

**Hidde Straasheijm (Studentnummer 292348)**

**Cranes of Consciousness.  
Een bespreking van Daniel C. Dennett's Consciousness  
Explained.**



**Leerstoelgroep: Theoretische Filosofie  
Bachelorthesis Wijsbegeerte  
Scriptiebegeleider: Dr. T. de Mey  
Adviseur: Dr. A. Wouters  
Datum: 20-8-1012  
Studiepunten: 10 Ects.  
Woorden: 12450**

## **Dankbetuiging**

Op deze plaats wil ik graag Tim de Mey en Arno Wouters bedanken. Hun adviezen hebben het schrijven van deze bachelor scriptie een stuk makkelijker gemaakt.

## **Inhoud**

Inleiding P.4

I. Hijskranen en luchthaken.

- a. Darwin en de filosofen P.5
- b. Naturalisme P.6

II. Problemen voor een naturalistische opvatting van het bewustzijn.

- a. Het Mind-body probleem P.7
- b. Intentionaliteit: een test-case voor het behaviorisme P.9
- c. Kunnen machines denken? P.10

III. Het bewustzijn verklaard.

- a. Heterofenomenologie P.11
- b. Het Cartesiaans theater versus Dennett's meervoudige versies P.13
- c. Evolutie van het bewustzijn P.15
- d. De kleren hebben geen keizer P.17

IV. Conclusie: een gesprek met Otto P.19

Werken van Dennett P.22

Literatuur P.23

## **Inleiding**

Volgens Charles Darwin is alleen een goede theoreticus in staat tot het doen van zinvolle waarnemingen. Daniel Dennett is zich hier altijd van bewust geweest. Keer op keer hamert hij er in zijn filosofie op dat ons bewustzijn niet is wat we denken dat het is. Theoretische veronderstellingen over hoe onze natuur in elkaar zit vertroebelen ons blik. Dit is op zijn minst merkwaardig omdat niets in de wereld ons zo vertrouwd is als ons eigen bewustzijn. We hebben immers een directe toegang tot onze eigen gedachtewereld. Hoe zouden we ons in hemelsnaam over dit soort zaken kunnen vergissen?

De filosofie van Dennett gaat rechtstreeks in tegen sommige diepe intuïties die we allemaal koesteren. Het is de bedoeling van deze scriptie om duidelijk te maken hoe hij de ideeën van Darwin gebruikt om een geheel nieuwe visie op het bewustzijn te ontwikkelen die niet minder revolutionair is dan de evolutietheorie honderdvijftig jaar geleden was. Dennett ontwikkelt deze visie in een voortdurende dialoog met wetenschappers afkomstig uit diverse disciplines. Met zijn filosofie probeert hij een bijdrage te leveren aan de ontwikkeling van onze kennis door de houdbaarheid van concepten en metaforen te onderzoeken. Daar waar de uiterste houdbaarheidsdatum overschreden is doet hij voorstellen voor aanpassingen van het begrippenkader dat de wetenschap hanteert. Het feit dat hij zich niet van de wetenschappelijke ontwikkeling afkeert komt de relevantie van zijn werk alleen maar ten goede.

In dit opzicht is het werk van Dennett een verademing. Als student wijsbegeerte word je maar al te vaak geconfronteerd met voorbeelden van hoe het ook anders kan. Binnen sommige filosofische scholen is het helaas mode om je op eenzame hoogte terug te trekken en vanuit die verheven positie neer te kijken op alles wat er in de andere universitaire disciplines gebeurt. De bespiegelingen van deze filosofen missen hierdoor iedere relevantie. Hun werk heeft geen concrete toepasbaarheid en is alleen toegankelijk voor een select groepje ingewijden.

Dennett daarentegen staat met zijn poten in de modder en is bereid vuile handen te maken. Hij schrijft voor een groot publiek dat hij wil overtuigen om anders na te gaan denken over vragen omtrent ons bewustzijn. De impact die zijn werk heeft op de traditionele manier waarop we over dergelijke vragen nadenken is bijzonder groot. Zoals gezegd brengt hij een omwenteling te weeg in het denken over de geest die vergelijkbaar is met de revolutie die Darwin te weeg bracht in het denken over de natuur. Daarbij maakt hij dankbaar gebruik van de inzichten van Darwin. Om deze onderlinge relatie duidelijk te maken zullen we nu eerst ingaan op de filosofische implicaties van het Darwinisme.

## I. Hijskranen en luchthaken

### a. Darwin en de filosofen

*Few words in our language foreshorten intellectual history as much as does the word species. The Greeks, in initiating the intellectual life of Europe, were impressed by characteristic traits of the life of plants and animals; so impressed indeed that they made these traits the key to defining nature and to explaining mind and society. And truly, life is so wonderful that a seemingly successful reading of its mystery might well lead men to believe that the key to the secrets of heaven and earth was in their hands. The Greek rendering of this mystery, the Greek formulation of the aim and standard of knowledge, was in the course of time embodied in the word species, and it controlled philosophy for two thousand years.*

John Dewey (1910, blz.3)

In zijn boek *Darwin's dangerous idea* noemt Daniel Dennett de evolutietheorie van Darwin een 'universal acid' (DDI, p.63). Een zuur dat alles op zijn weg aanvreet, niets ongemoeid laat en de wereld onherkenbaar achterlaat. Wetenschappelijke uitgangspunten en filosofische vooronderstellingen worden erdoor uitgedaagd en slechts zelden ondergaan ze de werking ervan neutraal. Het universele zuur heeft eeuwenoude overtuigingen doen wankelen. Vrijwel iedere theorie met betrekking tot de mens, het leven en de wereld om ons heen heeft een antwoord moeten vinden op het darwinisme.

Deze bijzondere eigenschap van Darwin's theorie werd al vroeg onderkend door zowel medestanders als tegenstanders. Een halve eeuw na het verschijnen van *The origin of species* schrijft de filosoof John Dewey dat de omwenteling die door Darwin in gang gezet is nog maar net is begonnen en dat alleen iemand die de onbezonnenheid van de profeet combineert met de koppigheid van de partisaan een voorspelling zou durven doen over hoe de filosofie er uit zal komen te zien na Darwin (1910, blz.9). Hij waagt zich echter wel aan een eerste overzicht van slachtoffers die de theorie op dat moment al gemaakt heeft. Bepaalde concepten die voor de filosofie gemeengoed waren zijn hun vaste betekenis kwijt geraakt. Het tweeduizend jaar oude begrippenkader waarmee de filosofie probeerde de wereld te verklaren, heeft, nadat het in aanraking is gekomen met het universele zuur, zijn bruikbaarheid verloren.

Één van de concepten die niet opgewassen bleek tegen het universele zuur is het klassieke begrip 'species'. In de Aristotelische oorsprong van dit concept ligt de bewondering van de Grieken voor de universaliteit van het particuliere besloten. Het feit dat ieder exemplaar van een bepaalde soort beschikt over algemene trekken aan de hand waarvan we al het leven op aarde kunnen classificeren kwam hen voor als een wonderlijk gegeven. Op deze algemeenheid probeerden ze hun kennis over de natuur te funderen. Over het veranderlijke en eenmalige kunnen we niet veel met zekerheid zeggen. Het eeuwige en tijdloze leent zich daarentegen bij uitstek voor het doen van wetmatige uitspraken die ogenschijnlijk een absolute zekerheid bezitten. Daarnaast viel het hen op dat de universele eigenschappen waarover iedere soort beschikt op een bepaalde manier afgestemd lijken te zijn op de omgeving waarin de soorten voorkomen. De natuur is niet zomaar een mechanisch geheel maar kent een functionele samenhang die zich niet goed laat verklaren. Het kon welhaast niet anders of deze doelmatigheid was het werk van een intelligente ontwerper die verantwoordelijk was voor het scheppen van de natuur en het aanbrengen van deze wonderbaarlijke orde. Op deze manier kon het begrip 'species' dienst doen als fundering van zowel wetenschap als religie (Dewey, blz.10).

Dit conceptuele bouwwerk heeft, enkele pogingen tot renovatie ten spijt, dienst gedaan tot Darwin in 1859 zijn beroemde *Origin of species* publiceerde. Hierin liet hij zien dat

'soorten' geen statische eenheden zijn die zich buiten de geschiedenis bevinden maar dat iedere soort zijn eigen ontwikkeling doormaakt. Daarnaast droeg hij een mechanisme aan dat deze ontwikkeling kon verklaren: natuurlijke selectie. Hiermee kon hij de doelmatigheid van eigenschappen verklaren zonder dat hij verplicht was een intelligente ontwerper te postuleren. In één klap ontdeed hij zich dus van zowel het idee dat de natuur onveranderlijk is als het idee dat alleen intelligentie het ontwerp in de natuur kan verklaren. Middels natuurlijke selectie kunnen opeenvolgende generaties organismen zich in de loop van de tijd aan hun omgeving aanpassen zonder dat daar een vooropgezet plan voor nodig is. Het verklaart de doelmatigheid van bepaalde eigenschappen door middel van een louter mechanisch proces. Darwin's grootste prestatie bestaat er niet uit dat hij dit proces heeft beschreven, maar dat hij de wetmatigheid ervan heeft aangetoond. Dat natuurlijke selectie optreedt is onontkoombaar als er maar aan een drietal voorwaarden voldaan wordt (DDI, blz.48-60). Ten eerste moet er binnen een populatie een zekere variatie voorkomen. Niet alle leden van een soort zijn identiek, sommige beschikken over eigenschappen die bij andere ontbreken. Ten tweede moet er selectie plaats vinden. Individuen die een eigenschap bezitten waardoor ze zich met een groter gemak voort kunnen planten zullen meer nageslacht produceren. Ten derde moet deze eigenschap overerfbaar zijn. Dit betekent dat er in de volgende generatie meer individuen zullen zijn die in het bezit zijn van deze nuttige eigenschap. Op deze manier zullen soorten zich langzaam aanpassen aan hun omgeving. Deze drie elementen vormen samen een evolutionair algoritme, waarbij de uitkomst gegarandeerd is mits er aan de voorwaarden is voldaan.

## **b. Naturalisme**

Tot op de dag van vandaag kent Darwin enthousiaste medestanders en fanatieke tegenstanders. Daniel Dennett behoort zonder meer tot de eerste groep. Het Darwinisme is een rode draad die door zijn filosofie heen loopt. Zijn enthousiasme is hoofdzakelijk gebaseerd op het feit dat Darwin hem de instrumenten biedt om bepaalde problemen te lijf mee te kunnen gaan. "If I were to give an award for the single best idea anyone has ever had, I'd give it to Darwin, ahead of Newton and Einstein and everyone else. In a single stroke, the idea of evolution by natural selection unifies the realm of life, meaning and purpose with the realm of space and time, cause and effect, mechanism and physical law" (DDI, blz.21). Zoals we gezien hebben was het, voordat Darwin zijn ontdekkingen deed, nagenoeg onmogelijk om de functionaliteit van de levende natuur niet te zien als bewijs van intelligent ontwerp. Het was letterlijk onvoorstelbaar dat de grote mate van complexiteit die kenmerkend is voor levende wezens het gevolg zou zijn van louter toeval. Hoe kon al deze doelgerichtheid het resultaat zijn van een mechanisch proces dat zelf geen enkel doel kent? Darwin's grote prestatie is dat hij een algoritme heeft opgesteld dat functionaliteit, doelmatigheid en betekenis kan verklaren zonder te verwijzen naar een vooraf opgesteld plan.

Op deze manier werd het mogelijk een naturalistische verklaring van de levende natuur samen te stellen. In hetgeen volgt zullen we zien dat Dennett geïnteresseerd is in de vraag hoe een naturalistische theorie van het bewustzijn eruit zou kunnen zien. Daarbij maakt hij dankbaar gebruik van de handvaten die Darwin hem biedt. Een naturalistische theorie wil in dit geval zeggen een theorie die geen intelligentie postuleert die de theorie niet zelf kan verklaren. De klassieke filosofie deed dit door bijvoorbeeld te stellen dat er een intelligente schepper is die verantwoordelijk is voor al het ontwerp in de natuur. Dennett noemt dit een luchthaak (DDI, blz.74). Men postuleert een vorm van intelligentie om iets te verklaren maar heeft vervolgens geen antwoord op de vraag waar deze vorm van intelligentie zelf vandaan is gekomen. Het Darwinisme is daarentegen een goed voorbeeld van wat Dennett een hijskraan noemt (DDI, blz.75): het laat zien hoe intelligentie van de grond af opgebouwd kan worden door een proces dat zelf geen intelligentie bezit. Omdat deze twee termen belangrijk zijn voor de rest van het verhaal volgt hier de definitie van Dennett zelf:

It is time for some more careful definitions. Let us understand that a *skyhook* is a “mind-first” force or power or process, an exception to the principle that all design, and apparent design, is ultimately the result of mindless, motiveless mechanistic. A *crane*, in contrast, is a subprocess or special feature of a design process that can be demonstrated to permit the local speeding up of the basic, slow process of natural selection, and that can be demonstrated to be itself the predictable (or retrospectively explicable) product of the basic process. (DDI, blz.76 Cursiveringen zijn van Dennett zelf).

Een hijskraan is dus een mechanisme dat een stuk functioneel ontwerp (design) kan verklaren. Aan de menselijke geest (mind) is zonder twijfel een ongelooflijke hoeveelheid ontwerp vooraf gegaan. En het is bijna onvoorstelbaar dat een enkele hijskraan, zelfs een krachtige hijskraan zoals natuurlijke selectie, al dit werk in zijn eentje verzet heeft. Daarom introduceert Dennett hier ook de mogelijkheid dat natuurlijke selectie op bepaalde momenten hulp heeft gekregen (the local speeding up of the basic, slow process of natural selection). Maar deze hulp kan niet gekomen zijn van luchthaken. Het is zeer goed denkbaar dat het proces van natuurlijke selectie in de loop van de evolutie gestuit is op onafhankelijke mechanismen die het ontwerpproces versnellen. In dat geval heeft de hijskraan een nieuwe hijskraan gebouwd. Voor deze nieuwe hijskraan geldt natuurlijk ook dat het ontstaan ervan op een mechanische manier verklaard moet worden (itself the predictable (or retrospectively explicable) product of the basic process). Dennett wil dus laten zien dat de menselijke geest het product is van verschillende hijskranen die ieder voor zich een stuk van het ontwerp kunnen verklaren dat nodig is voor een dergelijk verbluffend stuk design. Maar daar laat hij het niet bij, de menselijke geest is namelijk zelf één van de meest indrukwekkende hijskranen die de evolutie heeft voortgebracht. In deel drie zal ik hier nader op ingaan, eerst wil nu enkele klassieke vraagstukken uit de filosofie of mind bespreken om duidelijk te maken met welke obstakels een naturalistische theorie van het bewustzijn moet afrekenen.

## **II. Problemen voor een naturalistische theorie van het bewustzijn**

### **a. Het mind-body probleem**

*Waar onze taal ons een lichaam doet vermoeden, en er geen lichaam is, daar, zouden we willen zeggen, is een geest.*

Ludwig Wittgenstein (1953, Blz. 57)

De vraag hoe geest en lichaam zich tot elkaar verhouden staat in de filosofie bekend als het mind-body probleem en het is een uitdaging voor iedere naturalistische theorie van het bewustzijn. Dit probleem wordt gerekend tot één van de klassieke vraagstukken in de filosofie en we kunnen een vroege formulering ervan terugvinden bij de zeventiende-eeuwse wiskundige en filosoof Descartes.

Met zijn mechanische filosofie probeerde Descartes een alternatief te bieden voor de Aristotelische oorzakenleer. Volgens deze opvatting moest de verklaring van een natuurkundig verschijnsel een antwoord geven op de vraag *waarom* iets gebeurde. Men wilde weten met welk doel een appel van de boom viel. Volgens Descartes was het onzin om je dit soort zaken af te vragen. Een verklaring kon volstaan met een antwoord op de vraag *hoe* de appel van de boom was gevallen. Het ging dus om het mechanisme dat de natuurverschijnselen veroorzaakte en niet om een of ander mysterieus doel dat achter de gebeurtenissen verscholen zat (Descartes, blz.202). Op deze manier kon de wetenschap volgens Descartes de gehele natuur verklaren, van de omwenteling van de planeten tot en met het gedrag van dieren. Alleen de menselijke geest onttrekt zich aan deze vorm van verklaren. Descartes hield vol dat de geest op een kwalitatieve manier verschilt van de materie. Waar de materie bestaat uit uitgebreidheid daar bestaat de geest uit denken (Descartes, blz.208). Onze gedachten, emoties en verlangens zijn dus opgebouwd uit een andere substantie dan de rest van ons lichaam en onze geest valt niet samen met ons materiele brein. Doordat Descartes de wereld op deze manier in tweeën had

opgedeeld creëerde hij echter een groot probleem voor zichzelf. Hij moest nu uitleggen hoe interactie tussen beide werelden mogelijk was. Hoe kan onze geest ons lichaam in beweging brengen als de geest zelf geen uitgebreidheid heeft? Op welke manier kan er sprake zijn van een causale relatie tussen lichaam en geest? Descartes en zijn volgelingen hebben verschillende antwoorden voor dit probleem aangedragen maar niemand heeft het overtuigend kunnen oplossen. Vandaar dat de theorie van Descartes, die tegenwoordig bekend staat als *dualisme*, door verreweg de meeste hedendaagse denkers wordt afgewezen. Ook Dennett staat er afwijzend tegenover. Het dualisme poneert immers een mysterieuze substantie die zich aan iedere wetenschappelijke verklaring onttrekt. De vraag wat het bewustzijn precies is wordt op die manier onbeantwoordbaar en een naturalistische theorie lijkt voorgoed buiten ons bereik te liggen (CE, blz.37).

Het Cartesiaanse idee van een immateriële geest die ergens schuil gaat in het lichaam als een geest in de machine heeft veel reacties los gemaakt. In de jaren veertig van de vorige eeuw hebben Ludwig Wittgenstein en Gilbert Ryle, ieder op eigen manier, een bijdrage geleverd aan de ontmanteling van dit idee. Hun alternatief wordt doorgaans aangeduid als filosofisch behaviorisme. Deze visie is verwant aan het behaviorisme dat omstreeks dezelfde tijd populair was in de psychologie maar het is niet hetzelfde (Churchland, blz.25). Wittgenstein en Ryle verleggen hun blik naar de taal. Ze proberen niet de vraag te beantwoorden hoe lichaam en geest zich tot elkaar verhouden maar willen door middel van een analyse van onze uitspraken over het bewustzijn aantonen dat het mind-body probleem een schijnprobleem is. Volgens Ryle zijn alle uitspraken die we doen over iemands mentale toestand te herleiden tot uitspraken over het gedrag dat we van deze persoon zouden kunnen verwachten. Als we van iemand bijvoorbeeld zeggen dat hij de wens heeft om op vakantie te gaan dan bedoelen we daar eigenlijk mee dat iemand, wanneer hij daartoe in de gelegenheid zou zijn, het gedrag zou vertonen dat toebehoort aan iemand die op vakantie wil gaan. Er is volgens Ryle dan ook geen sprake van een afzonderlijke wil, materieel of immaterieel, die ergens in het lichaam schuil gaat. De wil om op vakantie te gaan is niets anders dan de dispositie bezitten om bepaald verdrag te vertonen (Churchland, blz.23). Volgens Wittgenstein is onze manier van spreken wat dit betreft misleidend. Omdat we gewend zijn iemand bepaalde bewustzijnstoestanden toe te dichten en over deze toestanden te spreken alsof het zelfstandige entiteiten zijn die deze persoon bezit, kunnen we gemakkelijk gaan geloven dat deze bewustzijnstoestanden objecten zijn die we in de geest kunnen aantreffen. Wittgenstein spreekt daarom ook over het bewustzijn als een grammaticale fictie (Wittgenstein, blz.169). Omdat het mind-body probleem dus louter een grammaticaal probleem is kan het niet opgelost worden door wetenschappelijk onderzoek naar de status van het bewustzijn. Het kan alleen verdwijnen door een nauwkeurige filosofische analyse van het taalgebruik. Voor de filosofie van Dennett zijn Ryle en Wittgenstein van groot belang. Hij verslindt hun boeken tijdens zijn studie. Over Ryle, bij wie hij later zal promoveren, schrijft hij "he was obviously and refreshingly right about something deep" (AE, blz.23). Toch zal hij later vaak ontkennen dat hij zelf een behaviorist is (DB, blz.210). Verderop zullen we nog zien waarom.

Hoewel het dualisme dus op brede afkeuring mag rekenen zijn er toch zaken die in het voordeel ervan pleiten. Het is namelijk een heel natuurlijke manier van naar de wereld kijken. De scheiding tussen lichaam en geest komt overeen met enkele van onze diepste intuïties over hoe de wereld in elkaar zit. De ontwikkelingspsycholoog Paul Bloom heeft ontdekt dat zelfs zeer jonge kinderen een onderscheid maken tussen levende en levenloze dingen. Van een mens verwachten ze dat hij bepaalde motieven heeft voor zijn handelen. Het toekennen van motieven aan 'onbezielde' materie gebeurt daarentegen uitsluitend in metaforische zin (Bloom, 2005). Zoals Descartes aantoonde heeft het geen zin om te vragen waarom een appel van een boom valt. Het heeft echter wel zin om te vragen waarom een persoon bepaalde dingen doet en andere laat. Mensen kunnen redenen aanvoeren voor hun gedrag terwijl alledaagse objecten dat niet kunnen.



## **b. Intentionaliteit: een test-case voor het behaviorisme**

Het belangrijke verschil tussen een mens en een object als een appel is dat een mens beschikt over een mentale wereld die bestaat uit gedachten, verlangens en emoties. En dit zijn inderdaad dingen van een andere categorie dan de fysieke materie om ons heen die we kunnen beetgrijpen. In het filosofisch jargon wordt het verschil tussen mentale en fysieke fenomenen vaak aangeduid in termen van intentionaliteit. De stelling dat mentale fenomenen verschillen van fysieke omdat de eerste over intentionaliteit beschikken is afkomstig van de negentiende-eeuwse filosoof en psycholoog Franz Brentano, één van de belangrijkste grondleggers van de fenomenologie. Volgens hem verwijzen intentionele fenomenen zoals het verlangen naar een sappige appel naar iets dat buiten hen zelf ligt. Daarin verschillen ze van de appel zelf die helemaal nergens naar verwijst. Als we er echter van uitgaan dat dit een verschil is tussen objecten in de wereld dan zou Brentano ons op deze manier rechtstreeks terug voeren naar het Cartesiaans dualisme met zijn merkwaardige ontologische categorie van een denkende substantie. We hebben al gezien dat het poneren hiervan de doodsteek is voor het naturalisme.

Om deze reden vatten de meeste filosofen die zich met intentionaliteit bezig houden de stelling van Brentano op als een stelling over de taal en niet over de werkelijkheid (CC, blz.23). Ze hoeven dus niet te erkennen dat er twee verschillende substanties in de wereld zijn waarvan de één intentionaliteit bezit en de ander niet. In plaats daarvan stellen ze dat we op twee verschillende manieren over de wereld kunnen praten. Er is een volstrekt logische taal waarin de termen die we gebruiken op een eenduidige manier verwijzen naar objecten in de wereld. Dit is de taal waar de wetenschap idealiter gebruik van maakt. Deze taal bestaat slechts uit objectieve uitspraken vanuit de derde persoon. Wanneer we echter psychologische categorieën beschrijven zoals denken, verlangen etc. maken we gebruik van een andere taal. In deze taal doen we uitspraken over intentionele fenomenen.

Volgens Brentano is het niet mogelijk om uitspraken die gedaan worden in de ene taal te herleiden tot de andere taal. Dit betekent dat we geen wetenschappelijke uitspraken kunnen doen over intentionaliteit. Onze psychologische categorieën laten zich niet vertalen naar uitspraken over de fysieke werkelijkheid. Hierdoor is een naturalistische verklaring van het bewustzijn onmogelijk. De stelling van Brentano is dan ook één van de grootste uitdagingen voor het naturalisme. Deze stelling weerleggen kan op twee manieren: De eerste is aantonen dat intentionele uitspraken volledig vertaalbaar zijn in het idioom van de wetenschap, of, als dat niet mogelijk zou blijken te zijn, laten zien dat intentionele uitspraken, voor zover ze onvertaalbaar zijn, de wereld op een verstoorde en onbetrouwbare manier weergeven en dat we beter af zijn zonder deze misleidende vorm van spreken.

Het filosofisch behaviorisme is van mening dat alle intentionele uitspraken te vertalen zijn in uitspraken over gedrag. In de praktijk blijkt dit echter moeilijker dan gedacht. Om bijvoorbeeld een leerproces te omschrijven in termen van gedrag valt de behaviorist terug op de verhouding tussen stimulus en respons. Bij experimenten met proefdieren lukt het op die manier om een beschrijving te geven van het leerproces zonder daarbij te verwijzen naar een veronderstelde bewustzijnsinhoud. Na een positieve stimulus *x* zien we dat het proefdier steeds meer geneigd is tot gedrag *y*. Zo kunnen we dus het leerproces beschrijven maar, en hier zit volgens Dennett de crux, niet *wat* er nu precies geleerd is. "It is certainly not the case that what the rat in the maze learns is a sequence of skeleton motions (...). Nor does it learn to move through a series of spatial juxtapositions no matter what. What it learns, of course, is *where the food is*, but how is this to be characterized non-intentionally?" (CC, blz.36 *Cursiveringen* zijn van Dennett zelf). Volgens Dennett zijn alle pogingen om hiervoor een vertaling te geven in termen van stimulus en respons op niets uitgelopen. Dit is dan ook de reden dat we volgens hem beter op zoek kunnen naar een nieuwe manier om mentale uitspraken te vertalen in een naturalistische taal.

### c. Kunnen machines denken?

*The original question, 'Can machines think?' I believe to be too meaningless to deserve discussion. Nevertheless I believe that at the end of the century the use of words and general educated opinion will have altered so much that one will be able to speak of machines thinking without expecting to be contradicted.*  
Alan Turing (1950, blz. 49)

In zijn proefschrift betoogt Dennett dat wanneer we spreken over computers we een intentionele taal gebruiken die volledig te herleiden is tot de fysieke toestand waarin de computer op dat moment verkeert. Als we zeggen dat onze pc het antwoord ergens op 'weet' dan bedoelen we dat de informatie gecodeerd is opgeslagen op de harde schijf. Op deze manier kunnen we een vertaling geven van de 'mentale' toestand waarin het apparaat zich op dat moment bevindt in een zuiver wetenschappelijke taal (CC, blz.43). Deze manier van spreken over computers levert echter nogal wat weerstand op bij de meeste mensen. We zijn geneigd om te zeggen dat computers niet werkelijk beschikken over mentale toestanden en dat het toekennen hiervan niet meer is dan een vorm van projectie. Mensen daarentegen zijn behept met een authentieke intentionaliteit. Als we van een mens zeggen dat hij denkt is dit niet zomaar een projectie maar een adequate beschrijving van wat er in zijn hoofd omgaat.

Volgens het functionalisme is dit bezwaar echter ongeldig. Functionalisten gaan ervan uit dat een mentale toestand volledig gedefinieerd wordt door de functionele rol die deze toestand vervuld binnen het systeem waarvan het deel uitmaakt. Als we van een computer zeggen dat hij bijvoorbeeld 'weet dat een appel over een klokhuis bezit' dan bedoelen we dat deze informatie beschikbaar is voor verdere verwerkende processen. In deze zin verschilt dit niet met de kennis die een mens heeft over appels. Ook deze kennis kunnen we beschrijven louter aan de hand van de functionele relatie die de overtuiging vervult binnen een systeem van opvattingen. Het maakt voor de functionalist dus niet uit hoe de overtuiging fysiek gerealiseerd wordt. Of hij opgeslagen is in binaire code in de hardware van een computer of gerepresenteerd wordt door een samenspel van neuronen in de hersenen, in beide gevallen gaat het om een overtuiging die onderdeel is van een denkend systeem (Churchland, blz.36).

Mentale processen zoals denken kunnen dus onderzocht worden los van hun fysieke realisatie, dit betekend dat we zonder probleem kunnen spreken over de intentionele eigenschappen van machines. Dit betekent op zijn beurt weer dat we met de computer een geldig voorbeeld hebben van een geslaagde vertaling van een psychologisch naar een fysisch idioom. We mogen dus concluderen dat er niet zoets is als een principiële onvertaalbaarheid die ieder uitzicht op een naturalistische verklaring van het bewustzijn te niet zou doen, zoals Brentano dacht (CC, blz.43). Er zijn echter evidente verschillen tussen de intentionaliteit zoals we die aan een computer toekennen en onze eigen verlangens, emoties en ideeën. "Computers, if they are intentional, are only intentional in virtue of the intentionality of their creators. People and animals, however, are not designed and manufactured the way computers and their programmes are, nor are they essentially in the service of interpreting, Intentional beings. (One could turn the argument around; then it becomes a rather top-heavy argument for the existence of an Intentional God – none of your theistic abstract Gods – whom we are designed to serve.) If we are to avoid the God hypothesis, we must look elsewhere for a source of Intentionality in living systems; we must find something else to endow their internal states with content. (CC, blz.44)" De intentionaliteit van een computer is een afgeleide van de intentionaliteit van de maker. Het feit dat we bedoelingen aan een computer kunnen toeschrijven komt omdat de computer ontworpen is door een intelligente technicus die hem geprogrammeerd heeft om bedoelingen te hebben. Hoe zit dit met onze eigen intentionaliteit? Als we de computer analogie serieus nemen, en Dennett is

zeker van mening dat we dit moeten doen, dan zou de conclusie gerechtvaardigd zijn dat onze intentionaliteit niet van onszelf afkomstig is maar het resultaat van een ontwerpproces. We hebben echter al eerder gezien dat dit niet hoeft in te houden dat we ontworpen zijn door een intelligente schepper. Darwin's evolutionair algoritme waar we in deel één kennis mee hebben gemaakt kan prima verklaren hoe onze intentionaliteit tot stand is gekomen. Al onze intentionaliteit is slechts een afgeleide is van het proces van natuurlijke selectie. Voor Dennett is er in dit opzicht hooguit een gradueel verschil tussen de intentionaliteit van een computer, die een afgeleide is van de bedoelingen van de maker, en onze intentionaliteit die een afgeleide is van de bedoelingen van moeder natuur.

We hebben nu gezien waarom het behaviorisme er niet in is geslaagd om een volledige naturalistische theorie over het bewustzijn op te stellen. De dogma's van de behaviorist zitten hem in de weg bij het opstellen van een afdoende verklaring van intentionele fenomenen. Dennett hoopt echter dat de geest alsnog uit de machine gejaagd kan worden met behulp van een theorie die hijzelf teleofunctionalisme noemt (CE, Blz. 460). Dit is het functionalisme zoals we dit hierboven besproken hebben. Sterk leunend op Darwin's evolutietheorie probeert hij het bewustzijn te begrijpen als een natuurlijk artefact. Onze intentionaliteit is op dezelfde manier afgeleid van ons ontwerpproces als dit bij culturele artefacten het geval is. Nu we dit begrepen hebben kunnen we onze aandacht richten op Dennett's theorie van het bewustzijn zoals hij deze in *Consciousness Explained* uiteen heeft gezet.

### **III. Het bewustzijn verklaard**

In *Consciousness explained* wil Dennett laten zien dat ons bewustzijn een hijskraan is en geen luchthaak. We hebben echter allemaal bepaalde diep gewortelde intuïties met betrekking tot de werking van onze eigen geest die het ontwikkelen van een naturalistische theorie moeilijk maken. Als we spreken over ons bewustzijn maken we gebruik van een aantal misleidende metaforen die een dergelijke verklaring in de weg zitten. Dennett wil deze uit de weg ruimen en er zijn eigen metaforen voor in de plaats stellen (CE, Blz. 455). Zijn betoog staat dus op twee benen: enerzijds het ontmantelen van de traditionele manier van denken over het bewustzijn en anderzijds een nieuwe manier van denken ontwikkelen. In mijn bespreking hieronder zullen beide motieven af en toe door elkaar lopen. Dennett ontwikkeld zijn idee van het bewustzijn als hijskraan vaak in tegenspraak met de traditionele ideeën die op de één of andere manier schatplichtig blijven aan het idee van het bewustzijn als luchthaak.

#### **a. Heterofenomenologie**

*Talking about the mind, for many people, is rather like talking about sex: slightly embarrassing, undignified, maybe even disreputable. "Of course it exists," some might say, "but do we have to talk about it?" Yes, we do.*

Daniel Dennett (IS, blz.1)

Wat gebeurt er precies als we denken? Wat gaat er in mijn hoofd om wanneer ik me bedenken dat ik trek begin te krijgen in een appel? Of op het moment dat ik naar muziek luister? En nu, terwijl ik deze woorden typ? Al deze processen maken onderdeel uit van mijn bewustzijn. Hoe moeten we dit bewustzijn verklaren? Voor Dennett is het raadselachtige karakter ervan geen reden om terug te deinzen. Ons bewustzijn is een fantastisch fenomeen. Het is een bron van plezier, angst, lust en verwondering. Maar het is niet minder dan alle andere zaken die we in de wereld aantreffen een onderdeel van de natuur. Dit betekent dat het verklaarbaar moet zijn op dezelfde manier als alle andere natuurverschijnselen. Er zijn echter mensen die een uitzonderingspositie claimen voor het bewustzijn. Zoals we gezien hebben deed Descartes dit al in de zeventiende eeuw en ook vandaag de dag zijn er filosofen die zo onder de indruk zijn van het verschil tussen

onze geest en de rest van onze natuur dat ze onmogelijk kunnen geloven dat het bewustzijn ooit door de wetenschap verklaard kan worden.

Onze gedachten zijn afgesloten voor de rest van de wereld en alleen voor onszelf intuïtief toegankelijk. Wellicht dat wetenschappers er ooit in zullen slagen om via hersenscans vast te stellen waar iemand aan denkt. Door middel van het meten van de hersenactiviteiten van een proefpersoon zouden ze dan na moeten gaan of hij zich inderdaad een paarse koe inbeeldt zoals hij beweert. Maar ook in dit hypothetische scenario kunnen de wetenschappers de geprivilegieerde toegang tot het bewustzijn van hun proefpersoon niet evenaren. Vaststellen dat iemand aan een paarse koe denkt is immers altijd nog iets anders dan zelf aan een paarse koe denken. Volgens sommige filosofen is het om deze reden onmogelijk om een naturalistische verklaring te geven van het bewustzijn. Ieder wetenschappelijk onderzoek moet het noodzakelijk stellen zonder het eerste persoonsperspectief dat een bewuste ervaring maakt tot wat hij is. Een wetenschappelijke omschrijving van een bewuste ervaring blijft altijd slechts dat: de beschrijving van een ervaring. Dit kan nooit in de plaats komen van de ervaring zelf.

Volgens Dennett is het maar de vraag hoe erg dit is. Een wetenschappelijke beschrijving van de elektromagnetische straling die we in het dagelijks leven 'licht' noemen zal nooit meer zijn dan enkele woorden en formules op papier. De beschrijving zelf geeft geen licht en zal dit ook nooit doen. Net als in het geval van het bewustzijn is er dus sprake van een fundamentele tegenstelling tussen hetgeen we beschrijven en de beschrijving zelf. Maar dit is geen reden om te veronderstellen dat het fenomeen licht voor altijd buiten het bereik van de wetenschap zal blijven (CE, blz.71). Tegenstanders van het naturalisme zullen hier echter niet door overtuigd zijn. Als we een natuurkundig fenomeen als licht onderzoeken beschikken we immers over objectieve data. Er zijn waarneembare verschijnselen in de wereld waar iedere onderzoeker toegang toe heeft. Dit is essentieel. Iedere wetenschappelijke beschrijving van een bepaald fenomeen gebeurt vanuit het derde persoonsperspectief. De data waar een theorie over het bewustzijn gebruik van moet maken zijn echter alleen toegankelijk vanuit de subjectieve ervaring en dit maakt ze ongeschikt als input voor een objectieve wetenschappelijke theorie.

Het is de uitdaging voor Dennett om aan te tonen dat we een theorie kunnen samen stellen die het bewustzijn verklaart met uitsluitend de voorhanden zijnde data (CE, blz.71). De enige data die hij tot zijn beschikking heeft zijn de getuigenissen van proefpersonen over hun mentale toestand. De methode om deze data te verkrijgen noemt hij *heterofenomenologie*. Dennett accepteert de beschrijvingen die zijn proefpersonen van hun ervaringen geven als data die verklaard moet worden. Hij wijst echter de suggestie van de hand dat de geprivilegieerde toegang die proefpersonen hebben tot hun eigen bewustzijn onfeilbaarheid garandeert (CE, blz.68). Dit is een methodologisch uitgangspunt met verregaande consequenties voor de vraag of Dennett er uiteindelijk in zal slagen om het bewustzijn te verklaren. Vandaar dat we er hier even bij stil blijven staan.

Dennett vergelijkt de heterofenomenoloog die de getuigenissen van proefpersonen over hun bewuste ervaringen optekent met een antropoloog die de religie van een inheemse stam in beeld probeert te krijgen (CE, blz.82). Voor de antropoloog zijn de religieuze verhalen die de gelovigen aan hem vertellen zijn belangrijkste data. Als iemand aangeeft dat zijn God vier armen heeft en vleugels dan aanvaardt hij dit als nuttige informatie. Als ambitieus wetenschapper zal hij proberen dit godsbeeld van een verklaring te voorzien. Hij probeert het geloof in deze vierarmige, gevleugelde inheemse heilige te onderzoeken, maar dit betekent niet dat hij het bestaan ervan aanvaardt! Hij is in staat om wetenschappelijke uitspraken te doen op basis van de hem beschikbare data. Zo zal hij iedereen die beweert dat de God van deze mensen vijf armen heeft tegenspreken (CE, blz.79). Maar niets van dit alles is een argument voor het bestaan van een dergelijke God. De heterofenomenoloog bevindt zich in een vergelijkbare positie. Hij aanvaardt de

getuigenissen van zijn proefpersonen als nuttige informatie en onderzoeksdata maar dat betekend niet dat hij er zonder slag of stoot vanuit gaat dat de subjectieve ervaringen waarover ze rapporteren zich daadwerkelijk op de manier voordoen zoals zij ze beleven.

Hoe ons bewustzijn werkt zou immers wel eens heel anders kunnen zijn dan wij vanuit onze eerste persoonsperspectief denken dat het werkt. Een goede theorie over het bewustzijn moet dus het feit dat we beweringen kunnen doen over ons bewustzijn verklaren. Hij moet eveneens verklaren waarom we bepaalde zaken op deze manier ervaren en niet op een andere manier. Maar het is een open vraag of de manier waarop we denken dat onze ervaringen tot stand komen in overeenstemming is met de werkelijkheid. We hebben slechts in beperkte mate toegang tot onze eigen psychologie. Het is ons vaak niet duidelijk waar onze ervaringen eigenlijk uit zijn opgebouwd. Als we hier naar gevraagd worden dan gaan we ongemerkt theoretiseren. We rapporteren niet meer wat we daadwerkelijk waarnemen maar wat ons logisch lijkt dat er zou moeten gebeuren (CE, blz.68).

### **b. Het Cartesiaans theater versus Dennett's meervoudige versies**

Toen we het dualisme bespraken zagen we al dat Descartes en zijn volgelingen verstrikt raakten in het probleem hoe de interactie tussen lichaam en geest gestalte moest krijgen. Om dit probleem op te lossen heeft Descartes voorgesteld dat de geest zitting heeft in een bepaald gedeelte van de hersenen, de pijnappelklier, van waaruit hij opdrachten aan het lichaam verstrekt. Deze ideeën van Descartes worden tegenwoordig niet meer serieus genomen. Toch blijft de veronderstelling dat er ergens in ons hoofd een commandocentrum moet zijn voor veel mensen een heel verleidelijke gedachte. We stellen ons voor dat al onze handelingen het gevolg zijn van opdrachten die ergens in ons brein worden geformuleerd. Een terechte vraag wordt dan: geformuleerd door wie? Het idee van een commandocentrum impliceert een gezagvoerder. Een intelligente mini-ik die besluiten neemt en orders verstrekt. Hoewel niemand gelooft dat een dergelijke homunculus zich ergens in het brein schuil houdt, blijkt het moeilijk om over het bewustzijn te spreken zonder een intelligente opdrachtverstrekker te veronderstellen.

Alle waarnemingen die we doen komen binnen als informatie van onze zintuigen en ergens in de hersenen wordt deze informatie verwerkt zodat we ons handelen kunnen afstemmen op hetgeen er in de wereld gebeurt. Het is zeer verleidelijk om te veronderstellen dat er een centrale plek moet zijn waar alle binnenkomende informatie geëvalueerd wordt zodat er een beslissing genomen kan worden over wat de beste manier is om te reageren. Deze plek zouden we het Cartesiaans theater kunnen noemen (CE, blz.107). Alle informatie komt hier samen en wordt op een bepaalde manier gerepresenteerd zodat het publiek in het theater op basis van deze informatie opdrachten kan verstrekken aan het lichaam. Weinig onderzoekers zijn echter van mening dat het proces in de hersenen inderdaad op deze manier verloopt. Praktisch iedereen zal ontkennen dat er zoiets is als een Cartesiaans theater waarin er beslissingen worden genomen op basis van een soort film die door onze zintuigen wordt samengesteld. Desondanks blijft bijna iedereen vast houden aan het idee dat er ergens in het brein een grens is tussen input en output. Als ons gedrag een samenspel is van input, bewustzijn toestanden en output, dan moet we kunnen vaststellen waar precies de informatie van onze zintuigen ophoud slechts informatie te zijn en een bewuste ervaring wordt. Is het niet een logische noodzakelijkheid dat we ergens een dergelijke grens kunnen trekken? (CE, blz.108)

Volgens Dennett niet. Hij wil een alternatief presenteren voor het Cartesiaans theater dat hij het 'meervoudige versies' model noemt. Eerst zal ik hier uitleggen wat dit model precies inhoud en waarin het verschilt van het cartesiaans theater, daarna zal ik laten zien waarom het volgens Dennett een betere manier is om over het bewustzijn na te denken dan de traditionele manier.

Voorop staat voor Dennett dat er geen punt in het brein is waar alles samenkomt. Het brein is de controlekamer van het lichaam maar in het brein zelf is er niet zoiets als een commandocentrum van waaruit alle beslissingen genomen worden (CE, blz.106). De informatie die we middels de zintuigen verkrijgen wordt verwerkt in verschillende parallelle sporen. Op basis van deze input worden op lokaal niveau oordelen geveld met betrekking tot de inhoud (CE, blz.111). Maar zodra zo'n oordeel geveld is hoeft het niet ergens anders naar verstuurd te worden voor verdere inspectie. Ieder oordeel hoeft maar één keer gemaakt te maken (CE, blz.113). Dennett vervangt de metafoor van het theater voor die van het schrift. Er worden voortdurend verschillende verhaallijnen uitgewerkt door het brein. Er bestaan dus meerdere versies die voortdurend onderhevig zijn aan redactionele herzieningen. Het is echter niet zo dat één van deze verhalen de canonieke versie is. Zo'n versie, die ons precies zou kunnen vertellen hoe een persoon de werkelijkheid ervaren heeft, bestaat niet (CE, blz.113).

Deze opvatting breekt met verschillende intuïties die we hebben met betrekking tot ons eigen bewustzijn. Wellicht dat het daarom in eerste instantie overkomt als een nogal merkwaardige theorie. We krijgen echter een beter beeld van wat Dennett bedoeld als we inzoomen op de meest karakteristieke eigenschap van het meervoudige versies model. Volgens Dennett kunnen we in het dagelijks leven een uitspraak doen als:

(1) x werd zich bewust van y op tijdstip T1.

Zodra we echter ons vertrouwde perspectief verlaten en inzoomen op het microniveau waarop onze hersenen werken wordt (1) zinloos (CE, blz.102). Als er een Cartesiaans theater in ons brein aanwezig was dan zou het gemakkelijk zijn om vast te stellen wanneer iemand zich bewust werd van een bepaalde stimulus. Dit zou het moment zijn waarop de informatie beschikbaar werd in het theater. Maar omdat een dergelijke centrale plek niet bestaat is (1) onbeantwoordbaar. Er zijn domweg geen feiten die ons ervan zouden kunnen overtuigen dat x zich bewust werd van y op tijdstip T2 in plaats van T1 als het verschil tussen T1 en T2 maar klein genoeg is (CE, blz.136). Dit zou alleen kunnen als er ergens een grens was tussen bewust en onbewust en een dergelijke grens is er niet (CE, blz.126).

Het grote voordeel van het meervoudige versies model is dat het Dennett de mogelijkheid geeft om enkele belangrijke uitdagingen voor een naturalistische opvatting van het bewustzijn het hoofd te bieden. We zullen er hier één bespreken. Een onderzoek dat gedaan is door Benjamin Libet in de jaren zestig van de vorige eeuw naar de manier waarop het brein tijdsvolgorde representeert. Als we de resultaten hiervan duiden vanuit de opvatting dat onze geest alle informatie voor zich gepresenteerd krijgt als in een Cartesiaans theater komen we hopeloos in de knoei. In dat geval lijkt een naturalistische opvatting van het bewustzijn onmogelijk. Als we de resultaten echter opnieuw interpreteren met Dennett's theorie in ons achterhoofd dan opent zich een geheel ander perspectief.

Libet is er achter gekomen dat in bepaalde omstandigheden proefpersonen een verkeerd oordeel maken over de volgorde van twee opeenvolgende stimuli. Terwijl A aan B vooraf ging ervaren ze dit andersom. De details van het experiment zijn hier niet belangrijk. De conclusie die Libet hier uit trok was dat de ervaring van A blijkbaar ergens in het brein opgehouden wordt om pas na stimulus B aan de geest gepresenteerd te worden. Libet blijft dus schatplichtig aan het idee dat er ergens in het brein een punt is waar onze ervaringen samenkomen en bewust ervaren worden. Volgorde van aankomst op dit punt is de volgorde waarin we ons van A en B bewust worden. Hoewel stimulus A eerder werd toegediend dan B, was B eerder op het podium in het Cartesiaans theater waardoor de proefpersonen tot het oordeel kwamen dat B aan A vooraf ging.

Als heterofenomenoloog concentreert Dennett zich op de verklaringen van de proefpersonen. Deze geven aan dat ze eerst B hebben ervaren en toen pas A. Vanuit het

meervoudige versie model is dit onproblematisch. Het brein heeft de foutieve conclusie getrokken dat B aan A vooraf ging en dus is dit wat er wordt medegedeeld. Op het moment dat deze conclusie getrokken wordt is de kous verder af. Het oordeel hoeft niet herhaald te worden ten behoeve van een publiek ergens in het Cartesiaans theater. Libet komt in de problemen omdat hij een ervaring in het hoofd probeert te construeren op basis waarvan zijn proefpersonen tot hun conclusie gekomen zijn. Daar moet alles zich in real-time nogmaals afspelen. Voor Dennett daarentegen is de ervaring verklaard met de vaststelling dat het brein tot een oordeel kwam met betrekking tot de tijdsvolgorde. Het is onzin om te stellen dat een oordeel daarna nog een keer ervaren moet worden. Dankzij het meervoudige versies model hoeft Dennett alleen het oordeel over de tijdsvolgorde te verklaren terwijl Libet de tijdsvolgorde zelf moet zien te verklaren. Dit is het grote probleem met het idee van het Cartesiaans theater: het veronderstelt dat alle conclusies die door het brein gemaakt worden nog een keer in de bioscoopzaal afgedraaid moeten worden alvorens we ze daadwerkelijk ervaren (CE, blz.144-170).

Het Cartesiaans theater is een probleem voor Dennett omdat het een overtreding is van de spelregels die we aan het begin opstelden. Het gaat uit van een autoriteit ergens in het brein maar door deze autoriteit te postuleren wordt er niet werkelijk iets verklaard. Het probleem wordt alleen maar verschoven. Eerst moesten we uitleggen hoe het komt dat wijzelf bewuste ervaringen hebben, nu moeten we uitleggen hoe het komt dat het publiek in het Cartesiaans theater een bewuste ervaring heeft. De theorie postuleert een vorm van intelligentie, een homunculus, waarvan het bestaan door de theorie niet verklaard kan worden. Dennett is geen tegenstander van het beschrijven van bepaalde functies binnen het brein als werden ze uitgevoerd door homunculi, maar hij waarschuwt ervoor om deze te intelligent te maken (CE, blz.14). Keer op keer hamert hij op het feit dat er geen intelligente autoriteit in het brein aanwezig is. We zien dit terug bij zijn bespreking van het spraakvermogen (CE, blz.227-252) en de architectuur van het brein (CE, blz.253-282). We zullen daar verder niet op ingaan. In alle gevallen zijn het niet meer dan luchthaken die het werk moeten doen dat in werkelijkheid door hijskranen wordt uitgevoerd. We zullen nu onze blik naar deze hijskranen verplaatsen.

### **c. Evolutie van het bewustzijn**

In het eerste deel hebben we gezien dat het Dennett's ambitie is om een naturalistische theorie van het bewustzijn samen te stellen. Hij wil laten zien dat het bewustzijn van de grond af opgebouwd kan worden door louter mechanische processen. In deel twee hebben we gezien dat eerdere pogingen hiertoe gestrand waren. Na de twee voorgaande paragrafen begrijpen we ook waarom dit zo is. Men ging altijd uit van een verkeerd idee over wat het bewustzijn precies was en hoe we het moesten onderzoeken. Dennett heeft daarom een eigen methode en een eigen theorie geïntroduceerd die we nu eveneens besproken hebben. Dit is het moment om de schuld uit het eerste deel in te lossen en te laten zien hoe het bewustzijn tot stand is gekomen. Welke hijskranen zijn er verantwoordelijk voor dit fenomenale stukje ontwerp van de natuur? En waarom stelt Dennett dat het bewustzijn zelf één van de meest bewonderenswaardige hijskranen is die door de evolutie zijn voortgebracht?

De eerste hijskraan die Dennett bespreekt zijn we al eerder tegengekomen: het proces van natuurlijke selectie. De eerste micro organismen die de zeeën bevolkten vele miljoenen jaren geleden waren zich niet bewust van hun eigen lot en alleen in overdrachtelijke zin kunnen we nu vast stellen dat ze bepaalde belangen hadden. Vanuit ons standpunt is het duidelijk dat het in hun belang was om zich zoveel mogelijk voort te planten, om uit de buurt van gevaar te blijven en efficiënt om te gaan met beschikbare hulpmiddelen. We kunnen er echter zeker van zijn dat ze om al deze dingen zelf niets gaven. Om de eenvoudige reden dat ze nog niet geprogrammeerd waren om ergens om te geven. Door natuurlijke selectie werden echter de organismen bevoordeeld die reacties vertoonden waarvan we achteraf kunnen zeggen dat ze in overeenstemming waren met hun belang. Voor het eerst ontstond er zoiets als een rudimentaire betekenis.

Sommige impulsen kregen de betekenis 'slecht voor me, weg wezen' andere de betekenis 'goed voor me, aan tafel' (CE, blz.176).

Voor de natuur is leren een kwestie van do or die. De evolutie is chronisch bijziend. De enige manier om vooruit te komen is door selectie op basis van de toevallig aanwezige differentiatie in eigenschappen. Er is nooit sprake van voorkennis. In dit tekort is echter voorzien door de tweede hijskraan die Dennet bespreekt: fenotypische plasticiteit. (CE, blz.184) Wat moedernatuur niet kan dat kunnen sommige van haar kinderen inmiddels wel: vooruit denken alvorens te handelen. Dit vermogen is tot stand gekomen dankzij natuurlijke selectie en heeft op zijn beurt het proces van natuurlijke selectie in een stroomversnelling gebracht. Een organisme waarvan alle reacties bij de geboorte vast liggen kan alleen maar hopen dat zijn bedrading op een dusdanige manier is aangelegd dat hij alle gevaren uit zijn omgeving aan kan. Feitelijk kan hij alleen die bedreigingen de baas die zijn voorgangers ook de baas konden. Voor nieuwe bedreigingen heeft de evolutie hem niet uitgerust. Een organisme bij wie de bedrading echter nog niet helemaal vast ligt kan tijdens zijn leven aanpassingen maken al naar gelang de situatie dat van hem vraagt. Zo kan hij zelf op een nieuwe originele manier op dreigingen reageren. (CE, blz.186) De mogelijkheid om tijdens je leven je eigen ontwerp verder te ontwikkelen is de basis voor de capaciteit om te leren. Deze ontwerpflexibiliteit is een belangrijke motor voor de evolutie. Kon er eerst maar één ontwerp per individu uitgeprobeerd worden, nu is dat een veelvoud hiervan. Het is nu ook mogelijk om van andere soortgenoten te leren. Ontwerp kan nu horizontaal doorgegeven worden in plaats van alleen verticaal.

Het ontstaan van organismen die in staat zijn om te leren betekent dat voor het eerst in de geschiedenis informatie belangrijk wordt. Daarmee zijn we ook getuige van de geboorte van informatiemanagement. Op het juiste moment moeten de juiste gegevens beschikbaar zijn. Specialisten vergaren voortdurend informatie over van alles en nog wat. Het probleem wat nu opgelost moet worden is hoe deze informatie op een overzichtelijke en nuttige manier centraal beschikbaar gesteld kan worden. Het is misschien verleidelijk om te veronderstellen dat er ergens in het brein een coördinator is die deze vacature heeft vervuld. Maar zo werkt de evolutie niet. Waar zou deze opeens vandaan moeten komen? Dennett stelt daarom een pandemonium model voor (CE, blz.189). Hierin vervullen tijdelijke coalities van specialisten de functie van coördinator ad interim. Zodra er behoefte is aan informatie omdat er bijvoorbeeld gevaar dreigt, grijpt een bepaalde coalitie de macht. Ze coördineren de reactie van het organisme tot dat de coalitie uit elkaar valt omdat het gevaar is geweken of omdat er een kleine paleisrevolutie heeft plaats gevonden en een ander coalitie de troon heeft beklommen.

Informatie overdracht tussen de specialisten binnen het organisme zelf is dus van belang. Maar ook informatie overdracht tussen soortgenoten gaat nu een rol spelen. Zodra er een rudimentaire taal ontstaat, worden deze mogelijkheden tot uitwisseling exponentieel vergroot. Het is denkbaar dat hier ook de oorsprong ligt van de gewoonte om tegen onszelf te praten. Zoals we andere om informatie vragen zo kunnen we ook onszelf om informatie vragen. Zo zou onze stream of consciousness weleens een antwoord kunnen zijn op het probleem: hoe krijgen we informatie van de ene specialist naar de ander? (CE, blz.196)

Dit alles is echter slechts een opmaat naar de derde hijskraan die Dennett beschrijft. Alleen bij mensen heeft het proces van informatieverwerking zo'n hoge vlucht genomen dat dit een eigen evolutionair proces opgang gebracht heeft: onze culturele evolutie. Ons bewustzijn wordt bevolkt door ideeën. Ideeën die ons gedrag sturen, onze overpeinzingen begeleiden en onze leefwereld domineren. Ideeën die ons moeten helpen te overleven, soortgenoten van het andere geslacht het hof te maken, en de wereld om ons heen naar onze behoeften in te richten. De samenstelling van onze ideeën is voortdurend aan verandering onderhevig. We zijn immers altijd bezig nieuwe ideeën te verwerven. In navolging van Richard Dawkins omschrijft Dennett deze ideeën als het basismateriaal voor een eigen evolutie geheel volgens Darwinistische principes (CE,



blz.199-208; Dawkins 1976, blz.189-201). Ook onze ideeën (Dawkins en Dennett spreken over memen, analoog aan genen) zijn onderhevig aan het principe van natuurlijke selectie. Alleen die ideeën die in staat zijn om zichzelf van geest naar geest te kopiëren hebben een kans om te overleven in een volgende generatie ideeëndragers.

Door het voortdurend uittesten van onze eigen grenzen middels zelfstimulatie zijn onze voorouders waarschijnlijk op verschillende nieuwe probleem oplossende strategieën gestuit die de interne communicatie van het brein op een radicale manier herzien hebben (CE, blz.209). Als memen hebben deze strategieën hun weg gevonden naar andere breinen. Een goed idee kopieert zich immers snel. Op deze manier zijn onze ideeën verantwoordelijk voor het vormen van onze geest. Dennett vergelijkt dit proces met de werking van een computer (CE, blz.211). Net als onze hersenen bezit een pc over een grote mate van plasticiteit. Afhankelijk van het programma dat we erop installeren kunnen we onze computer veranderen in een tekstverwerker of een schaakcomputer. Deze structuren bestaan uit regels en niet uit draden (rules rather than wires, CE, blz.211). Computerprogrammeurs spreken van een virtuele machine. Deze regels bepalen het gedrag van de computer maar genereren ook zijn 'bewustzijnstoestanden'. Ze verlenen de computer als het ware een virtuele identiteit. Het is nog steeds dezelfde machine met draadjes en elektronische componenten maar we kunnen hem nu ook begrijpen als een functionele entiteit die erop uit is om, bijvoorbeeld, mijn koningin te slaan bij de volgende zet. Iets vergelijkbaars doen de memen die onze geest bevolken. Volgens Dennett is ons bewustzijn niets anders dan "a huge complex of memes (or more exactly meme-effects in the brain) (CE, blz. 210)" die als een virtuele machine geïmplementeerd zijn in ons brein en de capaciteiten van ons brein vergroten zoals een computerprogramma de capaciteiten van een computer vergroot. Dit memencomplex vormt een hijskraan die in omvang en hefvermogen alle andere ver achter zich laat (CE, blz.221).

#### **d. De kleren hebben geen keizer<sup>1</sup>**

Er zijn veel filosofen voor wie het evident is dat de bovenstaande beschrijving van het bewustzijn onvolledig is. Daarbij beroepen ze zich op een krachtige intuïtie. Zoals we al zagen toen we het functionalisme bespraken staan veel mensen wantrouwig tegenover een mechanische verklaring van het bewustzijn. Ze hebben het gevoel dat een dergelijke verklaring het allerbelangrijkste over het hoofd ziet. Het zijn niet de processen in mijn hoofd die verklaard moeten worden maar het bewustzijn zelf. Het meervoudige versies model kan dan een hele fraaie metafoor zijn van hoe de informatie verwerking in mijn brein verloopt, het allerbelangrijkste wordt er niet door verklaard: dat ik me bewust ben van de uitkomst van dit informatie verwerkende proces. Dit is een hardnekkig bezwaar dat keer op keer opduikt. Wie het plausibel vindt klinken zit waarschijnlijk nog steeds gevangen in de veronderstelling dat er zich ergens in ons hoofd een Cartesiaans theater moet bevinden.

We zullen deze intuïtie eens nader bekijken. Wat gebeurt er volgens onze intuïtie als we tegen onszelf zeggen "ik heb honger"? In dat geval hebben we blijkbaar een leeg gevoel in onze maag dat allereerst door ons opgemerkt wordt en waarover we vervolgens deze uitspraak doen. We zouden de honger een eerste orde idee kunnen noemen en de uitspraak die betrekking heeft op dit idee een tweede orde idee (CE, blz.307). We zijn nu misschien geneigd om te veronderstellen dat een dergelijk tweede orde idee noodzakelijk is voor bewustzijn. Een ervaring op zich is niet voldoende, de ervaring moet opgemerkt worden. Dit opmerken gebeurt op zijn beurt ook weer onbewust, tenzij het begeleid wordt door een derde orde idee van het type: "ik denk dat ik honger heb." En zo kunnen we verder gaan met een vierde orde idee, een vijfde orde idee et cetera (CE, blz.308). We hebben dus onze ervaringen en onze opvattingen die betrekking hebben op deze

---

<sup>1</sup> De titel van dit gedeelte ontleen ik aan Bloom (2005), die hem op zijn beurt overgenomen heeft van Stewart Guthrie. Een mooi voorbeeld van een 'meme' in actie.

ervaringen. Dit lijkt een redelijke veronderstelling en het verwijt aan Dennett is dat hij dit onderscheid niet serieus lijkt te nemen in zijn theorie.

Dennett aanvaardt deze uitdaging en ontkent dat dit onderscheid werkelijk zo belangrijk is als zijn tegenstanders beweren. Hoewel hij toegeeft dat het soms mogelijk is om te spreken van een ervaring en een afzonderlijk idee over deze ervaring is dit lang niet altijd juist. In de gevallen waarin we inderdaad ons een idee vormen over een bepaalde ervaring kan zijn theorie dit verklaren door specialisten die het werk van elkaar in de gaten houden (CE, blz.316). In alle andere gevallen is het domweg niet zo dat de twee van elkaar te scheiden zijn. Mijn overtuiging is dat ik op dit moment ervaring x beleef zijn in dat geval niet twee separate mentale toestanden maar één en dezelfde. In het genoemde voorbeeld waarin we ons bedenken dat we honger hebben is het moeilijk om y van x los te zien. Als we beide werkelijk van elkaar konden onderscheiden dan zouden we zonder probleem een uitspraak kunnen doen als "ik denk dat ik honger heb maar misschien vergis ik me en heb ik alleen slaap" (CE, blz.319).

Het idee dat werkelijk bewustzijn iets anders is dan de informatie verwerkende processen die Dennett beschrijft is echter hardnekkig. Om aan te tonen dat hij met zijn theorie iets fundamenteels over het hoofd ziet zijn er verschillende voorbeelden gegeven van processen waarin alle informatie aanwezig was maar waar we niet kunnen spreken van bewuste ervaring. Een mooi voorbeeld hiervan is blindzicht. Mensen die aan deze aandoening lijden hebben een gedeelte van hun zicht verloren naar aanleiding van een hersenbeschadiging. Er bevindt zich een blinde vlek in hun visuele veld. Toch zijn ze in bepaalde omstandigheden in staat om ware uitspraken te doen over zaken die zich in deze blinde vlek afspelen. Bepaalde informatie wordt dus wel verwerkt maar leidt niet tot het vermogen om te kunnen zien. Het lijkt er dus op dat alle functionele eigenschappen van zicht aanwezig zijn zonder dat we van zien kunnen spreken (CE, blz.325).

Voor Dennett's tegenstander is het belangrijk om te kunnen aantonen dat er een principieel verschil is tussen het bewust kunnen waarnemen van de wereld en er alleen informatie over verkrijgen. Volgens Dennett is dit verschil echter niet principieel maar gradueel. Hij stelt zich voor hoe iemand met blindzicht zijn vermogen om voorspellingen te doen over wat er in zijn blinde vlek gebeurt verder zou kunnen ontwikkelen. Is het ondenkbaar dat als iemand deze capaciteit maar ver genoeg ontwikkelt, hij over een vermogen zou bezitten dat functioneel equivalent is aan zien en dat we daarom in alle relevante manieren ook zouden moeten beschouwen als een vorm van zien? (CE, blz.343)

Vele zullen nog steeds van mening zijn van niet. Volgens hen valt een ervaring niet samen met de functionele rol die de ervaring speelt in het netwerk van bewustzijnsinhouden, input en output maar is er iets specifiek dat een ervaring tot een ervaring maakt. Iedere ervaring zou beschikken over een kwalitatieve inhoud. Ze hebben hier zelfs een naam voor: Qualia. Met dit begrip wil men doorgaans het verschil aanduiden tussen de subjectieve ervaring en de elementen waaruit deze ervaring is opgebouwd. We kunnen dit illustreren met een voorbeeld. We weten dat bepaalde dingen ons rood toeschijnen omdat het oppervlak van deze voorwerpen de meeste lichtfrequenties absorbeert en het licht maar gedeeltelijk weerkaatst. We weten, kortom, dat de kleur rood bestaat uit elektromagnetische straling van een bepaalde frequentie. Maar zo ervaren we dit niet! We ervaren geen lichtfrequentie maar hebben een heel specifieke ervaring van een felle kleur die voor ons direct herkenbaar is als rood. Deze ervaring noemen we een quale.

Zijn qualia dit gehoopte extra's dat onze bewuste ervaring onderscheidt van de louter informatie verwerkende processen in ons brein? Dennett is natuurlijk van mening dat dit niet het geval is. Om het bestaan van een zo specifieke ervaring als die van de kleur rood te verklaren verwijst hij naar de functie van kleuren in het dagelijks leven. De dingen hebben geen kleuren van zichzelf, ze beschikken over eigenschappen die ervoor zorgen

dat ze bepaalde lichtfrequenties weerkaatsen en andere niet. Dat wij ze als gekleurd herkennen heeft er alles mee te maken dat het voor ons belangrijk was om ze te leren herkennen. De evolutie heeft die individuen bevooroordeeld die de fel rode kleur van een giftige bes duidelijk konden onderscheiden tegen een groene achtergrond. Tegelijkertijd waren ook die exemplaren van de bes in het voordeel die ons het meest rood toeschenen. Kleur en kleurwaarneming zijn dus gezamenlijk geëvolueerd en zeer fijn op elkaar afgesteld. Dit is volgens Dennett de reden dat kleur ons voorkomt als een absolute kwaliteit die niet verder analyseerbaar is. De evolutie heeft ons alert gemaakt op bepaalde vormen van informatie maar dit is geen reden om aan te nemen dat er nog iets onverklaarbaars is dat deel maakt van onze waarneming naast deze informatie (CE, blz.382).

Mensen die blijven volhouden dat ons bewustzijn niet kan bestaan uit de mechanische processen zoals Dennett deze beschrijft doen dit vaak omdat ze vrezen dat als hij gelijk heeft we niets meer zouden zijn dan voorgeprogrammeerde machines, we zouden onze specifieke menselijkheid op deze manier wel eens kunnen verliezen. Zoals we gezien hebben is het eerste volgens Dennett zonder meer het geval, maar de bijbehorende angst is volgens hem ongegrond. Ja, we zijn biologische machines zonder een immateriële geest of iets dat de functie van een dergelijke geest zou kunnen vervullen. Veel mensen zijn bang dat met het verlies van een dergelijke autoriteit die de oorsprong is van ons denken, spreken en handelen we ook onze identiteit kwijt raken. Als al mijn gedachten het resultaat zijn van een pandemonisch proces waar niemand toezicht op houdt, wie garandeert dan nog de eenheid ervan? Wie garandeert mijn ondeelbaarheid als individu?

Dennett sluit zijn boek af met de vraag wat dit ondeelbare 'ik' precies is dat we zo koesteren (CE, blz.412-430). Hij concludeert dat onze identiteit bestaat uit de vele verhalen die we vertellen over onszelf. Zoals een spin een web om zich heen produceert zo produceren wij verhalen. Het 'ik' bestaat uit die verhalen en in die verhalen als een narratief zwaartepunt. Het is dus niet de auteur van de verhalen maar het resultaat ervan. Het ik is een fictie maar een heel bijzondere vorm van fictie. Het heeft dezelfde functie als de kleren van de keizer: zolang we er in geloven staan we niet naakt.

#### **IV. Conclusie: Een gesprek met Otto**

In het eerste deel hebben we de filosofische implicaties van Darwin's evolutietheorie bekeken en we hebben gezien hoe we Dennett's streven om een naturalistische theorie van het bewustzijn te ontwikkelen konden formuleren in behulpzame termen (hijskranen in plaats van luchthaken). In deel twee hebben we gezien dat het Cartesiaans dualisme één van de grootste obstakels is voor het naturalisme. Het behaviorisme faalde echter als alternatief omdat het niet om kon gaan met intentionaliteit. Dankzij Darwin is Dennett echter in staat geweest om een alternatief te ontwikkelen waarbij hij onze intentionaliteit beschouwd als een afgeleide van al het ontwerp dat door de hijskranen van de evolutie verzet is. We hebben toen kennis gemaakt met het idee van het bewustzijn als een natuurlijk artefact. Dit alles was slechts de opmaat voor het derde deel: Dennett's verklaring van het bewustzijn. We zagen daar het idee van een geest in de machine terug komen in de vorm van het Cartesiaans theater. Dit idee ligt als impliciete veronderstelling ten grondslag aan onze intuïtieve ideeën over het bewustzijn. Het is echter niets anders dan een luchthaak. Dennett daarentegen legt het bewustzijn uit als het gevolg van drie hijskranen: natuurlijke selectie, fenotypische plasticiteit en culturele evolutie. Uiteindelijk is het bewustzijn inderdaad niets meer dan een natuurlijk artefact, een verzameling informatie verwerkende processen in ons brein, tot stand gekomen door een lang proces van Darwinistische evolutie. Dit idee stoot mensen tegen de borst maar alle bezwaren die ze maken voeren linksom of rechtsom terug naar een luchthaak in plaats van een alternatieve hijskraan.

In *Consciousness explained* duikt te pas en te onpas Dennett's sceptische tegenspeler Otto op. Otto denkt zo het zijne van de merkwaardige theorie die Dennett in dat boek ontvouwt en zoekt in de latere hoofdstukken herhaaldelijk de dialoog met de auteur. Tot nu toe heeft hij zich gedeisd gehouden maar dit is het moment waarop ook wij hem niet meer kunnen negeren en enige aandacht zullen moeten schenken aan de bezwaren die hij al zo lang voor zich heeft moeten houden.

Otto: Dank je. Het is jammer om te zien dat de ideeën van Dennett zo'n sterke invloed op je lijken te hebben. In de hele scriptie kom ik nauwelijks enige vorm van kritiek tegen en dat terwijl de filosofie van Dennett toch enkele bizarre consequenties heeft. Zoals ik ook al tegen hem zelf gezegd heb (CE, blz.362-368) kan ik zijn filosofie moeilijk een verklaring van het bewustzijn noemen. Hij lijkt eerder te ontkennen dat er zo iets als een bewustzijn bestaat.

HS: Hmm, ik wil niet muggenziften over details. Zou het je tevreden stellen als ik akkoord ging met de vaststelling dat er ook bij Dennett nog steeds sprake is van een bewustzijn, maar dat dit bewustzijn iets heel anders is dan het bewustzijn dat we kennen vanuit onze directe ervaring? Bepaalde zaken die we als vanzelfsprekend ervaren blijken anders te zijn dan we gedacht hadden. Zo is het voor mij heel natuurlijk om te veronderstellen dat de bedoeling van mijn gedachten voor mezelf volledig transparant is. Ik weet wat ik bedoel als ik iets zeg omdat ik het zelf ben die betekenis aan mijn eigen woorden geef. In de werkelijkheid ontstaan mijn gedachte echter als het gevolg van een anarchistisch proces waar niemand de baas over is. Zo is het volgens Dennett's theorie maar dat is niet hoe ik het ervaar. Er is dus wel sprake van een bewustzijn, alleen een ander bewustzijn dan we gedacht en misschien ook wel gehoopt hadden.

Otto: Daar ga ik zeker niet mee akkoord! Dit is precies mijn punt! Zonder de veronderstelling dat *ik* het ben die mijn gedachten denkt, kunnen we niet van een bewustzijn spreken. In dat geval zijn mijn gedachten niet meer dan gecodeerde informatie die door een mechanisch proces in mijn brein bewerkt worden maar er is niemand die ze ook daadwerkelijk *denkt*! Ik ben geen computer zoals Dennett en jij lijken te veronderstellen. Van een computer kunnen we alleen in metaforische zin zeggen dat hij bepaalde bedoelingen heeft. Zijn bedoelingen zijn een afgeleide van de bedoelingen van de maker. Onze bedoelingen daarentegen zijn authentiek, ze ontspringen aan niemand anders dan aan onszelf omdat wij over een bewustzijn beschikken.

HS: Ik zie waar je bezwaar zit. Op verschillende plekken in mijn scriptie heb ik geprobeerd dit te weerleggen. Laat ik nog een laatste poging wagen aan de hand van een gedachtenexperiment dat Dennett zelf heeft voorgesteld in een eerder essay. Daarbij zal ik tevens proberen alle losse eindjes van mijn verhaal aan elkaar te binden en zal duidelijk moeten worden wat nu precies de erfenis van Darwin is op de filosofie van Dennett.

Otto: Ga je gang, maar ik waarschuw je: ik ken de trucjes van Dennett inmiddels wel zo'n beetje. Weinig kans dat je me met een nieuw gedachtenexperiment kunt overtuigen.

HS: We zullen zien. Stel je voor dat je bezeten bent door de vurige wens om de toekomst te leren kennen. Een manier om die wens in vervulling te laten gaan is om jezelf in te vriezen tot bijvoorbeeld het jaar 2401 en dan weer te laten ontdooien en tot leven te laten wekken. De techniek om een dergelijke onderneming tot een succesvol eind te brengen is natuurlijk nog niet ontwikkeld, maar stel je eens voor dat dit wel het geval is. Hoe zou je een dergelijke onderneming dan aanpakken? Er zijn verschillende ontwerpproblemen waar je een oplossing voor moet vinden. Om te beginnen moet je de vrieskist waarin je ligt veilig stellen voor alle mogelijke bedreigingen die in de toekomst je pad zouden kunnen kruisen. Je kunt dit op twee manieren aanpakken. Je zou een geschikte locatie kunnen zoeken, daar een bunker bouwen en hopen op het beste. Je zou

echter ook een mobiele vrieskist kunnen ontwerpen die zelf op een actieve manier toekomstige dreigingen uit de weg probeert te gaan. Zo'n intelligent voertuig vraagt een hoop ontwerp. Hij moet bijvoorbeeld in staat zijn om zijn eigen brandstof bij elkaar te sprokkelen. Het liefst zouden we willen dat hij kleine mankementen zelf kan repareren en zo zijn er nog een boel vereisten waar hij aan zal moeten voldoen. Één van de dingen waar je rekening mee moet houden is dat je idee aan zal slaan en dat er meer mensen zijn die hun eigen overlevingsrobot gaan bouwen. Jouw robot zal dan de competitie aan moeten gaan met al deze andere robots om de schaarse hulpbronnen. In zo'n situatie doe je er verstandig aan hem het vermogen te geven om te leren zodat hij flexibel kan inspelen op nieuwe omstandigheden. Je zult hem dus ook uit moeten rusten met de mogelijkheid om zijn eigen pad te zoeken. Hij moet zijn eigen intenties kunnen onderzoeken en in staat zijn om bepaalde subdoelen tot de zijne te maken al naar gelang de situatie dit vraagt. Het is niet ondenkbaar dat een dergelijke robot gedurende zijn bestaan besluiten zal nemen die indruisen tegen hetgeen je oorspronkelijk van plan was. Bepaalde missies die hij onderneemt kunnen wellicht ronduit suïcidaal zijn vanuit jouw oogpunt. Toch is een dergelijke robot die bezit over voldoende vrijheid om zelfstandig afwegingen te maken geen slechte keuze. Al met al is het misschien wel de enige optie die je hebt als je wilt dat je robot zijn kostbare lading veilig bewaard tot 2401 (IS, blz.295-298).

Otto: Ik zie waar je naar toe gaat. Toch blijf ik erbij dat een dergelijke robot niet beschikt over authentieke bedoelingen. Zelfs als hij zijn eigen bedoelingen kan herzien dan nog komt dit omdat dit de bedoeling van zijn ontwerper was.

HS: Dat had ik wel zo'n beetje zien aankomen. Maar dan moet je de conclusie onder ogen zien dat wij zelf ook niet over zogenaamde 'authentieke bedoelingen' beschikken. Het verhaal dat Dennett hier verteld is namelijk precies hoe onze eigen evolutie verlopen is vanuit het oogpunt van onze genen (IS, blz.298). Dawkins heeft laten zien dat wij niets anders zijn dan overlevingsmachines ontworpen door zelfzuchtige genen (Dawkins, 1976).

Otto: Er is een verschil. De robot in Dennett zijn gedachte-experiment is ontworpen omdat zijn ontwerper een bepaald doel bewust voor ogen stond. Onze genen hadden geen doel toen ze ons ontwierpen, in ieder geval geen bewust doel.

HS: Dat klopt, maar er is een duidelijke overeenkomst. Laten we nog even terug gaan naar onze bespreking van natuurlijke selectie in deel drie. We zagen toen dat de eerste micro-organismen die de wereld bevolkten geen werkelijke belangen hadden. We kunnen achteraf gezien echter wel vaststellen dat enkele exemplaren succesvoller waren dan hun soortgenoten doordat ze toevallig over eigenschappen beschikten die van dienst kwamen in de strijd om het bestaan. Deze eigenschappen ontstonden zonder reden maar werden geselecteerd vanwege het nut dat ze hadden. We kunnen dus zeggen dat ze een bepaald doel dienden. Naar mate de organismen complexer werden en zich meer gingen richten op een specifieke tactiek om te overleven werden deze doelen beter te onderscheiden voor de buitenstaander (IS, blz.31). Maar het principe blijft hetzelfde. Alle bedoelingen die we in de natuur aantreffen zijn afgeleide bedoelingen. Het mooie aan de theorie van Darwin is dat we dit alles kunnen verklaren zonder dat we een intelligente schepper hoeven te veronderstellen. Dankzij Darwin kunnen we nu begrijpen dat al het ontwerp dat er nodig is om zoiets complex als een menselijk brein te produceren uiteindelijk verklaard kan worden door een mechanisch proces als natuurlijke selectie. Dennett wil laten zien dat hetzelfde, *mutatis mutandis*, van toepassing is op hoe bedoelingen en intentionaliteit ontstaan in onze geest. Er is niet sprake van één auteur aan wie al onze gedachten ontspringen en die garant kan staan voor een juiste betekenis. In plaats daarvan komt ons bewustzijn tot stand via een mechanisch proces en is alle betekenis een afgeleide van dit proces. Een proces dat bovendien gehoorzaamt aan dezelfde Darwinistische wetten. De radicale conclusie die Dennett hieruit trekt is dat we niet beschikken over zoiets als 'authentieke bedoelingen' (IS, blz.300). Al onze intentionaliteit

is slechts een afgeleide van het proces van selectie dat ons gevormd heeft, zo als dit ook het geval is voor de robot wiens intentionaliteit een afgeleide is van de bedoelingen van de maker.

Otto: Toch blijft het moeilijk te accepteren...

HS: Het is een gedachte waar veel mensen afwijzend tegenover staan. Darwin probeerde de intelligente ontwerper uit de natuur te verdrijven en veel mensen vonden dat al een verontrustend idee. Dennett gaat nog een stapje verder en komt tot de conclusie dat er ook in ons hoofd geen plaats is voor een dergelijke autoriteit. In beide gevallen blijkt er geen luchthaak nodig te zijn om bedoeling, functie en intentionaliteit te verklaren maar zijn het gewoon een stel hijskranen die al het werk doen. Mensen vinden dit verontrustend omdat hun zelfbeeld hierdoor flink op zijn kop gezet wordt. Natuurlijk levert zo iets weerstand op. Dit is ook de reden waarom ik in de inleiding de theorie van Dennett minstens zo revolutionair noem als die van Darwin. Natuurlijk kan ik niet in de toekomst kijken om te voorspellen welke impact het werk van Dennett zal hebben op een nieuwe generatie filosofen. Maar ik zie wel dat de theorie over het vermogen beschikt om vast geroeste ideeën uit te dagen, precies zoals dit ook bij Darwin het geval was.

Otto: Ik blijf er wel bij dat je weinig kritisch bent waar het de theorie van Dennett betreft.

HS: Vergeet niet dat dit slechts een bachelor scriptie is. Er volgt nog een master en wie weet ooit een proefschrift...

Otto: Ik ben dus nog niet van je af...

HS: Reken er maar niet op.

Otto: Tot ziens dan.

HS: Tot ziens en bedankt.

### **Werken van Dennett:**

**CC** 1969 Content and Consciousness. New York: Routledge

**IS** 1987 The Intentional Stance. Cambridge, MA: MIT Press.

**CE** 1991 Consciousness Explained. London: Penguin Books.

**DB** 1993 Back from the Drawing Board. In: Dahlbom 1993. Blz. 203-235.

**DDI** 1995 Darwin's Dangerous Idea. Evolution and the Meaning of Life. London: Penguin Books.

**FE** 2003 Freedom Evolves. London: Penguin Books.

**AE** 2008 Autobiographical Essay. Oorspronkelijk gepubliceerd in drie delen in: Philosophy Now edities: July/ August, September/ October, November/ December 2008.  
Geraadpleegd op 12-07-2012 via: <http://ase.tufts.edu/cogstud/incpages/publctns.shtml>

## **Literatuur:**

Bloom, Paul (2005). *Descartes' Baby. How the Science of Child Development Explains What Makes Us Human*. New York: Perseus Books.

Buskes, Chris (2006). *Evolutionair Denken. De Invloed van Darwin op ons Wereldbeeld*. Amsterdam: Uitgeverij Nieuwezijds.

Churchland, Paul M. (1988). *Matter and Consciousness*. Revised edition. Cambridge, MA: MIT Press.

Dawkins, Richard (1976). *The Selfish Gene*. Oxford: University Press.

Dawkins, Richard (1987). *The Blind Watchmaker*. London: Penguin Books.

Dahlbom, Bo ed. (1993). *Dennett and his Critics*. Cambridge, MA: Blackwell Publishers.

Descartes, R. (1644). *Principia Philosophiae*. Vertaling opgenomen in: Idem (1985). *The Philosophical Writings of Descartes*. Volume I. Cambridge: University press. Blz. 177-291.

Dewey, John (1910). *The Influence of Darwin on Philosophy*. In: Idem (1997) *The Influence of Darwin on Philosophy and Other Essays*. New York: Prometheus Books, Blz. 1-19.

Elton, Matthew (2003). *Daniel Dennett. Reconciling Science and our Self Conception*. Cambridge: Polity.

Putnam, Hilary (1960). *Mind and Machines*. In: Idem (1975). *Mind, Language and Reality*. Philosophical papers volume 2. Cambridge: University Press. Blz. 362-385.

Turing, Alan M. (1950). *Computing Machinery and Intelligence*. In: Boden, Margaret A. ed. (1990) *The Philosophy of Artificial Intelligence*. Oxford: University Press. Blz. 40-66.

Wittgenstein, Ludwig (1953). *Philosophical Investigations*. Vertaald als: Idem (2002). *Filosofische Onderzoekingen*. Amsterdam: Uitgeverij Boom.