts dan geldig als de bij de algemene naleving ervan te voorziene gevolgen en nevenwerkingen voor de belangen en de waarden van een ieder door alle betrokkenen gezamenlijk en zonder dwang aanvaard kunnen worden*"⁸⁸.

De bron van het morele handelen wordt net als in die van Kant ook in deze theorie in de praktische rede gevonden. Dat maakt nu ook duidelijk waarom Habermas vrije wil expliciet koppelt aan de vooronderstelling dat morele agents worden bewogen door redenen. Moraliteit in deze neo-Kantiaanse zin vooronderstelt dat we redelijk handelen. Dat is de wijze waarop onze morele oordelen, normen en handelingen tot stand behoren te komen. De veronderstelde redelijkheid van de deelnemende participanten wordt gegarandeerd doordat ze de ongedwongen dwang van het beste argument accepteren. Daarmee is de discoursethiek deontologische en is er net zoals in de Kantiaanse moraaltheorie sprake van een zichzelf

opgelegde plicht, en wordt onpartijdigheid of universaliseerbaarheid van normen centraal gesteld.

Maar in plaats van een hypothetische instemming, zoals bij Kant's Categorisch Imperatief, moet de onpartijdigheid van een norm bij Habermas worden vastgesteld in een werkelijke instemming van alle participanten in een moreel discours. In het feitelijk debat, als een proces van rolovername, moet men zich inleven in het perspectief van de gesitueerde concrete ander. De discoursethiek is daarmee een formele moraaltheorie, dwz geeft alleen de formele procedures waarin morele normen tot stand komen, maar vertelt over de inhoud van deze normen verder niets. Voordeel is dat het toegeeft aan particulariteit van deelnemers in hun sociale context. Morele normen worden hier dus niet door een filosoof vanuit zijn leunstoel op een stoffig zolderkamertje gedefinieerd. Om tot formele normen te komen is een feitelijke deliberatie tussen alle betrokkenen nodig. Dit maakt de moraaltheorie in democratisch opzicht aantrekkelijk. Idealiter wordt er zo met ieders particuliere perspectief rekening gehouden.

Maar hoe verhoudt de discoursethiek zich t.o.v. moraaltheorieën die de moraal in de emotie funderen? Als een cognitivistische moraaltheorie is de morele argumentatie het uitgangspunt van de discoursethiek. Dit is niet compatibel met het emotivisme van Russel en Ayer. Deze reduceert morele normen tot uitingen van morele gevoelens en zal derhalve argumentatie als deze betrekking heeft op de bron van de moraal als een dwaling opvatten.

De discoursethiek is ook niet compatibel met de contracttheorie. Daarin worden mensen louter opgevat als actoren die hun eigenbelang willen maximaliseren. Habermas meent dat we niet langer in staat zijn het verplichtende karakter van normen te verklaren als we morele redelijkheid tot een dergelijke instrumentele rationaliteit reduceren. Hoe zal Habermas oordelen over de Humeiaanse moraaltheorie? Ik vermoedt dat hij goedkeurt dat Hume de regels van de moraal in regels van morele instituties plaatst die door consensus zijn bereikt. Maar Habermas zal zich niet kunnen vinden in de idee dat de moraal gefundeerd is in het gevoel, omdat daarmee de bron van de moraal buiten het bereik komt te liggen van morele argumentatie.

Maar de bron van de moraal in de emotie, of wel 'fellow-feeling' dwz de empathie, opent wel een ander pad waarlangs consensus kan plaatsvinden. Doordat mensen met elkaar delibereren, hun eigen positie uit de doeken doen, en naar die van de ander luisteren, kan er een toestand ontstaan waarin de actoren zich door empathie emotioneel verplaatsen in de ander. Dit kan de kloof tussen participanten in conflict verkleinen. Overeenstemming in morele kwesties komt vanuit deze visie niet primair tot stand op grond van de kracht van argumenten maar door het opwekken van gevoelens van medemenselijkheid. Het door Habermas geschetste pad langs de weg van de soevereine rede is dus niet de enige weg waarlangs consensus betreffende morele kwesties kan worden bereikt. Het hangt er maar vanaf wat je aanneemt als ultieme bron van de moraal.

Laten we nu ter vergelijking eens kijken naar Dennett's reconstructie van het ontstaan van moraal, vrijheid en morele verantwoordelijkheid. De totstandkoming uit de quasi moraal van de volledige moraal van een autonome zelf wordt door hem met behulp van de verrijzende cultuur verklaard.

Zoals gezegd neemt Dennett in navolging van David Hume de natuurlijke drijfveren als uitgangspunt voor de moraal. Deze drijfveren zijn: liefde voor kinderen, beperkte liefdadigheid, belangstelling, wrok en seksuele begeerte. Dennett schets een weg waarlangs

we met behulp van de door de natuur aan ons geschonken reflectie, de natuurlijke instincten kunnen hebben aangepast en uitgebreid tot kunstmatige deugden van de moraal, zoals gerechtigdheid.

Habermas' bezwaar tegen de toepassing van de memetiek als verspreider van culturele informatie heb ik eerder besproken. Ik wil nu kijken naar de andere aspecten van Dennett's benadering door me te concentreren op de laatste stap die hij maakt. Hij kijkt hoe wij als psychische ingenieurs het fundamentele concept van een verantwoordelijke morele actor kunnen waarborgen in een tijdperk waarin wetenschappelijke studies naar de aard van ons handelen ook de nodige resultaten produceren.

In tegenstelling tot Habermas verwerpt Dennett deze resultaten niet, maar omarmt ze als belangrijke bron om meer over onszelf te weten te komen. Vanuit een naturalistische houding corrigeert hij wel regelmatig de interpretaties van onderzoeksresultaten als ze aanleiding geven tot onrechtmatige conclusies over ons mensbeeld. Bijvoorbeeld wanneer in zijn ogen ten onrechte het beeld van de mens als irrationeel wezen (wiens handelen aan diens bewuste controle ontsnapt) uit neurowetenschappelijke en psychologische onderzoeksresultaten wordt afgeleid.

Dennett corrigeert dit beeld aan de hand van zijn eigen bewustzijnstheorie. Er is wel degelijk vóór mentale veroorzaking te argumenteren als je de idee van het Cartesiaans ego loslaat. Dennett stelt voor dat we bij het loslaten van het door Kant veronderstelde mythologische beeld van de volledige autonome onveroorzaakte individu, we nu niet door moeten slaan naar een door irrationaliteit gedomineerde zelfbeeld. We houden volgens hem nog steeds een rationele zelf over die de moeite waard is.

We moeten stoppen met die vrije-wil-paniek en (uit angst voor het spook van het determinisme) de verwerping van alles wat uit de wetenschap aan onderzoeksresultaten over onze menselijke cognitie beschikbaar komt. Als we blijven vasthouden aan een geest die onaantastbaar maar mysterieus is, verliezen we tegelijkertijd de mogelijkheid om de morele normen aan te passen aan een op reëlere vermogens gebaseerd mensbeeld. Voor alles moeten we zien dat er vrijheid in onze rationele vermogens kan worden gevonden. We kunnen het, dwz. ons aanpassen aan onze sociale omgeving. We hebben het vermogen om te voldoen aan de eisen die in het kader van de aan ons geschonken verantwoordelijkheid, aan ons zijn opgelegd. Nee, we hebben niet die onveroorzaakte zelf zoals in het geconcentreerde tijdstip van een zelfvormende handeling van Kane⁸⁹, maar als we ons zelf groot genoeg maken in tijd en ruimte dan blijkt er wel degelijk voldoende mogelijkheid om van amoreel vrij kind morele actor te worden.

Met de factoren geluk, een omgevingskader, wat hulp van vrienden en de loop der tijd, groeien we naar een moreel actorschap. Het begint met geluk, de ontvangen genen, en een stabiel gezin om in op te groeien. Hoewel geluk niet binnen onze controlesfeer is gelegen, is de marathon lang en is een klein verschil in startpositie derhalve niet onrechtvaardig. En als we de wetenschap niet buiten spel zetten (iets wat we soms onrechtmatig doen om onze mythologische zelf in bescherming te nemen), kunnen we haar kennis juist gebruiken om te compenseren bij diegenen waar het geluk aan voorbij gegaan is. Het morele actorschap gaat om handelen naar door ons gekozen redenen, en om het afleggen van verantwoording voor (de redenen van) ons gedrag. Maar we worden hierop voorbereid door in onze opvoeding aangemoedigd te worden om te communiceren en om mee te doen aan onze gewoontes om redenen te vragen en te geven en vervolgens te beargumenteren waar we mee bezig zijn. Aldus bouwt Dennett aan een constructie die ook één van de pijlers vormt voor de

definitie van Habermas' moreel actorschap; het wegen van redenen en bewogen worden door ons toegeëigende redenen.

De kern van de zaak is dat we goed moeten zien wat Dennett hier telkens met vrijheid bedoelt. Niet de vrijheid van alternatieve mogelijkheden in de enge zin (een wonderbaarlijk vermogen om een gevorkte toekomst te hebben in één mogelijke wereld), maar de vrijheid die samen gaat met onze rationele vermogens. Het is de extra vrijheid die de betere schaakcomputer heeft omdat hij rokeert waar de mindere computer had kunnen rokeren maar het niet deed. Het vermogen gevaar af te wenden. Een weg te begaan die niet mogelijk zou zijn als we de verstandelijke vermogens zouden missen. Het is de viering van ons rationeel vermogen. Het is de vrije wil die het waard is om te hebben. Nog eenmaal: om te begrijpen en aan te voelen over welk soort van vrijheid Dennett het hier heeft. In 'Elbow room' beschrijft Dennett het mechanisme van de graafwesp Sphex. Dit insect volgt bij de voorbereiding van het eieren leggen een serie van geprogrammeerde stappen. Als een onderzoeker één van de stappen onderbreekt dan herhaalt de wesp deze stap. Voor een dier als een wesp kan dit proces van zinloze gedragsherhaling eindeloos doorgaan. De wesp lijkt niet te beseffen wat er gaande is. Dit is het soort van gedachteloos gepredetermineerd gedrag wat wij mensen kunnen vermijden. Mensen kunnen de nuttelosheid van eindeloos herhaald gedrag inzien en door een daad van vrije wil iets anders doen. Beschouw dit als een operationele definitie van vrije wil.

Dennett meent dat als we onszelf willen definiëren in het licht van wat we wetenschappelijk gezien over onszelf te weten zijn gekomen, we dan pragmatisch kijken naar wat we hebben en op basis daarvan de idee van verantwoordelijkheid en moreel agency vestigen. Voor filosofen die de mythische onveroorzaakte zelf met metafysische luchtkastelen en 'skyhooks' willen redden heeft hij de volgende boodschap:

"Oefeningen in ethisch theoretiseren die weigeren te buigen voor de empirische feiten over de menselijke situatie moeten wel fantasieën voortbrengen; misschien van enig esthetisch belang, maar als praktische aanbevelingen niet serieus te nemen." [Dennett⁹⁰]

Maar voor de streng deterministen, de extremisten aan de andere zijde van het vrije wil debat, heeft hij ook een boodschap. Waar zij roepen dat, nu gebleken is dat vrije wil niet bestaat het goed zal zijn om nu ook maar met de moraal af te rekenen, werpt Dennett tegen dat los van het feit dat determinisme verantwoordelijkheid niet uitsluit, schuld en woede goed harmoniëren. Dreiging van schuld helpt handelingen voorkomen die het oproepen.

In antwoord op de vraag hoe we een verantwoordelijkheidsbegrip kunnen construeren wat sterk genoeg is om alle nieuwigheden uit de wetenschap te weerstaan, doorloopt Dennett een aantal stappen.

Allereerst kiest hij voor Harry Frankfurt's theorie van eerste, en tweede-orde wensen. De tweede-orde wensen hebben de eerste orde wensen als inhoud. Ze keuren de eerste orde wensen goed of af. Volgens Frankfurt is het de rol van de actor om na te denken over motieven die om de heerschappij van zijn gedrag strijden, en de uitkomst van sommige motieven boven andere te verkiezen. En wie of wat verkiest hier dan de ene motieven boven de anderen? Niet het cartesiaans ego, een res cogitans, maar de rede zelf. En waar komt dit vermogen van de rede vandaan? Uit de opvoeding die ervoor zorgt dat een kind zich bezighoudt met redenen vragen en geven.

“Het zelf is het systeem dat mettertijd verantwoordelijkheid wordt gegeven, zodat je er op kan vertrouwen dat er iemand verantwoordelijkheid neemt, zodat er iemand thuis is wanneer er vragen over aansprakelijkheid opkomen. Kane en de anderen hebben gelijk om te zoeken naar een plek waar de verantwoordelijkheid niet verder kan worden afgeschoven. Ze hebben alleen naar het verkeerde gezocht.” [Dennett⁹¹]

De tweede stap die Dennett in zijn constructie van een morele agent onderneemt, is zich afvragen wat het dragen van verantwoordelijkheid inhoudt. Het antwoord luidt: verantwoordelijkheid krijgen toegewezen betekent niet alleen gestraft worden voor de verkeerde dingen die je doet, maar ook lof krijgen voor de goede dingen die je verricht. Het is een ‘package deal’; of je neemt het beiden of je krijgt geen van twee. Het aanbod wat we elkaar telkens doen is “als je vrij wilt zijn, moet je je verantwoordelijkheid nemen”. En dat is fair als je ervan uitgaat dat we inderdaad het rationele vermogen hebben om te voldoen aan de sociale normen die we elkaar opleggen. Blijft natuurlijk de vraag wat je doet met die arme drommels die dit niet lukt. We moeten ergens een grens trekken tussen diegenen die we wel en die we niet verantwoordelijk voor hun handelen kunnen houden. De grens die we hier trekken is er een waarvan gebleken is dat die meeschuift met hetgeen we leren uit wetenschappelijk onderzoek betreffende de menselijke besluitvorming. Dennett’s advies past bij zijn naturalisme. “Het gevaar dat een verantwoordelijkheidsbegrip niet overeind blijft is uiteindelijk veel groter als we de systemen op mythes baseren.” Onder mythes verstaat Dennett hier ook metafysica. Bij deze opstelling sluit het voorstel van Stephen White in “The unity of the Self” goed aan. White adviseert geen metafysica te gebruiken om de ethiek op te baseren maar “doe het andersom; bepaal wat we behoren te bedoelen met ons ‘metafysisch criterium”.

Laat om te beginnen eerst zien hoe er een interne rechtvaardiging kan zijn voor iemand die in zijn straf berust, en feitelijk zegt ‘bedankt dit had ik nodig!’, en gebruik dat begrip vervolgens om de lezing van onze cruciale uitdrukking ‘een actor had iets anders kunnen doen dan hij of zij gedaan heeft’ vast te leggen en te ondersteunen: “Een actor had iets anders kunnen doen dan hij of zij gedaan heeft in het geval dat het toerekenen van verantwoordelijkheid en schuld aan die actor voor de onderhavige reactie gerechtvaardigd is⁹².”

Het feit dat vrije wil wenselijk is kan gebruikt worden om ons concept van vrije wil te verankeren op een manier waarop metafysische mythes dat niet kunnen. De ideale strafinstelling volgens White is dat elke straf in de ogen van de bestrafte gerechtvaardigd is (en antwoord met ‘bedankt dit had ik nodig!’). Dit vereist een zekere intelligentie van die personen die voor straf in aanmerking komen, willen ze terzake kundig genoeg zijn om bekwame rechters van die bedoelde rechtvaardiging te zijn. Hun (verbeelde) berusting dient dan als referentie of draaipunt om de plaats en hoogte van de drempel te bepalen tussen verantwoordelijk en niet-verantwoordelijk voor eigen handelen. Dus als we de verschillende mogelijke reacties nalopen van de beschuldigde van een delict die zijn straf krijgt opgelegd, dan kunnen we het volgende onderscheid aanbrenge:

- Zij die dit kunnen begrijpen zijn onproblematische voorbeelden van schuldige boosdoeners. Ze zeggen het zelf, dus hen houden we verantwoordelijk.
- Zij die dit niet kunnen begrijpen zijn zeker niet bekwam genoeg om zonder toezicht van de vrijheid van het burgerschap te genieten, dus stellen we ze niet verantwoordelijk.

- En dan heb je nog degenen die klaarblijkelijk bekwaam zijn maar zich tegen de straf verzetten. Ze zeggen daarmee impliciet dat ze wel verantwoordelijk willen zijn, dat ze de voordelen van bekwaam verantwoordelijk burgerschap willen, maar tegelijkertijd onthullen ze dat ze te klein zijn (Als je je echt klein maakt kun je alles externaliseren).

Deze strategie lijkt veel beter te werken dan het antwoord ‘had je anders kunnen doen’ op de vraag of je verantwoordelijk bent. We zoeken dan geen getuigenis bij een metafysicus of een kwantumfysicus, maar naar specifieke bewijzen van iemands bekwaamheid, of naar verzachtende omstandigheden. De staat nodigt je uit te berusten in je straf. Je hoeft er natuurlijk niet in te berusten, maar als de staat zijn werk goed heeft gedaan, behoor je het te doen. Dwz. de staat kan je redenen geven die zij zonder te behoeven blozen kan verdedigen. Als je het niet begrijpt dan is dat jouw probleem. Begrijpen een heleboel mensen het niet dan is dat het probleem van de staat die de drempel te laag heeft gemaakt.

Het principe is dus behoren impliceert kunnen, (ought → can), waarbij kunnen in de ‘ruime’ en niet de ‘enge’ zin wordt opgevat. Dus zoals in ‘de computer had kunnen rokeren’. In de relevante betekenis hadden actoren iets anders kunnen doen. Dat is dan hetgeen ze hadden behoren te doen.

Zo hebben we dan een verantwoordelijkheidsbegrip gecreëerd dat geen beroep op metafysica doet. In plaats daarvan kijken we naar het algemene beeld van iemands vermogens, en naar de specifieke omstandigheden van de situatie waarin de overtreding is begaan, om vast te stellen of iemand ‘anders had kunnen doen’. Feitelijk haalt Dennett met deze constructie de vraag naar “had iemand anders kunnen doen” weg uit het domein van de filosofisch metafysische haarkloverij en plaats het pragmatisch in de alledaagse praktijk. Hij erkent daarmee dat we een theorie nodig hebben die onze praktische problemen m.b.t. vrije wil en verantwoordelijkheid kunnen oplossen.

Hoe zit dat dan met de pragmatiek van de oplossing van Habermas? Voor een beantwoording van deze vraag verplaatsen we ons eens in het perspectief van Dennett, en beoordelen van daaruit Habermas’ onderneming om het zelfbeeld van de mens t.o.v. de wetenschappelijke onderzoeksresultaten te positioneren. We zouden dan kunnen zeggen dat wat Habermas doet, hoe goed dan ook bedoeld, toch enigszins gekunsteld aandoet.

Habermas blijkt vast te willen houden aan een sterke definitie van vrije wil (ceteris paribus alternatieve mogelijkheden hebben, ofwel een gevorkt toekomstperspectief binnen onze actuele wereld) en aan een zelf welke ontrokken is aan de causale orde van de natuur. Om dit te bewerkstelligen is Habermas bereid een sterke metafysica te adopteren waarmee hij net als Kant een van causaliteit ontkoppelde zelf wil blijven koesteren. Maar dat brengt ‘m in problemen waarvan er enkele zijn besproken in sectie 2.4. Niet moet worden vergeten dat Habermas het causaliteitsbegrip nodig blijft houden om zijn emancipatoire programma te kunnen verdedigen. Want zo kan bijvoorbeeld vrijheid van politieke dwang (bijvoorbeeld door fysieke onderdrukking) niet worden gedefinieerd zonder daarin een beroep te doen op het causaliteitsbegrip. In de relatie vrije wil met fysisch determinisme het causaliteitsbegrip vervolgens te clausuleren doet dan aan als een ad hoc oplossing.

Maar het beeld van een agent die in vrijheid wordt bewogen door redenen om met andere consensus te vinden over het beste argument in een discours betreffende de aanvaarding van morele normen, is te combineren met een antireductionistisch fysikalisme als men maar bereid is om twee dingen los te laten; (1) de fixatie op de enge manier van ‘had ook anders kunnen handelen’, en (2) de enge blik in de zoektocht naar één moment in de tijd, één plaats

in de ruimte, waar een zelf wordt gedefinieerd die de volledige controle heeft. Als je daar toch naar op zoek blijft gaan dan moet je je realiseren dat, om met Dennett te spreken, ‘als je jezelf maar klein genoeg maakt, je altijd alles kunt externaliseren!’. Op die wijze krijg je nooit een degelijk verantwoordelijkheidsbegrip van de grond, en hoef je nooit verantwoordelijk voor iets te zijn.

En wat is er eigenlijk beter aan zo’n mythologische zelfbeeld dan het beeld wat ontstaat uit correct geïnterpreteerde wetenschappelijke resultaten van onderzoeken verricht naar onze cognitieve vermogens en onze conaties (natuurlijke driften). Bedenk wat we met een beter wetenschappelijk inzicht in ons zelf op voorhand kunnen vermijden wat zonder correctie tot veel leed en penitentiaire correcties achteraf zou leiden. Die kans laten we liggen als we de wetenschap in toom houden om onze mythologische maar mysterieuze zelf in stand te houden.

Resumerend betreffende de opvattingen van Habermas en Dennett aangaande moraal en vrije wil, staan hiertegenover elkaar resp. een Kantiaans deontologische en een utilistische moraal. Overigens interessant is te zien dat Habermas en Dennett dezelfde uitgangspunten voor de criteria van het verantwoordelijk actorschap proberen te vestigen: een agent die vrij kan kiezen voor de redenen om zijn handelen door te laten leiden. Waar ze van mening over verschillen is wat ze onder vrije wil verstaan, en meer in het algemeen over de epistemische status die moet worden toegekend aan het objectiverende perspectief.

Habermas meent vanuit de Kantiaanse traditie dat mensen zich niet in het beeld kunnen vinden waarin ze worden beschouwd als objecten. In de visie van Kant en ook Habermas heeft het verstand resp. de instrumentele rationaliteit van het objectiverende perspectief betrekking op manipuleerbare objecten die ingezet kunnen worden als middel voor het bereiken van een doel wat het eigenbelang dient. Kant’s categorisch imperatief verbod mensen te behandelen als middel. Werden mensen gereduceerd tot middel, tot instrument, dan werd mensen het vermogen ontnomen de vrije wil uit te oefenen om zichzelf de morele wet te kunnen stellen. Die gedachte van Kant lijkt ook Habermas nog te sturen in zijn verlangen om de mens een uitzonderlijke status te geven in de natuur.

Dit is een nobele gedachte. Dit Roept dan ook de vraag op of we, indien we alleen het objectiverende perspectief hanteren, de waardigheid van de mens geen geweld aan doen. Daarmee verband houdende vragen zijn; in hoeverre zijn we gewapend met alleen het objectiverende perspectief in staat om normatieve uitspraken te doen, kunnen we een kritisch standpunt ten opzichte van maatschappelijke standpunten innemen etc. Dit zijn kwesties die ik bewaar voor de laatste sectie. Nu komt nog eenmaal de overweging van de epistemische status van het objectiverende-, en participerende perspectief aan bod.

6.5 Over de noodzaak van het participerend perspectief!

Het is inmiddels duidelijk geworden dat Habermas door de invoering van een eerste filosofie, en de status die hij aan het participerende perspectief verleent, de oppermacht van de objectiverende (natuur)wetenschappen betwist. Dennett schrijft er echter het hoogste epistemische gezag aan toe. We zullen dit verschil in opvatting eens nader beschouwen.

Volgens Dennett’s fysicalisme is de wereld geconstitueerd door louter fysische processen. Een Laplaciaanse demon kan de wereld volledig beschrijven in fysische termen. Hiermee toont Dennett zich een erfgenaam van de empiristische traditie gevormd door Locke en

Hume. David Hume was een naturalist wiens primaire filosofische doel het was om ons zelf weer te geven als deel van de natuurlijke orde. Hij meende dat zelfkennis tot stand kon komen door middel van wetenschappelijk begrip van de menselijke aard, op een wijze die gelijk is aan de Newtoniaanse natuurwetenschap. Zijn opvattingen weergegeven in de "Treatise" laten zien dat we in grond wezens van instincten en gewoonten zijn, wiens mentale levens eerder worden gedomineerd door de passies dan door de rede, wiens overtuigingen worden gevormd door de mechanismen van associatie en gebruiken, in plaats van door een a-priori reflectie, en waarvan de morele levens het product zijn van gevoelens die getraind zijn door conventies. "Dat we zo in elkaar zitten is een feit en dat kunnen we niet veranderen noch verklaren." Rationalistische systemen misinterpreteren onze aard verkeerd waar ze suggereren dat de menselijke ziel vreemd is aan de wereld waarin ze zichzelf aantreft.

Hume's epistemologie is een gedetailleerde toepassing van zijn claim dat overtuigingen meer een act van de zintuiglijke dan van de cognitieve delen van onze aard zijn. De inhoud van de geest noemde hij 'percepties', die hij onderverdeelde in impressies (sensaties en gevoelens) en ideeën (de kopieën of beelden hiervan). De geest heeft zijn eigen wetten, namelijk die van associatie, waar ideeën uit elkaar en uit de impressies voortkomen, en waar de gevoelens (of passies) uit elkaar voortkomen, zoals genot naar liefde en liefde naar welwillendheid leidt. Onze belangrijkste overtuigingen zijn de producten van deze associatieve wetten van de verbeelding, en niet van de rede zoals rationalistische filosofen menen.⁹³

In tegenstelling tot Hume en Locke verwerpt Dennett (overigens net als Habermas) het cartesians mentalisme waarin de idee (Dennett noemt het cartesians theater) als één-op-één kopie van de werkelijkheid, centraal staat.

Daarmee kan voor Dennett de cognitieve toegang tot de wereld louter gerechtvaardigd worden door de relatie subject-object. Ook de relatie tussen subjecten onderling en de relatie tussen het subject en zichzelf, (dat wat aanleiding geeft tot zelfkennis) verloopt via dit pad.

De vermeende geprivilegieerd toegang tot een subjectief innerlijk, via introspectie (het eerste-persoonperspectief) wordt door Dennett afgewezen. Zelfkennis vindt eveneens plaats via de objectiverende waarneming.

Dit is herleidbaar tot Dennett's theorieën over het bewustzijn en intentionaliteit. In zijn theorie van bewustzijn is het zelf niet een geestesoog, maar een narratieve gebruikersillusie. En in zijn theorie betreffende intentionaliteit worden mentale toestanden even zogoed toegeschreven aan objecten als subjecten, niet omdat ze verwijzen naar iets in hun innerlijke toestand maar omdat ze een relatie hebben (verwijzen is daarbij het verkeerde woord,) met bepaalde door hen vertoonde patronen, die wij herkennen vanuit een door ons ingenomen interpreterende houding: de intentionele houding. Hiermee lijkt Dennett overigens toe te geven aan de idee dat de waarneming theoriegeladen is. Echter de bron evenals de arbiter die oordeelt over het succes van de theorie is gelegen in de zintuiglijke waarneming, meer in het bijzonder in de relatie subject-object. En de meest betrouwbare institutie die zich bezighoudt met het abstraheren van kennis uit het publiekelijk waarneembare is de objectiverende (natuur)wetenschap.

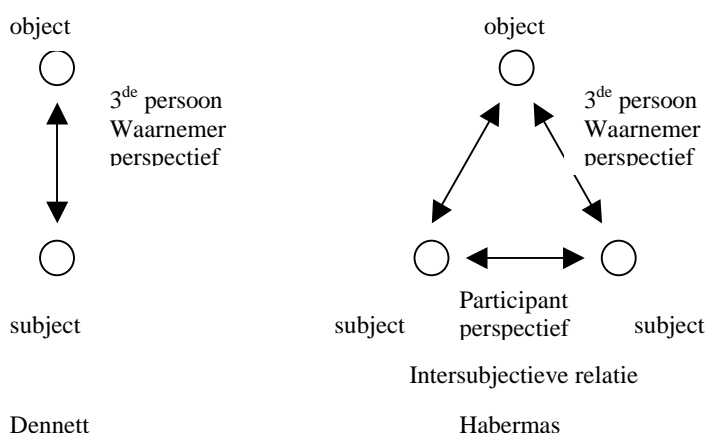
Als Dennett een evolutionaire rechtvaardiging geeft voor het primaat van de subject-object relatie in onze kennis, en de correspondentietheorie van waarheid (maar dan nu bedoeld als een propositie die correspondeert met een stand van zaken in de werkelijkheid), dan zal hij de nadruk leggen op de consequentie van een mogelijke kloof tussen verschijning en werkelijkheid voor de overlevingskansen van een wezen.

We kunnen wel de correspondentie van onze theorieën met iets in de werkelijkheid gaan relativeren, maar als je naar de ontstaansgeschiedenis van het leven kijkt, dan kan het niet anders zijn dan dat de mate waarin er een kloof aanwezig is tussen verschijning en werkelijkheid het verschil kan maken tussen leven en dood. Als een primitief wezen een compleet ‘vals beeld’ heeft van hetgeen er in zijn omgeving gaande is, dan kan hem dat fataal worden.

Maar dit soort van evolutionaire redeneringen kunnen ook gebruikt worden om een andere visie, die van het pragmatisme, te verdedigen. Het kan denkbaar zijn dat niet de waarheid als correspondentie met iets in de werkelijkheid in het voordeel van ons overleven is, maar iets anders. Maar de vraag blijft wat dat ‘iets anders’ in plaats van correspondentie kan behelzen als je bedenkt dat de theorie alleen maar instrumenteel succesvol gebruikt kan worden om iets in de wereld gedaan te krijgen, als deze op een of andere manier inhaakt op een stand van zaken in die wereld.

Tegenover het bezwaar dat we zonder een linguïstische bemiddeling geen toegang tot de werkelijkheid hebben, er dus geen archimedisch punt is om te beslissen of onze theorieën ook werkelijk corresponderen met de werkelijkheid, moet er een verdediging van de correspondentietheorie van de waarheid gevonden worden in iets wat prelinguïstisch is.

Kunnen we het succes waarmee we onze (objectiverende) theorieën lijken te kunnen toepassen, prelinguïstisch koppelen aan de visie dat deze theorieën (die ons de verschijningen oplevert) daadwerkelijk corresponderen met een stand van zaken in de werkelijkheid? Onze aanwezigheid in het hier en nu zou dat kunnen doen. Zou er geen goede match zijn tussen verschijning en werkelijkheid, hadden we het niet zover gebracht. Zonder correspondentie tussen theorie en iets in de werkelijkheid is er geen overleven. Als de theorie over objecten en subjecten niet een zekere mate van correspondentie vertonen met feiten in de werkelijkheid, hou zouden ze dan instrumenteel kunnen zijn, hoe zouden we de feiten er dan mee kunnen manipuleren?



Figuur 1; Schematische weergave van cognitieve toegang tot de wereld .

Habermas, in tegenstelling tot Dennett meent dat de voorwaarden voor de cognitieve toegang tot de wereld van de objecten niet door diezelfde subject-object relatie kan worden gevestigd. Habermas hanteert een aanpassing van de opvatting van Kant, dat de ‘objectieve’ wereld

wordt geïnterpreteerd met de hulp van een geraamte van basiscategorieën. De aanpassing van Habermas bestaat er uit dat dit geraamte niet zoals bij Kant a-priori is, maar de socio-culturele regels van de leefwereld reflecteert. De in de praktijken van de leefwereld belichaamde socio-culturele regels leveren de voorwaarden voor onze cognitieve toegang tot de 'objectieve' wereld. Deze leefwereldpraktijken, die niet zijn te reduceren tot objectieve beschrijvingen (gevormd door de subject-object relatie), betreffen het performatief, dwz. communicatief door argumentatie vestigen van overeenstemming m.b.t. de waarheid, morele juistheid en oprechtheid van onze taaldaden. Waarheid en morele juistheid van taaldaden worden intersubjectief (subject-subject) in een ideale spreesituatie gevestigd. De consequentie hiervan is dat, omdat het objectieve perspectief (relatie subject-object) het participantperspectief (relatie subject-subject) vooronderstelt, een volledige beschrijving van de wereld niet mogelijk is vanuit het objectieve perspectief. Vandaar dat Habermas de opvatting afwijst dat de wereld uitputtend in fysische termen valt te beschrijven. Het plaatje, zo meent hij, wordt pas compleet met het complement ervan, de interpreterende beschrijvingen van het participerende perspectief. M.a.w. verklaren verenigd met verstaan. Vanuit zijn visie is kennis over de werkelijkheid dus immer een construct van ons denken, en niet noodzakelijk een één-op-één afspiegeling van iets in de werkelijkheid. Objectieve uitspraken zeggen niet alleen iets over de wereld, maar ook iets over ons.

Elke taalhandeling is opgebouwd uit een 'propositioneel' gedeelte waarin verwezen wordt naar een bepaalde stand van zaken, en een performatief gedeelte waarin de spreker door argumentatie duidelijk maakt hoe deze stand van zaken moet worden opgevat. Als een mededeling, als een vraag, als een belofte, etc. Uitgangspunt is dat de betekenis van elke uitspraak afhankelijk is van een interpretatiekader (denk aan de gestaltswitches, plaatjes waarin we afhankelijk van de interpretatie een oud vrouwtje of een jonge dame herkennen, die de theoriegeladenheid van de waarneming doen bevestigen). Wat sprekers en toehoorders doen door te argumenteren is hun interpretatiekaders synchroniseren. En gezien het feit ook dat kennisaccumulatie nimmer passief schouwend is maar immer actief probleemoplossend, vereist een volledig beeld van de wereld een performatief participantperspectief op het communicatief handelen.

Resumerend kun je zeggen dat waar Dennett het primaat legt bij de relatie subject-buitenwereld, Habermas het primaat bij de communicatie tussen subjecten onderling plaatst. Voor Dennett gaan taal en communicatie niet zonder een ideaal van waarheid. Voor Habermas gaat waarheid niet zonder een ideaal van taal en communicatie.

Zoals eerder gezegd leunt de epistemologie van Habermas sterk op het pragmatisme van Peirce en Dewey. Voor hen was een propositie waar als deze bevredigend werkt. De betekenis van een propositie moet worden gevonden in de praktische consequenties die de acceptatie ervan heeft. Onpraktische ideeën moeten worden verworpen.

Typische karakteristieken van het pragmatisme zijn: het anti-cartesianisme en een Kantiaans geïnspireerde afwijzing van de toeschouwer theorie van kennis en de correspondentie theorie van waarheid. Het anti-cartesianisme bestaat uit de verwerping van de opvatting uit de erfenis van Descartes, dat de geest in een in zichzelf opgesloten ruimte bestaat, waarvan de inhoud (ideeën of impressies) subjectief en privaat is, en geheel afgezonderd van de publieke en objectieve wereld. Als we dit beeld eenmaal hebben geaccepteerd, dan worden we vervolgens

geconfronteerd met een reeks van problemen betreffende solipsisme, scepticisme, anti-realisme en idealisme, waarmee empiristen lang hebben geworsteld.

Maar hierbij hoeven we voor een onderlinge vergelijking tussen de opvattingen van Habermas en Dennett, niet te lang stil te staan, want we hebben gezien dat ook Dennett, net als Peirce, James, Dewey, Popper, Wittgenstein, Davidson en Rorty, dit cartesianisme verwerpt. De Kantiaanse erfenis welke Pragmatisten accepteren betreft de conceptuele geladenheid van de waarneming. De waarneming levert nooit directe kennis, dat is de zgn. 'myth of the given'. Altijd is er een vorm van interpretatie op onze waarneming betrokken. Dennett hanteert zoals we hebben gezien ook een vorm van interpretationisme, doch in tegenstelling tot Habermas blijft het binnen het bereik van het perspectief van de waarnemer. Alle betekenisvolle kennis is geobjectiveerde kennis. Wat Habermas' verwerping van de correspondentietheorie van de waarheid betreft, bevindt Dennett zich toch in een ander kamp. Hij vindt deze neiging naar een epistemologisch relativisme onaanvaardbaar. Dennett ageert met name tegen Richard⁹⁴ Rorty's eindeloze poging om de 'illusoire constructie van objectiviteit' te ontmaskeren.

"The sad fact is that in some intellectual circles, inhabited by some of our more advanced thinkers in the art and humanities, this attitude passes as a sophisticated appreciation of the futility of proof and the relativity of all knowledge claims. In fact opinion, far from being sophisticated, is the height of sheltered naiveté, made possible only by flatfooted ignorance of the proven methods of scientific truth-seeking and their power. Like many another naïf, these thinkers, reflecting on the manifest inability of their methods of truth-seeking to achieve stable and valuable results, innocently generalize from their own cases and conclude that nobody else know how to discover the truth either...Among those who contribute to this problem, I am sorry to say, is an earlier Amnesty Lecturer in Oxford, my good friend Richard Rorty." [Dennett⁹⁵]

Als onze theorieën ons in staat stellen om de meest gecompliceerde technische hoogstandjes uit te voeren, mensen naar de maan te sturen, met elkaar over de hele planeet via internet te communiceren, onze eigen DNA helemaal te ontrafelen, met succes gecompliceerde hersenchirurgie te verrichten, dan kun je toch niet beweren dat wat voor objectiviteit doorgaat niets anders is dan een contingent en (inter)subjectief construct. Als ik in de wirwar van waterwegen in de Biesbosch verdwaal, heb ik toch liever een navigatiekaart van de scheepsvaart dan een wegenkaart waar toevallig de Biesbosch ook op staat, omdat de eerste beter correspondeert met de feiten in de Biesbosch.

Zo ver als Rorty in zijn epistemisch relativisme gaat, gaat Habermas niet, doch men kan zich wel afvragen in hoeverre geldigheidsclaims betreffende een stand van zaken beschreven door de natuurwetenschap ook het perspectief van de participant nog als aanvulling behoeft. Dit perspectief zou je als verdisconteert mogen beschouwen in het objectiverende perspectief als onderdeel van de wetenschappelijke praktijken. Je zou zeggen dat verstaan toch niet van toepassing is op de waarheid van een wiskundige vergelijking in de snaartheorie.

Argumentatie binnen de wetenschappelijke gemeenschap kan in dit domein gescheiden zonder enige aanspraak op een waarde als morele juistheid, en lijkt uitputtend te kunnen worden beschreven in louter objectiverende termen. Maar deze kritiek gaat natuurlijk wel voorbij aan gevallen waarin de natuurwetenschappelijke methoden worden aangewend voor onderzoek naar de menselijke psyche en daaraan verwante onderwerpen van de

menswetenschappen. Dat we de mens niet mogen objectiveren heeft, zoals we hebben gezien, voor Habermas ook nog een morele grond.

7 Pragmatische beoordeling van antireductionistisch fysicalisme en zwak naturalisme

We stoppen hier het theoretisch filosoferen in argumenteren vóór en tegen de theorieën van Habermas en Dennett en kiezen tot slot een pragmatistische aanpak. Dat doen we door te kijken naar de praktische gevolgen die het voor waar houden van de theorieën van Habermas en Dennett hebben. Meer in het bijzonder komt dan aan de orde welke praktische implicaties de adoptie van de theorieën hebben voor de wijze waarop we om zullen gaan met de relatie tussen manifest mensbeeld en een mensbeeld gebaseerd op (natuur)wetenschappelijke kennis. Centraal voor Habermas staat het vermogen van kritische theorieën om er normatieve praktijken mee te kunnen beschrijven, en er kritische analyses mee uit te kunnen voeren. Volgens hem stellen de op het natuurwetenschappelijke paradigma gebaseerde theorieën ons daartoe niet in staat. Achtereenvolgens ga ik in op de volgende vragen:

- Biedt de theorie een goede mogelijkheid voor de beschrijving van normativiteit?
- Kan men met de theorie een kritisch standpunt vertolken?
- Wat is de implicatie van een adoptie van de theorie voor de omgang met nieuwe (neuro)wetenschappelijke onderzoeksresultaten in relatie tot het manifeste mensbeeld?

7.1 Biedt de theorie een goede mogelijkheid voor een beschrijving van normativiteit?

Uitspraken zijn alleen betekenisvol indien ze getoetst kunnen worden aan de hand van het verificatieprincipe, dwz verifieerbaar aan de hand van publiekelijke waarneembaarheden. Aldus baseerden Logisch positivisten, fysicalisten ofwel objectivisten zich op de idee dat alleen publiekelijk verifieerbare zintuiglijke waarneming betekenisvolle kennis kan leveren. De uit betekenisvolle uitspraken bestaande kennis moest over een stand van zaken in de wereld gaan. Het enige geschikte perspectief op kennis zou het objectieve ofwel derdepersoonperspectief zijn. Vanuit dit perspectief zouden we in staat gesteld worden om alles wat we waarnemen en waarover we een uitspraak doen, publiekelijk te laten bevestigen. Uitspraken die niet op deze manier geverifieerd konden worden waren betekenisloos. Zo werd beschrijvende ethiek wel maar normatieve ethiek niet tot betekenisvolle wetenschappelijke kennis gerekend. Immers een stand van zaken vertelt ons niet hoe de dingen behoren te zijn.

Logisch positivisten meenden dat alle wetenschappen (ook de menswetenschappen) dit verificatie-ideaal en het deductief nomologische model⁹⁶ van de empirische, dwz natuurwetenschappen moesten volgen. Ze streefden naar een eenheidswetenschap. Doel van de wetenschap is het vinden van wetmatigheden uit hetgeen we empirisch in de zintuiglijke waarneming kunnen vaststellen.

Dit lokte protesten uit van menswetenschappers die zich volgens de criteria van logisch positivisten bezighielden met het bedrijven van 'pseudo-wetenschap'. Volgens Weber en Dilthey konden de empirische wetenschappen helemaal niet tot één wetenschap (het fysicalisme), teruggebracht worden. Natuur-wetenschap is waardeneutraal, houdt zich bezig met verklaren en houdt er een objectivistisch perspectief op de wereld op na, dwz vat alle waargenomen processen op als causaal nomologisch te verklaren processen.

Menswetenschappen houden zich bezig met verstaan, door inlevingsvermogen en empathie. Ze hanteren een perspectief van waaruit de waarneming wordt geïnterpreteerd. Betekenis van menswetenschappelijke uitspraken is betekenis-als-relevantie die is geconstrueerd uit de interpretatie van de waarneming als een voelen en beleven, met een beroep op het (inter)subjectieve perspectief.

De logisch positivisten, niet onder de indruk van deze kritiek, kwamen uiteindelijk toch onder druk te staan door kritiek van empiristen uit eigen gelederen en daarbuiten. Het verificatieprincipe was te sterk. Natuurwetten bleken niet verifieerbaar te zijn. Het voorgestelde alternatief, het confirmatieprincipe was te zwak, waardoor pseudo-wetenschappen niet konden worden uitgesloten. Wittgenstein wees op de theoriegeladenheid van de waarneming. Feiten in de wereld mogen dan misschien wel neutraal zijn, maar de observatie van die feiten zijn niet neutraal en voor ieder weer anders. Iemands observatie is altijd afhankelijk van diens achtergrondkennis. Quine wees op het feit dat wetenschappelijke theorieën immer onderbepaald zijn door de feiten (ik kan wel zeggen dat alle raven zwart zijn aan de hand van de waarneming van duizend raven maar om de hoek kan er wel eens een witte raaf zitten). De Duhem Quine these vertolkte de idee dat een theorie nooit in isolatie kan worden getest, waardoor je niet alleen de theorie aan de hand van de werkelijkheid test, maar ook de eraan ten grondslag liggende veronderstelling. Popper bracht de nekslag toe door te wijzen op het inductieprobleem en de theoriegeladenheid van de waarneming. Inductie als principe is niet betrouwbaar omdat het zelf op inductie is gebaseerd, en zuivere feiten bestaan niet maar zijn altijd in theoretische concepten gehuld. Poppers voorgestelde falsificatieprincipe onderbouwd door het principe van weerlegbaarheid kon het demarcatieprobleem beter hanteren dan het op inductie gebaseerde verificatieprincipe.

Zo laat de wetenschapsfilosofie met betrekking tot de waarheid sinds het begin van de twintigste eeuw een terugtrekkende beweging zien. Maar daarbuiten liggen de zaken anders. Ondanks de vernietigende kritiek van filosofen is het logisch positivisme 'op de werkvloer' zeker in de natuurwetenschappen, nog de standaardopvatting over het reilen en zeilen van de wetenschap.

Daarmee is het fysicalisme en de idee dat het objectivistische perspectief het enige perspectief is wat als gezaghebbend voor kennisaccumulatie moet worden beschouwd, immer nog dominant. En het is deze stroming waar Habermas tegen strijdt. Dat de beschrijving van normativiteit niet op juiste wijze vanuit het objectivistische perspectief tot stand kan komen is zijn grootste bezwaar.

Om normativiteit in beeld te krijgen is, zoals voor elke waarneming, een interpretatiekader nodig. Het interpretatiekader van het objectivistische perspectief is ongeschikt voor het geven van een analyse van de mensen in hun onderlinge relaties omdat dit kader al hetgeen wordt waargenomen in termen van nomologische causaliteit verklaard.

Wat zo ontbreekt is de beleving van binnenuit. Wat nodig is om deze beleving naar voren te krijgen is een hermeneutisch perspectief waarin ik hetzelfde interpretatiekader aanneem als degenen waarvan ik een analyse wil maken. De strategie die ik daarbij moet volgens is dat ik de mensen in de samenleving waarvan ik een analyse maak, met argumenten kan overtuigen van de waarheid, waarachtigheid en morele juistheid van deze analyse. Alleen dan weet ik dat ik een zelfde interpretatiekader heb aangenomen als deze mensen. Dit ingenomen perspectief is het participantperspectief op het communicatief handelen.

We zouden nu kijken of de praktische gevolgen, dus de toepassing van Habermas' theorie inzake de beschrijving van normativiteit, anders is dan die voor de theorie van Dennett. Zoals

reeds besproken hanteert Dennett weliswaar een fysicalistische benadering, maar neemt daarbij wel een antireductionistisch positie in. Hoewel hij meent dat de wereld in grond bestaat uit fysische entiteiten en processen die in principe te verklaren zijn uit de eigenschappen van de elementaire deeltjes, erkent hij dat er verschillende perspectieven zijn waarin we naar deze wereld kunnen kijken. Dennett verwerpt niet de volkspychologie daar hij meent dat dit een effectieve manier is om het gedrag van anderen en jezelf te verklaren en beheersen. Dit beeld valt niet te reduceren tot bijvoorbeeld neurofysiologische beschrijvingen omdat de gepostuleerde mentale toestanden, hoewel handige rekeneenheden binnen de volkspychologische theorie, niet verwijzen naar innerlijke processen. Hoewel Dennett in zijn beschrijving van de evolutie van de vrije wil de memetiek als (niet door Habermas geaccepteerd) model voor cultuurverspreiding hanteert, schakelt hij uiteindelijk over op het intentionele perspectief en de taal van de volkspychologie om zijn verklaring te voltooien. Dennett's interpretationisme toont in ieder geval dat er op basis van een fysicalisme een theorie is te ontwikkelen die de basis kan leggen voor de epistemische rechtvaardiging van een intentioneel vocabulaire, en van de idee dat actoren worden bewogen door redenen, iets wat Habermas essentieel acht voor het kunnen beschrijven van normativiteit, en een hermeneutisch verstaan. Ik concludeer derhalve dat Dennett's theorie, waar het in zijn praktische gebruik betrekking heeft op de beschrijving van normativiteit in de samenleving, hoewel het een fysicalisme betreft, toch het belangrijkste gereedschap rechtvaardigt dat Habermas nodig heeft om zijn kritische analyse te geven; n.l. een volkspychologische benadering die het handelen in termen van intentionaliteit en het bewogen worden door zelfgekozen redenen en argumenten verklaart. In dit opzicht is er geen verschil⁹⁷. Beide theorieën scheppen ruimte voor beschrijvende normativiteit en ethiek. Maar dat men met een theorie ook kritisch kan analyseren en een normatief standpunt in kan nemen is een zwaardere eis. Dat bekijken we nu dan!

7.2 Kan men met de theorie een kritisch standpunt vertolken?

Kan met de theorie een kritische houding worden aangenomen ten aanzien van maatschappelijke ontwikkelingen en machtsverhoudingen binnen de samenleving? Habermas' bezwaar tegen theorieën gebaseerd op louter het objectiverende perspectief is dat ze normativiteit niet zichtbaar maken, en het vermogen missen er een kritische analyse mee te voltrekken. De theorie van Habermas is ontworpen om aan te tonen dat de verlichting ook een communicatief rationeel vermogen om kritiek te leveren heeft voortgebracht. Elke taalhandeling is volgens Habermas opgebouwd uit een 'propositioneel' deel, waarin wordt verwezen naar een bepaalde stand van zaken en een 'performatief' deel waarin de spreker duidelijk maakt hoe deze stand van zaken moet worden opgevat: als een mededeling, een vraag, een belofte, een bevel etc. De spreker maakt de communicatieve strekking van zijn taalhandelingen aan de toehoorders duidelijk via de geldigheidsaanspraken die hij via het performatieve gedeelte van zijn taalhandelingen naar voren brengt, namelijk de aanspraken op waarheid, juistheid en waarachtigheid. Deze geldigheidsaanspraken hebben in principe een kritiseerbaar karakter. De toehoorders kunnen de aanspraken afwijzen en eigen afwijkende aanspraken er tegenover stellen en hun redenen ervoor beargumenteren. De interne rationele structuur van dit communicatief handelen is zodanig dat de participanten zonder dwang of manipulatie zich rationeel laten overtuigen door de dwangloze dwang van het beste argument. Als allen overtuigd zijn is er consensus bereikt en mogen we deze

rationeel noemen. Deze consensusstheorie zou ook een rationeel gefundeerde normatieve consensus aannemelijk maken. Het verwijt van Habermas is nu dat objectivistische perspectieven het performatieve communicatieve deel niet in rekening brengen. Daardoor wordt het deel wat het kritisch vermogen met zich meedraagt niet verdisconteerd.

De vraag is echter of het gerechtvaardigd is aan te nemen dat normatieve en morele oordelen binnen het bereik van de cognitie (de praktische rede) en de dwangloze dwang van het betere argument zijn gelegen. Filosofen van de Humeiaanse traditie betwijfelen dit in ieder geval. Laat ons eens verplaatsen in Daniel Dennett's fysicalistische opvattingen. Dennett meent dat de wereld geheel uit fysische entiteiten bestaat, en dat het verloop van de processen in de wereld zich in principe (door een Laplaciaanse demon) aan de hand van de fysische wetten laat verklaren. Het objectivistische (subject-object) perspectief is, omdat de wereld geheel wordt geconstitueerd door fysische entiteiten en processen, een goede en feitelijke de enige manier om betekenisvolle kennis over de wereld te vergaren. Het eerste-persoonsperspectief is niet het perspectief van een cartesiaanse ego of innerlijk geestesoog. Het cartesiaans ego is product van een narratief zwaartepunt een gebruiksiillusie die ons een zelf heeft geleverd. Dennett maakt geen onderscheid tussen subject-subject of subject-object perspectief, (tenslotte kunnen we vanuit zijn optiek evenzogoed intentionaliteit toeschrijven aan dingen, bijvoorbeeld computers, als aan mensen). Voor Dennett is de mens uiteindelijk een object, ware het niet van wel een speciale soort. Maar Dennett erkent ook dat de zintuigen geen kant en klare informatie afgeven, maar telkens in theoretische concepten zijn gehuld. Gedurende de evolutie is de taal ontstaan, en deze taal heeft tevens aanleiding gegeven tot de vorming van een gebruikersinterface waarmee we de ander en onszelf konden gaan voorspellen en manipuleren. De subject-object toegang tot de wereld kent bij Dennett verschillende interpretatieniveaus, elk met een eigen taalgebruik. Voormist onze cognitieve vermogens dat toestaan is het fysische niveau het niveau van waaruit we de processen tot in de kleinste detail kunnen verklaren en voorspellen. Vanuit het intentionele perspectief interpreteren we (menselijke) gedragspatronen als mentale toestanden, en het handelen in termen van redenen, motieven en dergelijke. Het vocabulaire kent ook termen voor gevoelens, en emoties. Nu hanteert Dennett net als Hume de opvatting dat de moraal is gefundeerd in de natuurlijke drijfveren, de emoties, waaruit door reflectie in conventies de kunstmatige deugden van de moraal ontstaan. *“Als reflecterende rationele actoren die bezig zijn met de gezamenlijke activiteit van wederzijdse overtuigingskracht kunnen we uitgaande van onze huidige waarden nieuwe hogere doelen formuleren. Met de wetenschap als hulp kunnen we dan kijken of deze ethische aspiraties goed verankerd zijn [Dennett⁹⁸]”*

Wetenschap kan dienen als hulp maar behoort zelf dus geen uitspraken te doen over hetgeen we behoren te doen. Daar zijn andere instituties als bijvoorbeeld de politiek, of ethische commissies voor. Het formuleren van de hogere doelen kan worden bediscussieerd. Echter, omdat hier de bron van de moraal in de emotie ligt, is er behalve dat de natuurlijke drijfveren bij mensen wel enige uniformiteit vertonen, geen universalistisch principe (zoals de praktische rede of de formele discoursethiek) waarop men in het geval van onenigheid terug kan vallen. Consensus zal uitzondering en *modus vivendi* de regel zijn. In zoverre is er dus geen absoluut gefundeerde kritiek te leveren omdat men niet kan terugvallen op een universele morele waarheid.

We hadden al geconstateerd dat in tegenstelling tot wat Habermas stelt, Dennett wel ruimte maakt voor een benadering die een onderscheid aanbrengt tussen, en ruimte schept voor zowel beschrijvingen in termen van louter nomologisch causale processen, en handelingen

die in volkpsychologische termen van redenen en motieven worden verklaard. Maar het moet gezegd worden dat Dennett's theorie niet expliciet een kritisch vermogen verdisconteert. Wat daarvoor nodig is, is een criterium aan de hand waarvan we onze normen en waarden intersubjectief kunnen funderen. En gegeven dat Dennett als Humeiaanse afstammeling en empirist meent dat de waarden die we nastreven uit de natuurlijke instincten en de emoties ontspringen, hebben we er geen directe cognitieve toegang toe, zoals Kantianen aannemen dat we die wel hebben omdat zij ervan uitgaan dat de moraal in de praktische rede is gefundeerd.

We kunnen volgens Humeianen niet, zoals Kantianen beweren, met de praktische rede bepalen of iets moreel juist is zoals we met het verstand (en eventueel de waarneming) kunnen bepalen of iets waar is. Kennis over de natuurlijke drijfveren en de emoties, en dus kennis over de bron van de moraal en normativiteit, is volgens empiristen wel mogelijk langs de weg van de objectiverende wetenschappen die zich baseren op hetgeen publiekelijk zintuiglijk verifieerbaar of weerlegbaar is. Vanuit deze optiek verloopt dus de cognitieve toegang tot de bron van normativiteit en moraal via de weg van wetenschappelijk onderzoek naar de aard en de rol van de natuurlijke drijfveren, en de emoties⁹⁹ Aan de hand van de wetenschappelijk informatie over de aard en hoedanigheid van onze natuurlijke drijfveren kunnen we dan proberen onze morele regels hier zo goed mogelijk op aan te laten sluiten. In een discours kunnen we vaststellen welke doelen we het waard achten om nagestreefd te worden. Er bestaat echter geen universeel principe (zoals de praktische rede) die ons hierin kan leiden. Maar als we eenmaal de doelen hebben bepaald, door een consensus of *modus vivendi*, dan kunnen we aan de hand van resultaten uit wetenschappelijke onderzoek naar de aard van onze natuurlijke instincten, drijfveren en emoties, bepalen hoe we de normen het beste vorm kunnen geven om de overeengekomen doelen te bereiken.

Habermas' opvattingen verschillen in de praktijk hier wezenlijk van die van Dennett. Habermas wijst überhaupt de idee af dat de mens toereikend in causaal nomologische 'objectiverende' termen van de wetenschap te verklaren zou zijn. En dat de moraal is gefundeerd in de natuurlijke drijfveren en emotie zal hij ook niet accepteren. Als een Kantiaan meent hij dat deze in de praktische rede is gefundeerd. Vandaar dat hij geen moeite heeft om morele juistheid ook onderwerp te laten zijn van argumentatie en rationaliteit. Om met een theorie een kritische analyse uit te kunnen voeren moet er een (objectief of intersubjectief) criterium zijn aan de hand waarvan (de beschrijving van) een bestaande situatie kan worden getoetst. Humeianen en empiristen in het algemeen menen dat dit criterium niet kan worden gevonden in een stand van zaken in de werkelijkheid, en dat is wat ze als enige valide bron van betekenisvolle kennis beschouwen. Habermas kan zijn theorie van een kritisch vermogen voorzien omdat hij de universaliteit van de eraan ten grondslag liggende waarden (criteria) in de principes van zijn discoursethiek verantwoord. Vanuit dit kritisch vermogen geeft Habermas in zijn theorie dan ook meer aandacht voor de soms benarde positie van de mens in een door natuurwetenschap en technologie beheerst tijdperk.

7.3 Implicatie theorie adoptie voor omgang met wetenschappelijke onderzoekresultaten?

Wat is de implicatie van een adoptie van de theorieën voor de omgang van resultaten uit recent (neuro)wetenschappelijk en psychologisch onderzoek in relatie tot het manifeste mensbeeld? De vraag is dan primair hoe binnen de twee theorieën de onderlinge verhouding tussen manifeste en wetenschappelijke kennis wordt geschetst? Het onderscheid tussen

wetenschappelijke en manifeste kennis is benadrukt door Wilfrid Sellars. Dennett spreekt in ‘Sweat Dreams’ over het onderscheidt tussen wetenschappelijke kennis en ‘folk physics and folk psychology’. De laatst genoemden zijn handige instrumenten, maar de wetenschap als meest betrouwbare en grondige belichaming van de zoektocht naar (intersubjectieve) waarheid, moet het hoogste woord hebben.

“If we let the elaboration of naive physics and academic physics be our guide, we can expect that in due course much of what is intuitive in folk psychology will be vindicated, incorporated, explained by our advanced academic theories of the phenomena. After all the reason naïve physics is so valuable to us is that its deliverances are, in the main true. We should expect folk psychology to be similiary rich in retrospectively confirmed truth. But we can’t count on it.[Dennett¹⁰⁰]”

Dennett meent dat, ofschoon de wereld in grond fysisch dus causaal nomologisch deterministisch is, dit in geen geval uitsluit dat we actoren zijn wier handelen worden bewogen door zelf gekozen redenen. Dat laatste is de volkpsychologische interpretatie vanuit het intentionele perspectief en deze werkt als instrumentele benadering uitstekend.

Toegespitst op de kwestie van de vrije wil, hoeven onze opvattingen aangaande vrije wil niet te botsen met het determinisme van het allesomvattende natuurwetenschappelijke wereldbeeld. Als actoren hadden we in de niet-enge zin ook anders kunnen handelen. We zijn vrij omdat vrijheid vanuit het antropocentrische (intentionele) perspectief moet worden gelijkgesteld met vermijdbaarheid (vrijheid is het rationele vermogen te kunnen vermijden). Bovendien impliceert (actorcentrisch bezien) fysisch determinisme geen voorspelbaarheid omdat we niet kunnen weten welke wereld van de verzameling van alle mogelijke werelden onze actuele wereld is.

De praktische implicaties van deze theorie in ogenschouw genomen, wordt er voldoende ruimte ingeruimd voor de verantwoording van het concept van moreel actorschap terwijl er toch vanuit kan worden gegaan dat de allesomvattende natuur samenvalt met het object van de natuurwetenschap. Dit maakt de speculatieve metafysica van een metatheoretische verbinding tussen het manifeste en het wetenschappelijke perspectief overbodig. Bijzonder voordeel is dat de natuurwetenschap zich niet geconfronteerd hoeft te zien met problemen die zouden ontstaan als haar onderzoeksdomein zou worden gecompartmenteerd.

Volgens Habermas echter kan een agent niet op het zelfde moment én het wetenschappelijke beeld én het manifeste beeld voor waar houden als deze daarbij aanneemt dat de allesomvattende natuur het object van de natuurwetenschap is.

De waarneming van het eigen optreden interpreteren als een gebeurtenis die nomologisch oorzakelijk tot stand wordt gebracht, en tegelijkertijd interpreteren als een door redenen bewogen handelen in een fysische ontologie, is volgens hem contradictief.

Maar omdat de cognitieve toegang tot de wereld niet direct maar taalbemiddeld is, moeten we er vanuit gaan dat de natuur meer omvat dan de natuur van de natuurwetenschap, en moeten we het complementaire paar, manifeste, en het wetenschappelijke beeld coëxisterend, in één metatheorie onderbrengen.

Kijkend naar de praktische gevolgen van deze theorie heeft de natuurwetenschap een begrens domein en houden we vast aan een zelf dat metafysische is geconstrueerd.

Maar zoals ik heb aangetoond hebben we door het fysische domein te compartimentaliseren, een eigenschapdualisme of een eigenschappen-n-alisme geschapen waardoor we geconfronteerd

worden met bijkomende problemen, als epifenomenalisme of noodgedwongen verwerping van de these van de causale geslotenheid van het fysische domein. Dit lijkt me niet een wenselijke situatie. Bijkomend probleem is hier de vraag hoe in de praktijk moet worden omgesprongen met nieuwe resultaten uit het door Habermas gecompartmenteerd neuro-wetenschappelijk onderzoek.

Het onderscheid in epistemische status en het te leveren aandeel van wetenschappelijke kennis in de formulering van morele verantwoordelijkheid gaat terug op een verschil in opvattingen over de wortels van de moraal. Habermas brengt deze geheel binnen het bereik van de praktische rede, terwijl Dennett deze gelegen ziet in de natuurlijke instincten. De eerste plaatst het kennen van moraliteit binnen het bereik van de dwangloze dwang van het beste argument zonder noodzakelijk te appelleren aan empirische kennis, terwijl de tweede juist zijn fundament zoekt in empirische kennis.

Hoe zullen de theorieën nu omgaan met nieuwe resultaten uit wetenschappelijk onderzoek met betrekking tot het geadopteerde mensbeeld? Meer specifiek: hoe wordt moreel actorschap gerechtvaardigd in het licht van recente resultaten uit de psychologie en neurowetenschap?

Nu Habermas het manifeste beeld en het wetenschappelijke beeld een zelfde gezag heeft toegekend zal hij wetenschappelijke kennis altijd metatheoretisch moeten samenvoegen met het manifeste beeld van moreel agency (acting for reason en being able to do otherwise).

De vraag is of zo'n metafysica de druk aan kan van nieuwe bevindingen op neurowetenschappelijk gebied.

Voor Dennett ligt dit met zijn fysicalisme anders. Voor hem is de volkpsychologie een ruwe theorie, van het objectiverende beeld. De wetenschap is wegens zijn veel grotere nauwkeurigheid en strengheid van onderzoek gezaghebbend. Hij schuift wetenschappelijke data niet zomaar terzijde. Hij gebruikt de data om uiteindelijk onze morele normen vastgelegd in conventies beter aan te passen aan de empirisch aangetoonde feitelijkheid van onze geaardheid. Maar tegelijkertijd erkent Dennett ook dat onze epistemische horizon begrensd is en dat de volkpsychologie 'here to stay' is. Los daarvan is het helemaal niet praktisch om alleen maar vanuit het fysische niveau de wereld te beschrijven te verklaren en te voorspellen. Een hoger abstractieniveau en een daarbij behorend perspectief blijft noodzakelijk. Kortom, voor de praktijk maakt de benadering gerekend naar onze manifeste gebruiken voor het voor waar houden van Habermas' of Dennett's theorie weinig verschil, ware het niet dat Habermas veel terughoudender zal zijn om wetenschappelijke data mee te nemen in de overwegingen bij het vaststellen van de normen van de morele en gerechtelijke instituties. Maar ik denk dat je niet alsmaar met een metafysisch geconstrueerde zelf de druk van de wetenschappelijke resultaten kunt weerstaan.

8 Conclusie

Dan zijn we nu aan de eindbalans toegekomen. We waren begonnen met de vaststelling van een vermeende kloof tussen twee beelden. Het manifeste zelfbeeld van de morele agent, die zijn handelen bewust laat leiden door vrij gekozen redenen, en het natuurwetenschappelijk geschetste beeld van de mens als onderdeel van een causaal nomologisch te verklaren universum. Dit manifeste zelfbeeld lijkt door recente wetenschappelijke onderzoeksresultaten meer en meer onder druk te komen staan. Deze kwestie heb ik opgepakt door te kijken hoe

Habermas en Dennett deze problematiek benaderen. Zoals we hebben gezien [in secties 2 en 3] hanteren ze daarbij twee verschillende opvattingen van naturalisme.

Habermas heeft een zwak naturalisme geformuleerd waarin de allesomvattende natuur meer is dan de natuur van de natuurwetenschap. Het door het participantperspectief op het communicatief handelen gevormde manifeste beeld, krijgt een zelfde epistemische status toegekend als het objectiverende beeld van de natuurwetenschap door beide perspectieven metatheoretisch te verbinden. Habermas die deze metatheorie nog niet heeft uitgewerkt, doet slechts een voorstel hoe de constructie ervan de mogelijkheidsvoorwaarden voor de twee complementaire perspectieven kan verantwoorden.

Maar wij hebben alvast met een korte proefrit [in sectie 2.5] even geproefd aan de complicaties die door de compartimentering van de natuurwetenschappen optreden, met deze de-facto eerste filosofie van Habermas. We zagen dat zo'n metatheoretische verbinding uiteindelijk toch weer tot dezelfde problemen (zoals het epifenomenale van het mentale) lijkt te leiden waarmee het klassieke eigenschapdualisme zich eerder zag geconfronteerd.

Bovendien ontslaat de compartimentering van wetenschap je nog steeds niet van de taak om te bepalen wat je aan wetenschappelijke informatie meeneemt in de (her)formulering van een verantwoordelijkheidsbegrip. Andere bezwaren tegen deze metafysische antireductie zijn dat het tegennatuurlijk voor het wetenschappelijke ideaal van unificatie is, dat het speculatief is, en dus problemen met weerlegbaarheid oproept.

Deze problemen tezamen waren aanleiding om te kijken of er een mogelijkheid bestaat om het manifeste beeld van moreel actorschap te verenigen met een ontologisch en nomologisch fysikalisme. Hiervoor werd [in sectie 3] Dennett's filosofie als model genomen. Dennett weet met zijn anti-reductionistisch interpretationisme een coherente theorie te produceren waarin het gezag van de volkspychologie, en het manifeste beeld van een morele actor wiens handelen wordt bewogen door vrij gekozen redenen, kan worden gerechtvaardigd, zonder concessies te doen aan het domein van het objectiverende wetenschappelijke perspectief. Zijn interpretationisme lijkt een aantal problemen op te lossen: De wetenschap ziet zich er niet in geconfronteerd met een compartimentering, ze kan haar aanname handhaven dat alles in de wereld in theorie een fysische verklaring heeft omdat Dennett dit beeld toekent aan het perspectief op het grondniveau: het fysische perspectief.

Maar Dennett ontkent daarmee niet dat er een ander perspectief is dat we in kunnen nemen en waarin we waarneembare patronen interpreteren als intentioneel.

Dennett's interpretationisme levert daarmee een oplossing voor de verwarring die er aangaande de causale interactie tussen geest en lichaam bestaat, door die verwarring als een 'category mistake' te ontmaskeren [sectie 4]. Een fout die ontstaat als men één en dezelfde ontologische verzameling van entiteiten onder twee verschillende categorieën (interpretaties) van beschrijvingen kent, maar doet voorkomen alsof die twee ook in ontologisch opzicht van elkaar zijn te scheiden en naast elkaar bestaan.

Het feit dat in zijn interpretationisme volkspychologische beschrijvingen niet tot het neurofysische niveau teruggebracht kunnen worden omdat er geen één-op-één relatie tussen gedrags- en breinpatronen bestaat, geeft aan het beeld van de intentionele houding en de volkspychologie een status waarmee het zich niet makkelijk laat bedreigen door neurowetenschappelijke onderzoeksresultaten. Volkspychologie is een gedurende de evolutie tot ontwikkeling gekomen theorie waarin (gepostuleerde) mentale toestanden handige rekeneenheden vormen in het verklaren en voorspellen van het gedrag van anderen en jezelf.

Volkpsychologie is 'here to stay' omdat ze werkt voor haar doel, het snel ruwweg beschrijven, verklaren en voorspellen van andermans gedrag. Het honoreren van de interpretatie van menselijke gedragspatronen in termen van mentale toestanden, en door redenen bewogen handelen van de agent, als een instrumenteel model om de medemens en jezelf te begrijpen, zou zelfs de weg naar een honorering van de methode van hermeneutisch verstaan kunnen openen.

Maar dat Dennett alles in het oog van de objectiverende waarneming heeft geplaatst kan wel als bezwaar worden aangemerkt. Hiermee lijkt hij het subjectieve belevingsaspect uitgehold te hebben. Het cartesiaanse ego dat niet als bron maar als product van de verhalen wordt beschouwd, is geen geestesoog gericht op onze innerlijke gedachten. Bewustzijn, dus het subjectieve aspect, wordt verklaard als narratief mechanisme.

Beleefde emotie bijvoorbeeld moet op zo'n manier worden uitgelegd als een door de cognitie verrichte talige bestempeling van lichamelijke verandering als gevolg van emotioneel geladen stimuli (Schachter). Die bestempeling geschiedt niet door een cartesiaanse ego maar door de rede. Het subjectieve aspect is daarmee feitelijk ook een aspect van het objectiverende perspectief.

Waar Dennett het subjectieve perspectief aldus reduceert, is Habermas in zijn verwerping van het cartesianisme minder radicaal omdat hij het subjectieve perspectief niet tot het objectieve perspectief reduceert, maar het naar die van de tweede persoon verbuigt.

Met betrekking tot de kwestie van de vrije wil hebben we [in sectie 5] gezien dat Habermas meent dat het taalspel van verantwoordelijk agency niet reduceerbaar is tot het taalspel van de objectiverende (causaal nomologisch verklarende) natuurwetenschap. Participanten die zichzelf zien als morele agents die naar redenen handelen, veronderstellen daarin alternatieve mogelijkheden.

Op dezelfde wijze betoogt hij weer dat we deze perspectieven een gelijke epistemische status moeten toekennen, en in een metatheoretische theorie verenigen onder de aanname dat de allesomvattende natuur meer dan de natuur van de natuurwetenschap is. We hadden van Dennett interpretationisme reeds gezien dat er in een fysicistische ontologie een interpreterend perspectief mogelijk is waarin we mensen opvatten als actoren die handelen naar redenen. In aanvulling hierop ziet Dennett ook geen conflict tussen vrijheid op het niveau van de actor en determinisme op het fysische niveau. Een deterministische fysische ontologie impliceert geen onvermijdelijkheid.

Vrijheid is de vrijheid gezien vanuit de actor die dankzij zijn cognitieve vermogens gedragsstrategieën ontwerpt die hem in staat stellen zich te ontworstelen aan noodlot en onvermijdelijkheid.

Vanuit het objectief (alleswetend) standpunt, het fysische perspectief, is er inderdaad sprake van determinisme. Een mogelijke wereld heeft slechts één toekomst. Maar bekeken vanuit ons gezichtspunt als actoren met beperkte intelligentie en gegevens kunnen we niet weten in welke mogelijke wereld we ons bevinden. Praktisch, bezien vanuit hun perspectief, ligt de toekomst voor actoren dus open.

Maar deze interpretatie van 'alternatieve mogelijkheden' is volgens Habermas niet te verenigen met de noodzakelijke veronderstelling van vrijheid in het taalspel van verantwoordelijk agency. Habermas geeft hiermee blijk een sterke definitie van vrije wil te hanteren. Maar de vraag is of het redelijk is om de intuïtie van vrijheid te paren aan boven elke twijfel verheven overtuiging van alternatieve mogelijkheden *ceteris paribus*. Zeker als je

bedenkt dat Habermas een relativistische houding aanneemt als hij beweert dat we geen onbemiddelde cognitieve toegang tot onszelf en de werkelijkheid hebben.

En dan zijn er nog andere problemen die Habermas ziet als handelingen direct terug te voeren zouden zijn op nomologisch gedetermineerde gebeurtenissen in het deterministisch objectiverende beeld van een fysische ontologie. Propositionele attitudes zouden geheel worden gepasseerd waardoor mentale veroorzaking wel een fictie moet zijn. Maar in een geformuleerd antwoord hierop [in sectie 6.2] hebben we gezien dat een antireductionistisch fysicalisme zo'n directe herleiding tot het fysische perspectief verwerpt. En zelfs als reductie mogelijk zou zijn, kan men nog argumenteren tegen de stelling dat intentionaliteit zich principieel niet laat objectiveren, zeker als je bedenkt dat een thermostaat uitgerust met sensoren in de ruimte in een soort van interne representatie reeds een zekere gerichtheid hierop vertoont. Een psycho-semantiek voor het brein (hoewel in de praktijk waarschijnlijk zeer moeilijk en zeker niet zonder context vast te stellen) is in principe ook niet onmogelijk omdat er geen principiële belemmering is in het verschil in fysische aard van de drager die het symbool vormt, of het nu klanken, tekens op papier, of neurale patronen betreft. En Habermas argument aangaande mentale veroorzaking lijkt ook te worden ontkracht als je aanneemt dat er vanuit het interpreterende intentionele perspectief ook relaties zijn te ontdekken die je als causaal kan opvatten.

Moeten we met Habermas nu toch concluderen dat een fysisch deterministisch universum het taalspel van verantwoordelijk agency onmogelijk maakt? Dennett voert in ieder geval dus een reeks van argumenten tegen deze opvatting van Habermas aan, waarvan beknopt op een rijtje gezet de conclusies luiden: (a) Determinisme op het fysische niveau impliceert nog geen voorspelbaarheid en onvermijdelijkheid vanuit het antropocentrische intentionele perspectief. (b) We moeten de idee van het cartesische ego inruilen voor een coherenter beeld, zodat we niet in de val trappen te denken dat mentale veroorzaking een illusie is. (c) Als je jezelf groot genoeg maakt in tijd en ruimte kun je wel degelijk een zelfbeeld construeren waarin je zelfvormend vermogen hebt. (d) We hebben de rationele vermogens om vrij te zijn. Ons verstand maakt ons vrij, bevrijdt ons uit de ketenen van de onvermijdelijkheid door gevaar te pareren wat zonder onze cognitieve vermogens ons lot zou hebben kunnen bezegelen.

Ik vind het een redelijke poging van Dennett om een basis te leggen voor een antireductionistisch, fysicalistisch gebaseerde verzoening van manifest en wetenschappelijk mensbeeld.

Zou Habermas zich kunnen vinden in hetgeen tot aan hier van Dennett's filosofie is beschreven? Het lijkt mij dat een antireductionistisch fysicalisme als dat van Dennett zou kunnen voldoen aan een aantal eisen die Habermas met zijn zwak naturalisme aan een verklarende theorie stelt: (1) De theorie mag het beeld van agents wier handelen wordt bewogen door redenen niet reduceren tot een causaal nomologische beschrijving van neurofysiologische processen. De theorie van Dennett voldoet hier omdat deze ook dit reductionisme afwijst. Het zou de tools kunnen leveren voor een hermeneutische benadering. (2) De theorie moet de vrije wil van agents tot uitdrukking laten komen. Hoewel Dennett in zijn boek beweert dat hij dat doet, maakt Habermas bezwaar tegen de wijze waarop Dennett de idee van alternatieve mogelijkheden rechtvaardigt op grond van de beperkte epistemische horizon van de actoren. Maar er valt dus over te twisten of Habermas' eis (de intuïtie van boven elke twijfel verheven overtuiging van alternatieve mogelijkheden *ceteris paribus*) niet te zwaar is.

En dan is er nog Habermas' opvatting dat het objectiverende perspectief de normativiteit niet tot uitdrukking kan laten komen. We hebben hierbij gekeken naar de speciale rol die het participerende perspectief hierin speelt. Hoewel de performatieve en pragmatische benadering van kennisaccumulatie plausibele kenmerken heeft, hebben we ook gezien [in sectie 6.5] dat men twijfels kan hebben, zeker waar het de verwerping van de correspondentietheorie van de waarheid betreft.

De andere normatieve geldigheidsaanspraken aangaande morele juistheid en waarachtigheid van Habermas' discousethiek bleken [in sectie 6.4] weer te herleiden te zijn tot de Kantiaanse aanname dat de moraal is gefundeerd in de praktische rede.

De Humeniaanse moraaltheorie die de fundamenteën van de moraal in de emotie en natuurlijke instincten plaatst, ontkent dat normativiteit en moraal door middel van argumentatie en de praktische rede gekend zou kunnen worden. Hoewel open voor debat en intersubjectieve overeenstemming, is het de instrumentele rede die hier bepaalt hoe we de doelen het best kunnen bereiken. De doelen zelf worden primair bepaald door waarden die uiteindelijk in de natuurlijke instincten en de emoties wortelen. Natuurlijk kan men hierover discussiëren maar uiteindelijk kan men niet de (on)juistheid ervan beargumenteren. Deze Humeiaanse visie plaatst de mens binnen de natuur, en doet de scheidslijn tussen mens en de dieren vervagen. Wat mij betreft lijkt me dit empiristische beeld realistischer dan het rationalistische beeld van de mens die zowat ongenaakbaar op het hoogste punt in de pyramide van de schepping of evolutie in een oefening van ethisch theoretiseren en idealiseren de principes van de moraal vestigt. Ik denk dat we in het bepalen van onze morele instituties de empirische (dus wetenschappelijke) feiten over onze natuurlijke aard niet moeten negeren. We kijken naar wat we hebben om dan te kunnen zien hoe we hierop onze morele instituties het best kunnen laten aansluiten om de gezamenlijk vastgestelde (morele) doelen te kunnen bereiken.

Zo moeten we ook Dennett's suggestie begrijpen om White's voorstel over te nemen om een verantwoordelijkheidsbegrip niet te baseren op een metafysisch gecreëerd zelf, maar op de wenselijkheid van zoiets als vrije wil: "een actor had iets anders kunnen doen dan hij of zij gedaan heeft, in het geval dat het toerekenen van verantwoordelijkheid en schuld aan die actor voor de onderhavige reactie gerechtvaardigd is". In die rechtvaardiging zoeken we dan geen getuigenis bij een metafysicus om te bepalen of je anders had kunnen doen, maar kijken naar iemands bekwaamheid en verzachtende omstandigheden.

Ik meen met Dennett dat we het beeld van een Kantiaanse onveroorzaakte zelf naar beneden bij moeten stellen, om te zien dat we nog meer dan voldoende aan vrije wil en bekwaamheid over houden die het waard is om te willen hebben, en om een acceptabel verantwoordelijkheidsbegrip op te kunnen baseren. In dat geval zal zo'n beeld ook beter bestand blijken tegen nieuwe wetenschappelijke onderzoekresultaten. In plaats van ons af te zetten tegen de wetenschap moeten we haar onderzoekresultaten juist preventief gebruiken om aldus opgehelderde tekortkomingen vroegtijdig te signaleren en hier maatregelen tegen te kunnen nemen. Als we begrijpen hoe een zelf oorzakelijk tot stand komt door factoren in de omgeving die deelhebben aan de vorming van een persoon, dan kunnen we die voorwaarden aanpassen die bijvoorbeeld criminele persoonlijkheid vormt. Zodoende kunnen we ook ongelijkheid in verdeling van geluk compenseren, of het nu socio-economische of biologische deficiënties betreft. Daarbij bedoel ik niet dat bijvoorbeeld alle psychische stoornissen gelijk biomedisch moeten worden aangepakt. Hermeneutische verklaringen en behandelingen zijn net als de volkpsychologie gebaseerd op modellen die goed kunnen

werken. Maar al met al moet je in ieder geval wel accepteren dat een zelf mede gevormd wordt door de causale interacties gedurende de ontwikkeling van een individu. Dat wordt moeilijker als we vasthouden aan een metafysische (onveroorzaakte) zelf.

Ik deel de kritiek van Habermas dat in onze moderne samenleving de aandacht te veel op efficiëntie en te weinig op het evalueren van de (de-facto gedepolitiseerde) doelen is gericht. Maar dat is volgens mij niet te wijten aan de verwaarlozing van een bijzondere vorm van rationaliteit maar aan een te ver doorgeschoten liberalisering, en het misverstand dat als er geen objectieve waarden zouden zijn het ook niet meer mogelijk is, en het er toe zou doen, om met elkaar te bepalen waar we gezamenlijk onze doelen moeten plaatsen.

Maar de hiertoe door Habermas aangeroepen aanvullende vorm van communicatieve rationaliteit levert te veel een gerationaliseerde mensbeeld op. Het ontkent te veel dat we niet alleen mensen van ratio maar vooral ook van gevoel zijn. Niet alles is te vangen in de dwangloze dwang van het betere argument.

We zijn in de natuur geëvolueerde wezens die in groepen leefden. Reeds voorafgaande aan de ontwikkeling van de taal moet er al in de instincten iets als groepsverbandbevordering zijn gevormd. Diep van binnen zijn we nog steeds door gevoel gestuurde wezens. We kunnen dat ontkennen door te denken dat alles, emotie, passie, esthetica, binnen het bereik van het argument ligt maar ik denk dat dat onjuist is. Ik sluit me aan bij Hume's uitgangspunt dat niet de rede (die in zichzelf inert is) maar het gevoel immer onze drijvende kracht is. Juist door dat niet te onderkennen, en in het verstand te blijven zitten raken we in onze doelen verwijderd van de bron: de emotie.

De wortels van de moraal moeten daarom worden gevonden in de passie en de empathie.

Consensus in het gezamenlijk bepalen van doelen die het ons waard zijn om nagestreefd te worden, kan ook tot stand komen door invoelen via de mechanismen van de empathie (zoals ik heb laten zien aan de hand van de moraalpsychologie van Hume). Daarmee komt ook meer aandacht te liggen op een ethiek van zorg in plaats van een ethiek van plichten.

Mijn kritiek op Dennett's theorie is dat ook hij de rol van die empathie teveel onderschat. Mensen zijn in zijn reconstructie van de evolutie van de vrije wil en moraal veelal bezig met de maximalisatie van het eigen profijt. [sectie 6.3]

Maar voor de verdiscontering van emotie als bron van de moraal hebben we de centrale rol van de subjectieve ervaring wel nodig. En in dat opzicht lijkt Dennett deze weer te veel te hebben uitgehold. Hoe dat te repareren is echter het onderwerp voor een nieuw paper.

Want als we toegeven dat er zoiets als een subjectief perspectief is, hoe moet dit dan tot uitdrukking komen als je meent dat voor een beschrijving en verklaring van alles, het objectieverende perspectief voldoende is? Het antwoord zou kunnen zijn dat hoewel het verloop van de processen in de wereld vanuit een geobjectiveerd perspectief in principe is te voorspellen, dit wegens onze epistemisch beperkte horizon praktisch onmogelijk is

We worden dan automatisch gestuurd in de richting van andere modellen zoals diegene die rekening houden met een subjectief perspectief. Of dit perspectief nu een speciaal geestesoog is of narratief tot stand komt, in de praktijk houden we rekening met het gegeven dat ieders perspectief op de wereld uniek is, omdat ieder een eigen historie heeft. En om rekening te houden met al deze perspectieven, blijft communicatie met de concrete-ander en gezamenlijke instemming in de besluitvorming wel van groot belang.

9 Literaturopgave en voetnoten

¹ De informatie in deze paragraaf is afkomstig uit de volgende bronnen;

- Edwards A.R., Pröpfer I.M.A.M. ; “Pluriforme bestuurskunde: voorbij redeloos relativisme”, in Bestuurskunde, No 3, 1996, pp 156-166
Internet ; 07-2005 , "Een inleiding in de Sociologie" , KU Leuven ;
#http://www.pedagogie.student.kuleuven.ac.be/studiehulp/1kan_Soc.doc
T. Heysse, S. Rumens, R. Tinnevelt ; “Habermas / Een inleiding op zijn filosofie van recht en politiek”, (Pelckmans: Kapellen, 2007)

² Het formele karakter van de theorie komt tot uitdrukking in het gegeven dat ze alleen uitspraken doet over de formele procedures waarin (morele) normen tot stand moeten komen. De inhoud van die (morele) normen wordt overgelaten aan het feitelijk discours.

³ In feite was dat ook Habermas' kritiek op Adorno en Horkheimer. Hun opvattingen hoorden thuis in de Cartesiaanse 'filosofie van het subject of van het bewustzijn' waarin een louter subject-object perspectief wordt ingenomen. Zij vatte de maatschappij ten onrechte op als een reusachtige machine met de geschiedenis als ontwikkelingsproces waardoor ze een ontoereikende beschrijving van onze samenleving gaven. Daardoor trokken ze ten onrechte de conclusie dat kritiek niet meer mogelijk was. Het maakte ze blind voor die ander vorm van rationaliteit; die van het communicatieve handelen. In het zelfde licht beschouwd is het begrijpelijk dat Habermas de objectiverende benadering van de neurowetenschap met argusogen bekijkt, en bezwaar aantekent als deze wetenschap, louter op grond van haar eigen onderzoeksresultaten verklaart dat het mensbeeld de komende tientallen jaren aanzienlijk zal moeten worden bijgesteld. Het staat haaks op wat hij van een goede analyse van de mens en zijn samenleving verwacht. De mens tot ding reducerende, natuurwetenschappelijke bijstelling van het mensbeeld is nu precies wat een emancipatoire kritische filosofie zou moeten ontmaskeren en bekritisieren.

- ⁴ T. Heysse, S. Rumens, R. Tinnevelt “Habermas / Een inleiding op zijn filosofie van recht en politiek”, (Pelckmans: Kapellen, 2007), P36, 46

⁵ dat wat we moeten veronderstellen de onderliggende grond te zijn van wat we kennen door de zintuiglijke waarneming.

⁶ Wilhelm Dilthey (1833-1911) wist deze eenzijdigheid zo treffend te typeren toen hij voor een pleidooi voor zijn hermeneutiek het volgende schreef; “door de aderen van het kennende subject dat Locke, Hume en Kant construeerden, vloeit geen echt bloed maar enkel het verdunde sap van de rede als niet meer dan een denkactiviteit. Ikzelf echter ben, door me historisch bezig te houden met de mens als geheel, ertoe gebracht om dit welwillende, voelende en zich verbeeldende wezen in alle verscheidenheid van zijn vermogen ook tot het fundament van kennis te maken.” [Matthias Voigt, 'Wilhelm Dilthey' *Handboek van de Filosofie*, blz 226]

- ⁷ J. Habermas; “Truth and Justification / Realism after the Linguistic Turn”, Polity Press: Cambridge, 2003), P3

⁸ De representerende ofwel afbeeldende functie van taal heeft betrekking op het afbeeldende aspect ervan; de functie waarin een propositie een afbeelding van iets in de werkelijkheid vormt. Vanuit deze optiek begrijpen we een zin of propositie als we de voorwaarden kennen waaronder deze waar is. Willen we dus bepalen wat de betekenis is van de woorden of zinnen in onze taal, dan hoeven we alleen maar te bepalen wat in de wereld het woord of de zin bevredigt (waar maakt). En de zintuiglijke waarneming is daarvoor het middel. De enige relatie die we hier hebben is dus tussen een teken en iets in de werkelijkheid. Dat is wat de betekenis van termen fixeert. In zo'n benadering van betekenis laat betekenis zich uiteindelijk ook formaliseren.

9 Arnold Burns & Herman van Dijn ; “Meesters van de Westerse Filosofie, Ludwig Wittgenstein”, Teleac: Vonksteen, Langemark, 1988), P68

10 T. Heysse, S. Rumens, R. Tinnevelt ; “Habermas / Een inleiding op zijn filosofie van recht en politiek”, (Pelckmans: Kapellen, 2007), P37-38

¹¹ Habermas verbindt taal, rationaliteit, interactie en zelfs de interpretatie van vreemde samenlevingen en culturen met argumentatie, discussie en het bereiken van overeenstemming. Een discussie kunnen we rationeel noemen wanneer ze verloopt volgens de regels van de ideale gespreksituatie. In die zin drukken die regels Habermas’ interpretatie van (communicatieve) rationaliteit uit. De regels vormen de normen inzake de rationaliteit die mensen in onze maatschappij aanvaarden.

12 T. Heysse, S. Rumens, R. Tinnevelt ; “Habermas / Een inleiding op zijn filosofie van recht en politiek”, (Pelckmans: Kapellen, 2007), P81

13 Thomas Mautner ; “Dewey, Dictionary of Philosophy”, (Penguin books; London, 2000), P138

¹⁴ Dewey beweerde dat de theorie van onderzoek (inquiry) veel van de meest eerbiedwaardige filosofische woordraadsels oplossen. *’ Sommige van het ongegronde dualisme waarmee is afgedaan.. zijn die van, het objectieve en subjectieve, het werkelijke en het schijnbare, het mentale en het fysische, natuurwetenschappelijke objecten en waarnemingsobjecten, dingen van ervaring en achter de ervaring verborgen dingen-in-zichzelf, waarvan de laatste een ondoordringbare sluier zou zijn die de cognitieve toegang tot de dingen van de natuur verhindert. De bron van deze dualismes ... is isolatie van cognitieve ervaring en zijn subject-materie vanuit andere toestanden van ervaring en hun subjectmaterie, Deze isolatie leidt onvermijdelijk naar tegenstrijdigheid van de dingen van gewone kwalitatieve ervaring... [Dewey] M.a.w. veel traditionele problemen van de filosofie verschijnen niet in onvervalste ervaring maar reflecteren meer de verwarring van filosofen over de aard van de natuur. Zo suggereert Dewey dat het 'probleem van de existentie van een externe wereld', “kunstmatig gegeneerd is door de soort van premissen die ik epistemologisch noem. Wanneer we handelen en omgevingsobjecten in koppige oppositie vinden ten opzichte van onze verlangens en inspanningen, is de uitwendigheid van de omgeving ten opzichte van het ‘Ik’(het zelf) een directe constituent (het wezen bepalende) van een directe ervaring.’ [Wallace Matson;” A New history of Philosophy” (Harcourt College Publishers; 200)0, p532]*

15 J. Habermas; “Truth and Justification / Realism after the Linguistic Turn”, (Polity Press: Cambridge, 2003), P26

¹⁶ Transcendentale filosofie is voor Kant een onderzoek naar de noodzakelijke mogelijkheidsvoorwaarden van kennis. Het resultaat van het onderzoek is een theorie van het bereik en de objectieve geldigheid van menselijke kennis. Volgens Kant zijn deze mogelijkheidsvoorwaarde gelegen in het synthetisch a priori van a) ruimte en tijd als subjectieve aspecten van het verstand, b) de begripsordenende vormen (de categorieën) van het verstand, (beiden van toepassing op de vorming van het fenomenale beeld van de wereld van de natuur), en de praktische rede voor de morele handelingen. Contrasteer dat met Transcendent, dat wat buiten het bereik van enige mogelijke ervaring ligt; voorbij de beperking van de wereld van de ervaringen. Er kan volgens Kant geen ervaring zijn van iets wat transcendent is.

17 J. Habermas; “Truth and Justification / Realism after the Linguistic Turn”, (Polity Press: Cambridge, 2003), P15

¹⁸ Quine schrijft, “Er is geen betere methode dan de wetenschappelijke methoden voor het beoordelen van de claims van de wetenschap, en er is noch behoefte nog noodzaak aan plaats voor een eerste filosofie, zoals metafysica of epistemologie, welke de wetenschap of de wetenschappelijke methode kan rechtvaardigen. Daarom moeten filosofen zich ook vrij voelen om gebruik te maken van de resultaten van de wetenschap voor hun eigen doelen, terwijl ze zich ook vrij moeten voelen kritiek te leveren als die claims ongegrond, verwarrend, of inconsistent zijn. Naturalisme is niet een dogmatisch geloof dat de moderne wetenschap geheel correct is. In plaats daarvan beweert het simpel dat de

processen van het universum een wetenschappelijk verklaring hebben en dat het die processen zijn die de wetenschap in haar streven wil verklaren.[W.V. Quine]”

19 J. Habermas; “Truth and Justification / Realism after the Linguistic Turn”, (Polity Press: Cambridge, 2003), P21

20 J. Habermas; “Truth and Justification / Realism after the Linguistic Turn”, (Polity Press: Cambridge, 2003), P26

21 J. Habermas; “Truth and Justification / Realism after the Linguistic Turn”, (Polity Press: Cambridge, 2003), P24

22 J. Habermas; “Truth and Justification / Realism after the Linguistic Turn”, (Polity Press: Cambridge, 2003), P24

23 Habermas 2003, P24

24 Jürgen Habermas ; “Truth and Justification / Realism after the Linguistic Turn”, (Polity Press: Cambridge, 2003), P24

25 Indien een niet-fysisch (verklaarbaar) mentaal proces in staat is om fysische processen te beïnvloeden is dit in strijd met de causale geslotenheid van een systeem en op z'n minst de eerste en tweede hoofdwet van de thermodynamica (resp. de wet van behoud van energie en de stelling dat entropie in een gesloten systeem niet kan verminderen). Causaliteit heeft o.a. betrekking op krachten die arbeid verrichten en zo een rol spelen in de energiebalans. De aanname van causale geslotenheid is gebaseerd op het fundamentele idee van de natuurkunde dat de totale energie van een systeem behouden blijft.

26 Eigenschapdualisme is de doctrine binnen de filosofie van de geest die er van uitgaat dat ons brein twee verschillende klassen van eigenschappen bezit: mentale en fysische. Het mentale is afhankelijk van en wordt gedetermineerd door het organische substraat, zonder dat er sprake is van absoluut reductionisme als bij fysicalisme. Mentale fenomenen superveniëren op fysische processen. Kim spreekt in dit verband van ‘minimal physicalism’. Maar superveniëntie kan nog niet uitleggen hoe het mogelijk is dat mentale eigenschappen, hoezeer ook geworteld in de neurobiologie, maar er tegelijkertijd van onderscheiden, oorzakelijke invloed op de hersenen kunnen uitoefenen. Kim noemt drie principes die ons huidige wereldbeeld domineren. (1) Causale geslotenheid van het fysische. Iedere fysische gebeurtenis heeft een fysische oorzaak. (2) Causale exclusie. Het zijn uitsluitend fysische gebeurtenissen en geen andere, die veranderingen in de wereld kunnen bewerkstelligen. (3) Generatieve exclusie. Als gebeurtenis ‘a’ plaatsvindt, of eigenschap ‘b’ ontstaat, beide als gevolg van ‘c’, dan is het optreden van ‘a’ uitsluitend te begrijpen in termen van ‘c’ en niet op grond van iets dat losstaat van ‘c’. Als we dit aanvaarden is het volgens Kim volstrekt onmogelijk conform het wereldbeeld van de causale geslotenheid van het fysische, dat immateriële processen, (bijvoorbeeld superveniërende mentale fenomenen) energie overbrengen en fysische processen in de hersenen in gang zetten. Volgens Kim moet de conclusie zijn dat het mentale binnen het domein van het fysische moet worden ingelijfd. Wanneer een mentale gebeurtenis M een fysische gebeurtenis P veroorzaakt, dan kan dat omdat M superveniëert op het fysische proces P*. Maar in werkelijkheid is het P* dat ervoor zorgt dat P veroorzaakt wordt. Erkenning van mentale veroorzaking kan alleen als we de aanname van causale geslotenheid opgeven. Doen we dat niet dan moeten we in het eigenschapdualisme accepteren dat mentale fenomenen causaal impotent zijn. [Bronnen; Kim J, “How can my mind move my limbs? Mental Causation from Descartes to contemporary physicalism, *Philosophical Exchange, Vol 30 / Johan A. den Boer, ‘Neurofilosofie, hersenen, bewustzijn, vrije wil, Uitgeverij Boom Amsterdam, 2003*].

27 De informatie betreffende emergentisme is hoofdzakelijk afkomstig uit de volgende bronnen;

Internet ; 06-2007 , "Emergent Properties" , Stanford Encyclopdia ;

#<http://www.plato.stanford.edu/entries/properties-emergent>

Internet ; 06-2007 , "Emergentie, fysicalisme, superveniëntie" , Gerrit van Daltsen ;

#<http://www.home.planet.nl/~gdalfts/pages/gdh003.htm>

²⁸ J.S. Mill geloofde dat verschillende chemische reacties (die indertijd onvoldoende werden begrepen) voorbeelden van emergentisme waren. Later bleken deze reacties echter bevredigend reductionistisch te verklaren. De duale opvatting van licht, als deeltjes- of als golfverschijnsel, vormde aanvankelijk een probleem maar is uiteindelijk geünificeerd onder een brugwet. Aan het einde van de 19^{de} eeuw meende Bergson dat sommige eigenschappen van de levende organismen zo speciaal waren dat die niet uit de elementaire wetten van de fysica konden worden verklaard. Een emergerende eigenschap ‘elan-vital’ moest de processen van metabolisme, groei, zelfherstel, zelfverdediging en reproductie helpen verklaren. Inmiddels heeft de biologie afdoende aangetoond dat dit soort processen te verklaren zijn met in principe dezelfde fysische en (bio)chemische basiswetten waarmee we anorganische processen kunnen verklaren.

²⁹ De informatie betreffende intentionaliteit in deze paragraaf is hoofdzakelijk afkomstig uit de volgende bronnen;

Jan Sleutels; “Intentionaliteit / Verschenen in Wijsgerig Perspectief, ”, (Universiteit Leiden: 2001),
William G. Lycan ; “Instrumentalism / Mind and Cognition, An Anthology”, (Blackwell Publishing: Oxford, 2003), P73-74

³⁰ Brentano merkte op dat bewustzijnsakten zoals geloven, verlangen, oordelen en waarnemen steeds gericht zijn op iets wat wordt geloofd, verlangd, beoordeelt of waargenomen. Kenmerkend voor het intentionele object is dat het bestaan ervan buiten onze geest niet wezenlijk is voor de bepaling van wat er zich in onze geest afspeelt. Of wat je gelooft of verlangt ook werkelijk het geval is, is voor de act van het geloven of verlangen als zodanig niet van belang; de gerichtheid van de act wordt er immers niet minder op wanneer het beoogde object niet een werkelijk bestaan ding of een feitelijke stand van zaken is. De intentionele psychologie van Brentano legde, vooral via diens leerling Husserl, de basis voor de filosofische fenomenologie, waarin de vraag naar de betekenis of object-gerichtheid van de bewustzijnsinhouden vooraf gaat aan de vraag of het bewustzijn de ‘buitenwereld’ op correcte wijze weerspiegelt.

³¹ Daniel Dennett ; “True Believers / in William G. Lycan, Mind and Cognition, An Anthology”, (Blackwell Publishing: Oxford, 2003), P75-87

³² Voor een meer accurate en gedetailleerde beschrijving, zoals in de gevallen dat we de computer buiten zijn ontwerpspecificaties gebruiken (bijvoorbeeld bij een lage temperatuur) of in geval van een defect van een der onderdelen, moeten we afdalen naar een lager abstractieniveau, dus de fysische houding aannemen, om een voorspelling van zijn gedrag te kunnen doen. Alleen het ontworpen gedrag is voorspelbaar vanuit de ontwerphouding.

³³ Idem,

³⁴ Dit doet denken aan de opvatting van Churchland dat mentale termen beschouwd moeten worden als gepostuleerde innerlijke toestanden waarvan de onderlinge causale relaties en de relaties met stimuli en gedrag worden beschreven door de wetten van de volkspychologie. Het komt er volgens hem op neer dat we voor het gemak aannemen dat mensen overtuigingen en wensen hebben omdat we daarmee hun gedrag beter kunnen verklaren. Deze postulaten zijn handig in de theoretische afleiding van toekomstig gedrag, maar verwijzen verder niet naar iets bestaands. Voor Dennett zijn de toegeschreven mentale toestanden echter niet louter theoretische postulaten, maar een soort van abstracta – “berekening gebonden entiteiten ofwel logische constructen”.

³⁵ Daniel Dennett ; “Real Patterns / in William G. Lycan, Mind and Cognition, An Anthology”, (Blackwell Publishing: Oxford, 2003), P100-119

³⁶ Daniel Dennett ; “True Believers / in William G. Lycan, Mind and Cognition, An Anthology”, (Blackwell Publishing: Oxford, 2003), P85

³⁷ Voor sommige doeleinden moet je een systeem fysisch bekijken, bijvoorbeeld wanneer er iets aan de hardware stuk is. Dat perspectief is echter nutteloos wanneer je de logica van het gedrag wil

doorgronden. Voor dat laatste is het intentionele perspectief nodig, dat op zijn beurt nutteloos is wanneer er iets te repareren valt.

38 Daniel Dennett ; “Real Patterns / in William G. Lycan, *Mind and Cognition, An Anthology*”, (Blackwell Publishing: Oxford, 2003), P112

39 Een interpreterende en sturende homunculus postuleren heldert niets op. Alle van het brein te verklaren vermogens zijn nu verschoven naar de homunculus. Dan moet je die homunculus weer verklaren, met een homunculus in een homunculus ...? Zoiets leidt tot een oneindige regressie.

40 Deze metaforische kijk betreft de beschrijvingen van het mentale van de volkspychologie. Het gedrag van mensen wordt begrepen en gestuurd met de, grove, instrumentele regels van de volkspychologie. Daarin zijn de verwijzingen naar mentale gebeurtenissen of toestanden geen verwijzing naar een bestaande entiteit (geest). De volkspychologie is de taal behorende bij de intentionele houding, een perspectief waarin men aan mensen intentionaliteit toeschrijft. Volkspychologie zou men in Dennet's ogen als een korte en bondige ruwe "hogere programmeertaal" de taal van de gebruikersinterface, kunnen beschouwen.

41 Daniel Dennett ; “De evolutie van de vrije wil” (Uitgeverij Contact: Amsterdam / Antwerpen, 2004), P286

42 Fysicalisme, de moderne versie van het materialisme, is de visie dat alles wordt geconstitueerd door aan wetten gehoorzame entiteiten zoals die door de natuurwetenschappen worden aanvaardt. Niets ontrekt zich aan de natuur van de natuurwetenschap. De term wordt hoofdzakelijk in de filosofie van de geest gebruikt voor de visie dat al het mentale uiteindelijk iets fysisch is. In het verleden was het reductionistisch; de visie dat alle beweringen over niet-fysische entiteiten of eigenschappen konden worden vertaald in beweringen van de natuurwetenschap (of in beweringen die fysische predikaten toekennen aan fysische objecten). Sinds de 70er jaren zijn antireductionistische varianten ontwikkeld waarin alles wat er bestaat fysisch is zonder te claimen dat het vertalen in de fysicalistische taal altijd mogelijk is. Het heet dan dat mentale eigenschappen superveniëren op fysische eigenschappen. In de filosofie van de geest kan een onderscheid worden gemaakt tussen type- en token fysicalisme. Type-fysicalisme vertolkt de visie dat soorten van fysische toestanden identiek zijn met soorten fysische toestanden. Token-fysicalisme is de visie dat elke token van een mentale toestand identiek is aan een of andere fysische toestand, maar erkent dat het niet mogelijk hoeft te zijn om mentale typen te identificeren met fysische typen. [*Penguin Dictionary of Philosophy, p424*]

43 Daniel Dennett ; “De evolutie van de vrije wil” (Uitgeverij Contact: Amsterdam / Antwerpen, 2004), P28-34

44 Jürgen Habermas ; “Language Game of Responsible Agency/ How can epistemic dualism be reconciled with ontological monism?”, *Philosophical Explorations*, Vol. 10, No1, March 2007, (Routledge, Taylor & Francis Group), P41

45 De opmerking moet hier wel worden gemaakt dat als Habermas van mening is, maar dat is mij niet geheel duidelijk geworden, dat de subjectieve wereld van gevoelens, intenties en verlangens een geprivilegieerde toegang veronderstellen, (dit lijkt twijfelachtig omdat hij het eerste-persoonsperspectief heeft getransformeerd in het tweede-persoonsperspectief) het de vraag is of hij de aanspraak op geldigheid daarvan wel als discursieve geldigheidsaanspraak heeft bedoeld.

46 Daniel Dennett ; “Sweat Dreams; Philosophical Obstacles to a Science of Consciousness”, Bradford Book, (The MIT press: Cambridge, Massachusetts, London England, 2005)

47 Daniel Dennett ; “Sweat Dreams; Philosophical Obstacles to a Science of Consciousness”, Bradford Book, (The MIT press: Cambridge, Massachusetts, London England, 2005), P25

48 Spreken over op- en neerwaartse veroorzaking verwijst naar het plaatje waarin op een grondniveau het verloop van de causale proces lijn tussen opeenvolgende fysische gebeurtenissen wordt afgebeeld,

en daar boven parallel aan het verloop van de proceslijn van mentale gebeurtenissen wordt getekend. Dit plaatje doet het valse vermoeden bestaan dat beide processen binnen een ontologie naast elkaar bestaan. Voor Dennett zijn mentale toestanden een soort van abstracta die een relatie hebben met gedragspatronen van het subject of object waaraan deze mentale toestanden worden toegeschreven.

⁴⁹ zoals iemand die begaat als deze na een rondleiding langs de collegezalen, de studenten, de bibliotheek, de docenten en professoren, de kantine etc, zou vragen waar nu eigenlijk de universiteit is. Dit voorbeeld is van Gilbert Ryle, om de vergissing duidelijke te maken die iemand begaat als hij spreekt over lichaam en geest, i.p.v. lichaam of geest. Lichaam en geest zijn twee verschillende manieren om over één en dezelfde ontologie te spreken.

⁵⁰ Dr. Frank Rösler et al ; “Das Manifest; Elf führende Neurowissenschaftler über Gegenwart und Zukunft der Hirnforschung”, (Gehirn & Geist, 6/2004), P31

⁵¹ Dit naar de verwoording van Habermas. Ik herken me zelf geloof ik veel meer in het beeld van een primair door gevoel geleid wezen. Mijn gevoel bepaald de doelen die het waard zijn om na te streven, terwijl mijn ratio in dienst staat van mijn gevoel en optreedt als instrument om me te helpen mijn gestelde doel te verwezenlijken.

⁵² Internet ; “Galen Strawson; 1998, 2004, *routledge Encyclopedia of Philosophy* “ (2004, 1998)

⁵³ In zijn discoursethiek, een formalistische theorie die de principes en procedures voor de constructie van geldige morele normen op basis van argumentatie in een open discussie geeft, heeft Habermas eerder met behulp van de techniek van de performatieve contradictie de vooronderstellingen van deelname aan zo'n discussie gegrondvest.

De techniek van de performatieve contradictie bestaat er in om aan te tonen dat ontkenning van de vooronderstellingen in strijd is met de praktijk van de argumentatie zelf, en dus leidt tot een performatieve contradictie. Ik geef een voorbeeld van deze techniek; Eén van die vooronderstellingen is, dat alle participanten oprecht zijn. Iemand begaat nu een performatieve contradictie als hij een ander persoon ‘wil overtuigen met behulp van een leugen’. Iemand overtuigen betekent immers dat je iemand doet geloven in de waarheid van een bepaalde stelling door goede argumenten aan te brengen. Hierbij is cruciaal dat diegene die je tracht te overtuigen veronderstelt dat de argumenten die je aanbrengt waarheidsgetrouw zijn. Maar zodra iemand twijfelt aan de oprechtheid van zijn gesprekspartner, zal hij zich niet meer laten overtuigen. De oprechtheid van alle gesprekspartners is dus een noodzakelijke vooronderstelling van de praxis van de argumentatie.

⁵⁴ Jürgen Habermas ; “Language Game of Responsible Agency?”, *Philosophical Explorations*, Vol. 10, No1, March 2007, (Routledge, Taylor & Francis Group), P14

⁵⁵ Dr. Frank Rösler et al ; “Da Manifest; Elf führende Neurowissenschaftler über Gegenwart und Zukunft der Hirnforschung”, (Gehirn & Geist, 6/2004), P36

⁵⁶ Timothy Schroeder ; “Reflection, Reason, and Free Will”, *Philosophical Explorations*, Vol. 10, No1, March 2007, (Routledge, Taylor & Francis Group)

⁵⁷ Dr. Frank Rösler et al ; “Das Manifest; Elf führende Neurowissenschaftler über Gegenwart und Zukunft der Hirnforschung”, (Gehirn & Geist, 6/2004), P31

⁵⁸ Dr. Frank Rösler et al ; “Da Manifest; Elf führende Neurowissenschaftler über Gegenwart und Zukunft der Hirnforschung”, (Gehirn & Geist, 6/2004), P35

⁵⁹ Jürgen Habermas ; “Language Game of Responsible Agency?”, *Philosophical Explorations*, Vol. 10, No1, March 2007, (Routledge, Taylor & Francis Group), P33

⁶⁰ In de filosofie van de geest propageert Davidson een "non-reductive materialism" welke hij "anomalous monism" noemt. Ofschoon elke mentale gebeurtenis identiek is met een of andere neurofysiologische gebeurtenis, zijn er geen ware universele statements van de vorm "mentale

gebeurtenissen van soort F, zijn identiek met neurofysiologische gebeurtenissen van soort G". Op gelijke wijze bestaan er geen universele wetten die overtuigingen en verlangens verbinden met handelingen, ofschoon handeling wel door overtuigingen en verlangens worden veroorzaakt. Kortom Psychologie is niet reduceerbaar tot fysica.

⁶¹ Ceteris paribus (Latijn voor), onder overigens gelijke omstandigheden. In ieder geval voor Habermas ceteris paribus vanuit het participant perspectief. Een participant in het taalspel van verantwoordelijk agency vooronderstelt in het wege van redenen en argumenten de beschikking te hebben over alternatieve mogelijkheden.

⁶² Jürgen Habermas ; "Language Game of Responsible Agency?", *Philosophical Explorations*, Vol. 10, No1, March 2007, (Routledge, Taylor & Francis Group), P17

⁶³ Jürgen Habermas ; "Language Game of Responsible Agency?", *Philosophical Explorations*, Vol. 10, No1, March 2007, (Routledge, Taylor & Francis Group), P20

⁶⁴ Jürgen Habermas ; "Language Game of Responsible Agency?", *Philosophical Explorations*, Vol. 10, No1, March 2007, (Routledge, Taylor & Francis Group), P29

⁶⁵ Matthias Vogt ; "Pragmatisme / Handboek Filosofie", (Rebo International; 2005), P222

⁶⁶ De hermeneutische methode moet een rationele reconstructie geven van de structuren van de leefwereld vanuit het perspectief van de participanten. Ze gebruikt daarvoor de taal van volkspychologie en de rationaliteit van het communicatief handelen. Ze doet dit zodanig dat ze als het ware mede-participanten van het relaas van deze historie door middel van rechtvaardigende argumenten kan overtuigen, en zich door tegenargumenten zelf kan laten overtuigen. Doordat ze zich op deze wijze begeeft in de deze (fictieve) deliberatieve praktijk, verplaatst ze zich in de positie van de 'gesitueerde-ander' en verstaat daarmee de betekenis-als-relevantie voor de participanten. Door deze wijze van interpreteren komen alle relevante aspecten waaronder de subjectieve beleving van betekenis bloot te liggen.

⁶⁷ In een universum wat niet deterministisch is zijn de gebeurtenissen die plaatshebben onvoorspelbaar. In zo'n universum, waarin men niet zou kunnen anticiperen op wat komen gaat, is men overgeleverd aan toeval en onvermijdelijkheid. Determinisme van het universum is dus een voorwaarde voor vermijdbaarheid en dat laatste is voor Dennett synoniem met vrijheid.

⁶⁸ Stelling van Van Ingwagen.

⁶⁹ Laplace's demon concretiseert een interessant probleem. Geen enkel informatieverwerkend systeem kan een complete beschrijving van zichzelf hebben, hoe de laatste details te representeren van het representeren van het representeren...dus zelfs Laplace's demon heeft een epistemische horizon en als gevolg daarvan niet in staat zijn eigen handelingen te voorspellen op de manier waarop hij de volgende situatie van het universum (waar hij buiten moet staan) kan voorspellen. [*Dennett, 'Evolutie van de vrije Wil', p112*]

⁷⁰ Een universum is deterministisch als er natuurwetten zijn die precies bepalen welke toestandbeschrijving volgt op de daaraan voorafgaande. In een indeterminisch universum hebben de wetten een stochastisch karakter (op kans gebaseerd) en in nihilistische universum is er geen regelmaat.

⁷¹ In modale logica is zeggen dat determinisme waar is, is zeggen dat onze actuele wereld zich in een deelverzameling werelden bevindt met de volgende eigenschappen; er bestaan geen twee werelden die precies eender beginnen (als ze wel eender beginnen zijn het geen twee verschillende werelden), en wanneer twee werelden precies een of andere toestandbeschrijving delen, delen ze alle erop volgende beschrijvingen.

⁷² Gedetermineerde mogelijke werelden zijn mogelijke werelden waarin de overgangswetten (wetten die de overgang van de status van de voxels [driedimensionale pixels die of geheel gevuld of geheel leeg zijn] tussen opeenvolgende momentane toestandbeschrijvingen vastlegt) regelmatig zijn en niets aan het toeval overlaten. Bij indeterministische mogelijke werelden hebben de overgangregels een stochastisch karakter (op kans gebaseerd) en in nihilistische / rommelige mogelijke werelden is er geen regelmaat.

⁷³ Daniel Dennett ; “De evolutie van de vrije wil” (Uitgeverij Contact: Amsterdam / Antwerpen, 2004), P112

⁷⁴ Daniel Dennett ; “De evolutie van de vrije wil” (Uitgeverij Contact: Amsterdam / Antwerpen, 2004), P73

⁷⁵ Daniel Dennett ; “De evolutie van de vrije wil” (Uitgeverij Contact: Amsterdam / Antwerpen, 2004), P77

⁷⁶ Meme; Culturele replicator. De term werd geïntroduceerd door Dawkins om een voorbeeld te geven van een andere replicator dan de genen. Definitie van Meme; (afkorting van "mimeme" ... dat wat wordt geïmiteerd (mimiekt)). Krijgt de klank als in Gene (Engels). Meme is een element van een cultuur die kan worden opgevat als informatie die wordt overgedragen door niet-genetische middelen. Het is elke informatie die wordt gekopieerd van persoon tot persoon of tussen boeken, computers of enig ander opslagmiddel. Memen zijn net als genen onderhevig aan mutatie en selectie. Neem bijvoorbeeld een grap. Je hebt er waarschijnlijk duizenden gehoord in je leven maar slechts enkelen onthouden die je ook regelmatig verteld. Dus grappen moeten onderling concurreren om te overleven. Verder kunnen grappen bij het doorvertellen ook wijzigingen ondergaan. (*S. Blackmore, “An introduction to Consciousness”, Hodger & Stoughton, Oxford, P160*)

⁷⁷ Universeel Darwinism: Een door Dawkins (1976) geïntroduceerde term. Standaard Darwinisme betreft evolutie door natuurlijke selectie dat inwerkt op genen. Maar de evolutietheorie kan ook op andere terreinen worden aangewend. Het proces van natuurlijke selectie kan worden beschouwd als een algoritme. Als er sprake is van variatie, selectie en erven, dan moet er sprake zijn van evolutie. Dus als er sprake is van iets dat varieert en selectief wordt gekopieerd. Met ander woorden; je kan andere replicatoren en andere evolutionaire systemen hebben. Dit is wat Dawkins, universeel Darwinisme noemt. (*Suzan Blackmore, Consciousness p. 160*)

⁷⁸ Daniel Dennett ; “True Believers / in William G. Lycan, Mind and Cognition, An Anthology”, (Blackwell Publishing: Oxford, 2003), P75-87

⁷⁹ Jürgen Habermas ; “Language Game of Responsible Agency/ How can epistemic dualism be reconciled with ontological monism?”, Philosophical Explorations, Vol. 10, No1, March 2007, (Routledge, Taylor & Francis Group), P31

⁸⁰ Matthias Vogt ; “Pragmatisme / Handboek Filosofie”, (Rebo International; 2005), P241

⁸¹ Volgens Dennett’s interpretationisme kun je ook aan thermostaten intentionaliteit toeschrijven. “De thermostaat gelooft dat het te koud is dus zal hij de ketel wel in willen schakelen”. Intentionaliteit is niet zo’n hokus pokus als door Habermas wordt voorgesteld.

⁸² Dit wordt geïllustreerd door het bekende voorbeeld van Ryle die over een categoriefout (category mistake) spreekt als hij een vriend over het hele universiteitsterrein heeft rondgeleid, en de collegezalen, de bibliotheek, de campus, de kantine, de studenten en zijn collega’s heeft laten zien, en dan geconfronteerd wordt met de vraag, ‘maar waar is nu de universiteit?’; Naar analogie kunnen we niet in één zin de beschrijvingen van twee verschillende perspectieven als een conjunctie opnemen. Perspectieven binnen een ontologische beschrijving moeten niet als conjunctie, maar als disjunctie verschijnen. Beschrijf je ze als een conjunctie dan krijg je allerlei problemen, zoals bij het epifenomenalisme van het mentale in het eigenschapdualisme.

⁸³ Dus Dennett's oplossing is dat we of consequent een beschrijving van het menszijn geven vanuit het fysische perspectief of consequent vanuit het intentionele perspectief. Bij wijze van voorbeeld is het dus niet 'lichaam en geest' maar 'lichaam of geest'.

⁸⁴ Dennett schrijft in een voetnoot van 'Real Patterns'; "Verschillende reviewers van de draft van dit artikel hebben aangenomen dat de conclusie die ik hier aanvoer is dat het geloven (of zijn inhoud) epiphenomenen zijn (die geen causaal vermogen hebben), maar dat is een misinterpretatie van een simplificatie van de notie van veroorzaking. Als men een patroon vindt van de soort die ik zojuist heb besproken, dan heeft men ipso facto (uit de kracht van de daad, noodzakelijkerwijs) een causaal vermogen ontdekt – een verschil in de wereld welke een daaropvolgend verschil toetsbaar maakt door middel van standaard empirische methoden van variabele manipulatie. Beschouw het vermogen van menigte-trekkende (aandachttrekkende) bordje 'free lunch' geplaatst in de etalage van een restaurant , en vergelijk zijn kracht (vermogen / macht) in een restaurant in de etalage in New York en zijn vermogen in een restaurant in Tokyo. Het intentionele niveau is duidelijk het juiste niveau waarop men zulke causale vermogens kan voorspellen en verklaren; het teken produceert in de ene populatie op meer betrouwbare wijze een zeker geloof dan in de andere populatie, en variaties in de kleur en de typografie van het teken zijn niet zo voorspellend in de variaties in menigte aantrekkend vermogen als de variaties in waarneembare betekenis. Het feit dat de regelmatigheden waarop deze succesvolle voorspellingen zijn gebaseerd. Efficiënt worden ingevangen in intentionele termen en niet worden afgeleid van 'dekkende wetten' laat niet zien dat de regelmatigheden niet 'causaal' zijn; het laat gewoon zien dat filosofen vaak hebben vertrouwd op vastgepinde noties van causaliteit afgeleid van exclusieve aandacht voor enkele voorbeelden uit de fysica en de chemie. Smith wees me erop dat ik hier Aristoteles claim echo dat zijn voorgangers finale oorzaken hadden genegeerd. [*Dennett; 'Real Patterns', Lycan, p 113*]"

⁸⁵ Jürgen Habermas ; "Language Game of Responsible Agency/ How can epistemic dualism be reconciled with ontological monism?", *Philosophical Explorations*, Vol. 10, No1, March 2007, (Routledge, Taylor & Francis Group), P33

⁸⁶ T. Heysse, S. Rumens, R. Tinnevelt ; "Habermas / Een inleiding op zijn filosofie van recht en politiek", (Pelckmans: Kapellen, 2007), P18

⁸⁷ Daniel Dennett ; "De evolutie van de vrije wil" (Uitgeverij Contact: Amsterdam / Antwerpen, 2004), P213

⁸⁸ T. Heysse, S. Rumens, R. Tinnevelt ; "Habermas / Een inleiding op zijn filosofie van recht en politiek", (Pelckmans: Kapellen, 2007), P86

⁸⁹ Opdat we de ultieme schepper van onszelf kunnen zijn zoekt Kane, in bijvoorbeeld zoiets als een kwantumonbepaaldheid, een onderbreking in de dwang van de causale keten die anders tot vòr de geboorte terug te voeren zou zijn.

⁹⁰ Daniel Dennett ; "De evolutie van de vrije wil" (Uitgeverij Contact: Amsterdam / Antwerpen, 2004), P313

⁹¹ Daniel Dennett ; "De evolutie van de vrije wil" (Uitgeverij Contact: Amsterdam / Antwerpen, 2004), P320

⁹² Daniel Dennett ; "De evolutie van de vrije wil" (Uitgeverij Contact: Amsterdam / Antwerpen, 2004), P331

⁹³ De informatie uit deze alinea komt hoofdzakelijk uit:

Thomas Mautner ; "Hume, Dictionary of Philosophy", (Penguin books; London, 2000), P257

⁹⁴ Volgens Rorty bestaat er geen absolute waarheid en geen punt waar het met de wereld naar toe gaat of zou moeten gaan. Alles is bij toeval voortgekomen uit evolutie. Het begrip 'contingentie' neemt een centrale plaats in bij Rorty. 'Contingent' wil zeggen "product van toevallige historische omstandigheden". Al onze theorieën en ideeën zijn contingent. De filosofie moet zich niet bezig houden met zoeken naar universele waarheid maar met de interpretatie van wat eigenlijk culturele producten zijn. Filosofische kennis is volgens Rorty eindig, subjectief, historisch geconditioneerd en

waardegebonden. Het ontbreken van een absolute waarheid wordt door Rorty gezien als iets positiefs, als een bevrijding. Hij ziet streven naar absolute waarheid en het vervolgens leven naar deze waarheid als oorzaak van veel geweld en terreur. In plaats van een absolute waarheid moet er een 'gesprek van de mensheid' ontstaan waarin plaats is voor een pragmatische moraal. Dit gesprek moet leiden tot solidariteit met anderen die lijden en tot het terugdringen van leed. Rorty is een pragmatist. Hij ziet ideeën als instrumenten, als middelen om problemen op te lossen en doelen te verwezenlijken. Dit kunnen problemen en doelen van allerlei aard zijn: filosofisch, wetenschappelijk, moreel of politiek. Volgens Rorty is er geen verschil tussen mens- en natuurwetenschappen. Bij beide worden waarheden niet gevonden maar gemaakt door gezamenlijke inspanning van velen. Feiten ontstaan er evenveel als er talen zijn om deze activiteit te beschrijven. Kennis kan daarom niet los gezien worden van de praktische, sociale en historische context waarbinnen ze ontstaat. Dit geldt zowel voor wetenschappelijke en alledaagse kennis als voor filosofie. [Bron; "Filosofische bijsluiter wetenschapsfilosofie van de twintigste eeuw", #www.fss.uu.nl/wetfil/96-97/achter.htm (6-09-2007)]

⁹⁵ Daniel Dennett ; "Faith in Truth" (Amnesty Lecture, Oxford; 18 Feb 1997 ; P8
#http://ase.tufts.edu/cogstud/papers/faithhint.htm

⁹⁶ In het deductief-Nomologisch model (Hempel) impliceert het explanans a.h.v. natuurwetten de explanandum.

Explanans	C1,C2, C3,...Ck L1, L2, L3,...Lr -----	Beweringen over individuele feiten en initiële condities Algemene wetten
Explanandum E		Beschrijving van te verklaren gebeurtenis, wet, of feit.

[Bron; M. Curd & J.A. Cover, "Philosophy of Science", Norton & Company, New York, 1998, p 769]

⁹⁷ Dat Habermas wellicht in de analyse van onze samenleving Dennett niet (direct) zal overtuigen staat hier los van.

⁹⁸ Daniel Dennett ; "De evolutie van de vrije wil" (Uitgeverij Contact: Amsterdam / Antwerpen, 2004), P300

⁹⁹ Dit strookt ook met de neurowetenschappelijke hypothese dat de oorsprong van de emotie in de oudste hersenen, het zgn. limbisch systeem zetelt. Volgens deze hypothese ontspringen in dit systeem onze diepste drijfveren, die ultiem de waarden bepalen die we nastreven. De neo-cortex gelegen rondom het emotionele brein, vormt evolutionair gezien het jongste brein en is verantwoordelijk voor de hogere verstandelijke vermogens die samenhangen met taalvaardigheden, het plannen, nadenken etcetera. Het limbische systeem geeft onze richting aan van wat we willen met ons leven terwijl de neo-cortex ons de weg geeft (haar functie is instrumenteel) om ons zo verstandig mogelijk vooruit te brengen. Vanuit dit perspectief kunnen de ultieme waarden ontsproten aan het limbische systeem, die buiten het bereik van de beïnvloeding door het instrumentele brein (de neo-cortex) liggen, niet het onderwerp van kritiek vormen. Kritiek en discussie kan alleen betrekking hebben op de instrumentele processen en criteria.

¹⁰⁰ Daniel Dennett ; "Sweet Dreams; Philosophical Obstacles to a Science of Consciousness", Bradford Book, (The MIT press: Cambridge, Massachusetts, London England, 2005), P35