

# Een mathematisch-fysicus over Dennett en Metzinger

with English summary

Ferdinand Verhulst  
Mathematisch Instituut, Universiteit Utrecht  
Postbus 80.010, 3508 TA Utrecht  
email: f.verhulst@uu.nl

4 december 2024

## **Samenvatting**

This review of monographs by the philosophers D.C. Dennett and T. Metzinger starts with a summary of causality ideas and other tools like metaphors and models of modern science. We note that Dennett has produced a lot of stimulating material on evolution and consciousness. His statement that the notion of identity of humans is a “user illusion” is more a hypothesis than a qualification in line with modern dynamical systems theory.

The extensive use of concepts by T. Metzinger is confusing and completely ignores the achievements of modern science. The description of neural experiments is missing a systematic set-up. His conclusions on the so-called first-person-experience and the non-existence of SELF have no scientific foundation.

We kennen de globale gegevens over de evolutie. Tien miljard jaar na het begin van het heelal werd uit een reusachtige en dichte stofwolk een ster gevormd, onze Zon, In dezelfde periode ontstonden de planeet Aarde en nog meer planeten rond de Zon. Na één of twee miljard jaar ontstond op de jonge Aarde een vorm van leven, de eencellige organismen, na verloop van tijd ontstonden steeds meer vormen van leven, zie bijvoorbeeld [10].

Dat er levende wezens kwamen was opmerkelijk maar heel bijzonder is het verschijnsel dat een aantal levensvormen bewustzijn vertoont. Het menselijk bewustzijn is pas ontstaan in het laatste 5000ste tijdsdeel van het bestaan van de Aarde, ongeveer de laatste miljoen jaar. Het is verbijsterend dat we het ontstaan, de werking en het functioneren van bewustzijn nauwelijks of niet begrijpen.

Neurologie, psychologie, de natuurwetenschappen en filosofie leveren bijdragen aan de studie van wat bewustzijn is. In het volgende behandelen we uitsluitend hypotheses over

bewustzijn die kunnen worden getest. Levensbeschouwelijke overwegingen, hoe belangrijk ook, blijven buiten beschouwing.

## 1 Wetenschappelijk gereedschap

Bij het analyseren en verklaren van verschijnselen is het essentieel om inzicht te hebben in oorzaak en gevolg problematiek. Dat spitst zich toe bij vragen over het bewustzijn van de mens die voor zijn hersenen en centraal zenuwstelsel over meer dan 100 miljard componenten, zenuwcellen, beschikt. Voor zo'n reusachtig dynamisch systeem is nauwelijks theorie beschikbaar. De term "dynamisch systeem" wordt gebruikt voor een stelsel van met elkaar verbonden componenten dat door onderlinge en externe interacties evolutie in de tijd vertoont. Bij die interacties speelt de oorzaak en gevolg problematiek, in het algemeen 'causaliteit', een essentiële rol. Een tegenstelling bestaat er tussen degenen die zeggen "het geheel is niets meer dan de som van de delen" en zij die zeggen "het geheel kan door emergentie iets nieuws leveren in vergelijking met de som van de delen", beide uitspraken veronderstellen een reduktionistische denkwijze. Tot de reduktionisten kun je bijvoorbeeld de fysici Carroll [1] en Hossenfelder [11] rekenen, ze zijn voorzichtig in hun uitspraken. Ook de filosofen Dennett [5]-[6] en Metzinger [13] redeneren reduktionistisch, we zullen zien dat ze echter nogal voorbarig zijn met hun conclusies

Het inzicht in het onderwerp 'causaliteit' heeft zich de afgelopen eeuwen fundamenteel gewijzigd. We bespreken dat eerst. Daarnaast is het nuttig om een goed omschreven idee te hebben van de concepten 'metafoor' en 'model'. die beide een rol spelen in wetenschappelijke discussies.

### 1.1 Oorzaak en gevolg

Een basis inzicht dat Aristoteles had was het volgende. Indien we een verschijnsel waarnemen dan moet er iets zijn dat het verschijnsel veroorzaakt. Dat 'iets' noemt men de 'oorzaak', het verschijnsel het 'gevolg'. Het fundamentele inzicht van Aristoteles was dat verschijnselen niet willekeurig of door magie gebeuren maar een duidelijk fysieke oorzaak hebben. Van causaliteit in strikte zin is hier nog geen sprake, fysieke gebeurtenissen hebben bij Aristoteles een doel-oorzaak. "Een steen die wordt geworpen valt omdat hij bij de Aarde hoort en daar naar terugkeert".

Vanaf de Renaissance tijd ontstonden gedachten over oorzaak en gevolg in een moderne en vruchtbare formulering: De dingen die in onze wereld gebeuren zijn alle het gevolg van oorzaken die gezamenlijk gehoorzamen aan wetten. Zo werd het vallen van een steen begrepen als beweging onder invloed van de zwaartekracht. Dezelfde kracht bepaalt de periodieke beweging van een slinger.

Het begrip 'wet' dat ook al in de Griekse oudheid werd gehanteerd speelt een grote rol in de wetenschap. *We zoeken naar wetmatigheden om de werkelijkheid te beschrijven want we nemen aan dat er wetten bestaan die oorzaak en gevolg bepalen.* In de vorm waarin dit in de 17e en 18e eeuw werd toegepast, vooral in de klassieke mechanica van Newton, noemen

we dit *naïef determinisme*.

Dat gaf bij het begin van de moderne wetenschap, in de 17e eeuw toch al problemen. Isaac Newton formuleerde zijn bewegingswetten van de klassieke mechanica waarbij beweging van de planeten door zwaartekracht een belangrijke toepassing was. Maar niemand kon uitleggen waar die zwaartekracht vandaan kwam en hoe deze door luchtledig kon werken. Er was nog meer. In de 17e eeuw kwam het inzicht dat de uitkomst van een handeling of proefneming niet altijd eenduidig (*naïef deterministisch*) vastligt maar met een bepaalde kans kan plaatsvinden. Zo ontstond het begin van de kans- en waarschijnlijkheidsrekening. Christiaan Huygens beweerde in de 17e eeuw dat licht een golfverschijnsel is, hij ondersteunde dit met experimenten en lange, correcte berekeningen. Newton daarentegen meende dat licht bestaat uit deeltjes, dat kwam hem beter uit in verband met de passage van licht door luchtledig die anders moeilijk te verklaren is. Een theoretische verklaring gaf Newton niet. Wie had er nu gelijk?

Het naïeve begrip ‘oorzaak en gevolg’ werd in de moderne tijd, vanaf rond 1900, ernstig verzwakt en gewijzigd. In het begin van de 20ste eeuw werden geheel nieuwe experimenten gedaan en schokkende nieuwe theorieën ontwikkeld. Het interessante maar ook moeilijke antwoord over het meningsverschil tussen Huygens en Newton over de aard van licht is dat ze beiden gelijk hadden. Het moeilijke hiervan is dat we moeten accepteren dat de werkelijkheid, in dit geval het licht, zich op een manier voordoet die afhankelijk is van de manier van meten. Rond 1900 kwam Planck met het revolutionaire idee dat licht bestaat uit quanta, energiepakketjes. Dat was een geheel nieuwe metafoor voor het verschijnsel licht. In de periode 1900-1930 werd de quantummechanica ontwikkeld.

Er is nog meer verwarring. Het beeld dat we hebben van ‘deeltjes’, bijvoorbeeld over het electron dat elektrisch negatief geladen is en deel uitmaakt van een wolk van electronen die om een kern draait maar ook voorkomt als een elektrische stroom die door een draad kan vloeien is erg misleidend. De positie van een electron heeft een verwachtingswaarde, er is een kans om het op een bepaalde plaats aan te treffen; we geven de positie van een elementair deeltje als een electron aan met een wiskundige golf functie. Als deze opzet van discrete posities van deeltjes en de rol van waarschijnlijkheid en verwachtingswaarden ons simpele beeld van oorzaak en gevolg aantast dan is de onzekerheidsrelatie van Heisenberg wel de beslissende klap. De onzekerheidsrelatie zegt dat we van een materieel deeltje op een bepaald moment *nooit* positie en snelheid beide precies kunnen bepalen. Het product van de foutafwijkingen van positie en snelheid zal altijd een klein positief getal zijn.

Een opmerkelijk nieuw gebied van onderzoek is de chaostheorie die in de tweede helft van de 20ste eeuw werd ontwikkeld. Al in de tweede helft van de 19e eeuw gaf Henri Poincaré de aanzet voor het verschijnsel chaos maar dit werk werd vele decennia niet begrepen en daardoor genegeerd. De chaostheorie is in heftige tegenspraak met het gewoontedenken omdat deze de causaliteit ondermijnt in problemen waar we de natuurwetten goed kennen en naïef deterministisch toepassen maar waar toch na korte tijd onvoorspelbaarheid optreedt. Het bekendste voorbeeld is het gedrag van de Aardatmosfeer met steeds wisselende weerregimes. Het kan in W-Europa regenen, soms lang, soms maar enkele dagen, de

natuurwetten die hierbij een rol spelen kennen we zeer goed, toch kunnen we het weer niet op lange termijn voorspellen.

We zullen spreken van *complex determinisme* bij het bestuderen van verschijnselen indien bij de theorievorming stochastische, quantummechanische of chaos aspecten worden betrokken.

In verband met het gebruik van computers en algoritmen in de wetenschap noemen we nog twee fundamentele aspecten. Allereerst de complexiteitsklassen P en NP. Tot de complexiteitsklasse P horen problemen die rekenkundig in polynomiale tijd kunnen worden opgelost. Stel we hebben voor een probleem  $n$  stappen in de berekening nodig. De complexiteit van het probleem is polynomiaal als de rekestijd evenredig is met het aantal rekenstappen  $n$ , bijvoorbeeld lineair evenredig of kwadratisch evenredig. Tot de complexiteitsklasse NP horen problemen die niet in polynomiale tijd kunnen worden opgelost, bijvoorbeeld omdat er een exponentieel grote tijd nodig is. Als we een praktisch rekenprobleem als NP herkennen zal de rekestijd een grote en soms onmogelijke obstructie vormen.

Een tweede probleem zou kunnen worden veroorzaakt door de incompleetheids stelling van Gödel. Deze zegt dat in elk voldoende ingewikkeld stelsel van axioma's en stellingen er uitspraken zijn waarvan we niet kunnen bewijzen dat ze waar of onwaar zijn. Om misverstanden te voorkomen: we kunnen dat niet bewijzen omdat we te dom zijn, we kunnen principieel voor sommige resultaten geen logische redenering geven die tot waar of onwaar leidt. Het is niet duidelijk of de incompleetheids stelling van Gödel gevolgen heeft voor de beschouwing van menselijk bewustzijn, maar het heeft wellicht gevolgen voor onze kijk op het gebruik van computers met algoritmische rekenmethoden en voor AI.

## 1.2 Metaforen en modellen

We zullen het woord 'model' gebruiken in de zin van een precieze wiskundige beschrijving, zo precies als nodig is voor de beschrijving van een verschijnsel in de werkelijkheid. Het geeft niet dat er veel details zijn weggelaten, de essentie van de fysieke gebeurtenis wordt door het wiskundig model goed beschreven. Alle andere aspecten en details zijn afwezig in het model. Het is goed om te weten dat we bij zeer complexe verschijnselen al informatie en inzicht krijgen door middel van een sterk vereenvoudigd 'conceptueel model'. Een model heeft altijd zowel kwalitatieve als kwantitatieve aspecten.

Een model is ook een metafoor, maar dat is een veel ruimer begrip. Aristoteles schrijft in zijn *Poetica*: "Een metafoor is een beschrijving van iets in termen van iets anders." Een model is dus ook een metafoor.

Zoals je goede en gebrekkige modellen hebt, zo heb je ook goede en slechte metaforen. In het geval van metaforen is 'slecht' ernstiger omdat het kwantitatieve aspect en daarmee de mogelijkheid van testen ontbreekt. Bij modellen maakt het kwantitatieve aspect dat je door meting de beperkte waarde van het model kunt vaststellen. Metaforen kunnen bedrieglijk zijn. We zullen in het vervolg zien dat het gebruik van de computer als metafoor voor de menselijke hersenen niet goed bruikbaar is. In [?] worden metaforen voor processen in de menselijke geest voorgesteld. Zo zou de oorzaak van het optreden van dwangneuroses kunnen worden gezien als het bestaan van een 'vreemde attractor' in de

hersenen (een vreemde attractor in de theorie van dynamische systemen is een verschijnsel dat permanent aantrekkend en afstotend is). Deze metafoor levert nieuw taalgebruik en beelden maar testen kun je deze metafoor niet.

We noemen kort de eisen die je aan onderzoek van bewustzijn moet stellen en toetsen dat aan het werk van twee filosofen van naam.

1. Metaforen en modellen spelen een grote rol in de beschrijving van bewustzijn. Zo noemt de neuroloog-psycholoog Damasio [3]-[4] het resultaat van een externe ervaring door de mens een representatie van de ervaring die een plaats krijgt in een deel van de hersenen. Dat is een bruikbare metafoor, het geeft een naam aan een bepaalde werking van neuronen veroorzaakt door die ervaring. Bij model denken we aan een verder ontwikkelde metafoor die behalve kwalitatieve eigenschappen ook kwantitatieve heeft welke kunnen worden getest door experimenten.
2. Het naïeve determinisme dat tot ruwweg 1900 opgang deed is onder invloed van stochastiek, quantummechanica en chaostheorie vervangen door complex determinisme. Redeneringen, met name reductivistische, dienen niet te worden gebaseerd op naïef determinisme, zie ook punt 3..
3. De hersenen met meer dan honderd miljard neuronen en het centraal zenuwstelsel spelen een hoofdrol in het bestaan en de werking van het bewustzijn. Dit is een zeer complex dynamisch systeem waarover nog weinig bekend is. Het is dan zeker te verwachten dat *patroonvorming, emergentie en zelf-organisatie* bij zulke systemen een rol zullen spelen; zie de overzichten [12] en [2].
4. De empirie speelt een grote rol bij de theorievorming van het bewustzijn. Experimenten moeten worden voorzien van hypothesen, doelstellingen en conclusies. Experimenten moeten herhaalbaar zijn.

In het vervolg bespreken we het werk van de filosofen Dennett en Metzinger. Commentaren zijn te vinden in de paragrafen ‘Discussie en Conclusies’, bovendien in korte cursieve tussenvoegsels.

We geven hier vast de hoofdconclusies. Ondanks een aantal interessante opmerkingen van Dennett en Metzinger zijn er twee grote bezwaren. Hun taalgebruik van metaforen en modellen sluit niet aan bij de moderne wetenschap en is in het 19e eeuwse denken blijven steken. Dat laatste geldt ook voor hun onvermogen om het verband te leggen tussen lichamelijke neuronale processen en geestelijke bewustzijnsprocessen; dat onvermogen hangt samen met het gebruik van naïef determinisme en gebrekkige kennis van de hedendaagse theorie van dynamische systemen welke inzicht kan geven in de werking van de hersenen. Zie ook [2] en [12].

## 2 Dennett over bewustzijn

Er is in de loop der tijden veel nagedacht over de ontwikkeling van de menselijke geest. D.C. Dennett was, hij overleed in 2024, een gezaghebbende Amerikaanse filosoof die zich uitvoerig heeft bezig gehouden met de evolutie van de mens en het begrip bewustzijn. ‘Consciousness Explained’ [5] is een van zijn bekendste werken.

### 2.1 Dennett: Consciousness Explained

Wie de ruim 500 bladzijden van het boek leest wordt getroffen door Dennett’s verbale begaafdheid, zijn speelsheid en brede kennis. In deel 1 toont hij zich zeer strijdbaar in de discussie met creationisten. In deel 2 van het boek karakteriseert Dennett het beeld dat Descartes schetst van de band tussen de geest en de werkelijkheid buiten het lichaam als ‘het Cartesisch theater’. Dat beeld geeft de gedachten van Descartes goed weer, zie [7]-[8]-[9]., maar er ontbreekt iets essentieels aan die gedachte en dat beeld. We zijn niet een eenzame waarnemer van ons lichaam en de werkelijkheid daarbuiten, we zijn deel van de werkelijkheid in voortdurende interactie daarmee. Die interacties hebben veel kanten. Hoofdstuk 13 gaat over de werkelijkheid van het zelf. Net zoals de andere dieren zijn we ons bewust van grenzen die onze geest en lichaam afbakenen; daar gebruiken we kleding voor en zelfs afscheidingen en muren. Dennett noemt het geval van Meervoudige Persoonlijkheds Storing waarbij patiënten meer dan een zelve in zich herbergen. De vraag is hoe onze hersenen aan meer dan een persoonlijkheid ruimte kan bieden. (*In de theorie van dynamische systemen is het opreden van meervoudige attractoren een normaal verschijnsel. De vraag die zou moeten worden gesteld waarom niet veel meer mensen een Meervoudige Persoonlijkheds Storing hebben. Misschien is dat ook het geval en zijn bij veel mensen die andere persoonlijkheden (attractoren) zwak van aantrekking en klein van bereik. Er zijn dan speciale omstandigheden nodig om ze tot activiteit te brengen.*) Het beeld van deze storing past overigens slecht in het Cartesisch theater van de geest. Dennett merkt nog op dat ‘het zelf ’ geen permanente entiteit is, maar dit verbaast niet daar de mens zich ontwikkelt en voortdurend onder invloed van extern wisselende omstandigheden is. Wat is ‘het zelf ’ dan wel. Het is volgens Dennett een voortdurend verhaal over onszelf gebaseerd op een continue zelfrepresentatie.

### 2.2 Dennett: From Bacteria to Bach and back

Creationisten zijn in Europa actief maar veel meer in de Verenigde Staten. Dennett gaat in zijn boek ‘From Bacteria to Bach and back’ [6] uitvoerig in op de discussie met creationisten: hoe kan een gecompliceerd organisme dat in staat is tot reproductie zijn ontstaan zonder aanwezigheid van een ontwerper (‘creative designer’). De aanhangers van ‘Intelligent design’, de orthodoxe aanhangers van de traditionele godsdiensten, zeggen “Die moet er wel zijn”. De oplossing ligt echter in wat Darwin noemde “natuurlijke selectie”. Organismen ontwikkelen zich in een omgeving waar levensmogelijkheden zijn, eigenschappen van het organisme die overbodig zijn of zelfs een bedreiging van het voortbestaan zullen

verdwijnen. Dat betekent allereerst de noodzakelijkheid van verandering en aanpassing van een organisme, maar volgens Dennett houdt het ook optimalisering in. Namelijk, wat is de eenvoudigste chemische of fysieke structuur die geschikt is voor een bepaalde taak? Tevens stabiliteit zodat een bepaalde fysieke en chemische structuur zich in zijn omgeving kan handhaven. (*Die optimaliteit is twijfelachtig, tijdens de evolutie blijven restanten die geen nut hebben maar ook niet schadelijk zijn.*)

Dennett bespreekt een belangrijk proces dat een rol speelt bij levende wezens. Hij onderscheidt vaardigheden ('competences') en begrip ('comprehension').

Bekijk eens het programmeren van een automatisch werkende lift. Het programma dat de werking van de lift beschrijft heeft een broncode ('source code') die veel commando's bevat. Als het liftprogramma goed is heeft het de juiste ontologie, dat wil zeggen het heeft alle gegevens om goed te kunnen functioneren in zijn liftwereld. Het programma herkent bijvoorbeeld in zijn codes wat een deur is. Bij dit alles heeft het liftprogramma geen kennis van of toegang tot het ontwerp van zijn programma. We zien hier een voorbeeld van vaardigheid zonder begrip.

We kunnen overigens in de broncode feedback inbouwen dat op bepaalde tijdstippen of na een storing een controle laat plaatsvinden van de werking, het liftprogramma functioneert dan nog beter. Ook dat leidt echter niet tot begrip. In dit alles lijkt het liftprogramma oppervlakkig op het functioneren van een bacterie, een paddenstoel of een boom (*de gelijkenis is beperkt, ze zijn ongelooflijk veel complexer dan de geprogrammeerde lift*). Het goede liftprogramma is trouwens ontworpen met vallen en opstaan, daarin lijkt het wel op het ontstaan van dingen in de natuur.

Komt begrip tot stand gebruikmakend van vaardigheden? Het is duidelijk dat dieren over vaardigheden beschikken, termieten bijvoorbeeld kunnen prachtige nesten bouwen, mensapen gebruiken soms takken als gereedschap, maar hebben ze ook begrip? Dennett's antwoord is bevestigend, dieren kunnen begrip hebben. Het is niet van dezelfde orde als bij de mens, mensen kunnen door het denken, analyseren, ervaren, experimenteren en uitwisseling zinvol onderzoek doen. Dat is bij dieren veel beperkter. Bij het voorbeeld van de lift die vaardig is maar die we geen begrip toeschrijven is het wel duidelijk. (*Hetzelfde zou kunnen gelden voor kunstmatige intelligentie (AI) indien een overmaat aan vaardigheden niet tot begrip leidt*).

Dennett vraagt dan hoe menselijk begrip tot stand komt door de activiteit van neuronen die van zichzelf niets begrijpen. (*Een vraag die kennis eist van het gedrag van grootschalige dynamische systemen.*) Evolutie is een langzaam proces waarbij volgens Dennett competenties (vaardigheden) volgens bruikbaarheid worden geselecteerd zonder dat er sprake is van comprehension (begrip). Speelt bewustzijn van het organisme hierbij een rol?

Dennett introduceert dan 'hogere dieren' als 'intentionele systemen' met de 'emergentie van begrip'. Daarentegen is bij het beschouwen van het gedrag van planten of lagere dieren het vaststellen van een bedoeling, een intentie, niet onjuist tenzij je aan de intentie begrip koppelt. Een plant of lager dier als een worm met veel vaardigheden is begaafd door de natuur maar een plant blijft bij Dennett een levende robot.

Dennett keert op het laatst terug tot de discussie of ons gevoel van een bewust wezen te zijn een illusie is. Hij noemt de vanzelfsprekendheid waarmee Descartes zijn eigen identiteit en denken aanvaardt de ‘first person point of view’ en de ‘user-illusion’. Het bewust functioneren van onze hersens in relatie tot ons hele lichaam is in de visie van Dennett een illusie.

## 2.3 Dennett: commentaar en discussie

Dennett heeft enorm veel los gemaakt met zijn discussies over evolutie en bewustzijn. Dat is een grote verdienste.

Het gaat wat ver om met Dennett de evolutie zo procesmatig te vereenvoudigen. Dennett noemt evolutie een algoritmisch proces, eigenlijk zoals er zandrichels op het strand ontstaan als de vloed golven richting strand stuurt. Het evolutie gebeuren heeft veel meer wendingen en kanten. Het is duidelijk dat een mens met twee ogen meer overlevingskansen heeft, want daardoor ziet hij diepte en kan afstanden inschatten. Of de fysieke structuur van mensen en andere dieren echt optimaal is lijkt twijfelachtig.

Als een kind vraagt wat die wolken in de lucht zijn, dan is het antwoord in termen van concepten die ook weer uitgelegd moeten worden. Dennett doet hetzelfde, hij formuleert concepten die een zekere verdieping opleveren maar ook weer meer analyse nodig hebben.

Zijn conclusie dat we ons bewustzijn als een ‘user-illusion’ ervaren is meer een opinie dan een goed gefundeerde hypothese. De theorie van dynamische systemen geeft duidelijk aan dat elk menselijk wezen een unieke entiteit is. Elk mens heeft een unieke fysieke bouw en evolutie geschiedenis, het gevoel van identiteit dat mensen hebben is dan geen illusie.

Een grote vraag bij Dennett is verder: hoe komt begrip tot stand gebruikmakend van vaardigheden? Dit is niet alleen een vraag maar ook een uitspraak. Namelijk: vaardigheden gaan aan begrip vooraf. De vraag is waarschijnlijk niet goed gesteld. Als je kijkt naar de ontwikkeling van technische vaardigheden bij de mens dan zie je dat een primitieve vaardigheid kan leiden tot een spoor van begrip. Dat maakt dan een verbetering van de vaardigheid mogelijk die op zijn beurt leidt tot meer begrip. Een voorbeeld is het ontstaan van de optica. De Babyloniërs deden zo’n 2000 jaar vChr. al beperkte oogoperaties en gebruikten een soort vergrootglazen welke bestonden uit transparant materiaal zoals Beryl (rotskristal). De Arabieren pakten dit in hun bloeitijd (800-1200 na Chr.) op door kleine dingen met vergroting beter te kunnen zien. In dezelfde tijd heeft iemand de bril uitgevonden, dat woord komt van ‘Beryl’. Weer veel later, in 1608, vond de Nederlander Zacharias Janssen in Middelburg de verrekijker uit. Dit werd later gevolgd door de uitvinding van de microscoop en zo ging het verder met de optica.

De vraag of veel vaardigheden tot begrip gaan leiden zal waarschijnlijk met “nee” moeten worden beantwoord. Mijn laptop, een MacBook Pro, bevat heel veel vaardigheden in de vorm van een geavanceerde broncode en een groot aantal computer routines. Hoeveel geheugen en hoeveel routines ook worden toegevoegd, het blijft een machine zonder begrip, laat staan bewustzijn. De vaardigheden zijn door een mens toegevoegd. Computer programmering kan niet worden gebruikt als metafoor voor werking en ontwerp (design)



van de menselijke geest. De hersenen werken niet algoritmisch en vertonen onverwachte werking die een geïntegreerde indruk maakt..

Het is slordig taalgebruik van Dennett om een plant een levende robot te noemen. Een robot heeft een beperkt en eindig aantal ontwerp instructies, een levend wezen heeft niet een eindig aantal discrete ontwerp instructies en vertoont altijd evolutie en flexibiliteit. Het heeft weinig zin om bepaalde dieren vaardig maar verstoken van begrip te noemen; zo'n uitspraak heeft alleen zin als je helder voor ogen staat wat bewustzijn en begrip is.

### 3 Metzinger over het 'Selbst'

Het boek de Ego tunnel [13] van Thomas Metzinger is een samenvatting van zijn publicaties. Het boek telt drie delen en negen hoofdstukken. De inleiding is opnieuw een samenvatting waarvan we twee punten noemen. Op de eerste bladzij stelt de auteur: "Na alles wat we tegenwoordig weten is er geen ding, geen enkele ondeelbare entiteit die we ZELF zijn, en dat is noch in de hersenen, noch in een of andere metafysische sfeer buiten deze wereld". (*Dat lijkt een beetje op Descartes, zie [7]-[8],-[9], die het lichaam deelbaar en de geest ondeelbaar noemt. Als je zo ZELF definieert is het duidelijk dat het ZELF niet bestaat, zo'n omschrijving heeft geen zin in een discussie. Het ZELF is, als het bestaat, een dynamische, evoluerende entiteit en niet statisch.*)

Als tweede voert de auteur de 'egotunnel' in, een metafoor voor het bewuste en zeer selectieve beleven van de werkelijkheid. (*De term is niet gelukkig; het is juist dat onze waarnemingen en ervaringen van de werkelijkheid selectief en zeer beperkt zijn, maar het woord tunnel geeft een statische toestand aan terwijl wijzelf en alles om ons heen voortdurend verandert. 'Tunnel' is een gebrekkige metafoor.*)

#### 3.1 Het bewustzijnsprobleem

Het is iets bijzonders dat zoogdieren een vorm van bewustzijn hebben. Bewustzijn betekent bij een wezen als de mens zowel het besef als de waarneming van de wereld om ons heen dat innerlijk plaatsvindt. Metzinger denkt dat ook vogels, reptielen en vissen een vorm van bewustzijn hebben.

Bij de mens verschijnt dan het beleven van een IK of ZELF als fenomenaal ZELF-model (PSM afgekort). Dat stelt ons in staat om met deze eerste-persoons-ervaring ook over ons zelf te denken, over onze gevoelens en emoties en over onze gedachtenwereld. (*Het woord 'model' wordt hier niet in wetenschappelijke zin gebruikt.*) Het bewustzijn is tamelijk robuust als ervaring hoewel het tegelijk voortdurend aan verandering onderhevig is. (*Dit is niet opmerkelijk, elk robuust dynamisch systeem kun je door kleine storingen in een andere toestand brengen.*) De hersenen zijn een biologische machine die als werkelijkheidsgenerator ons vertelt wat wel en niet bestaat. We hebben kennelijk een manier om met deze verschillende werkelijkheden om te gaan.

Onze ervaring van dingen buiten ons zelf leveren een neuronale representatie op in onze

hersenen (Damasio) (*dat is in feite een projectie van een oneindig-dimensionale werkelijkheid op een veel lager-dimensionale representatie*). Dat beseffen we niet, we hebben het gevoel dat we aanwezig zijn in een wereld zoals onze geest dat vaststelt.

Metzinger merkt op dat er door de eeuwen heen twee hoofdpunten in het bewustzijns onderzoek naar voren komen. Allereerst moraliteit, alleen iemand die een geweten heeft bezit bewustzijn in eigenlijke zin. (*Er wordt niet omschreven wat een 'geweten' is*). Een tweede fundamenteel inzicht is dat het bewustzijn integreert, verschillende delen op verschillende tijdstippen worden gezien als deel uit makend van één werkelijkheid. Dat onze geest van de duizenden indrukken in ons leven een geheel maakt is een van de grote prestaties van de hersenen. Allerlei fysieke stoornissen kunnen deze vaardigheden beperken. De neuroloog Tononi veronderstelt dat vanuit de miljarden neuronen een ontwerp opduikt met een krachtige interne structuur dat voor eenheid zorgt (*deze hypothese sluit goed aan bij de theorie van dynamische systemen*). De miljarden neuronen zijn er en vuren (werken) op prikkels maar hoe de globale interactie dynamica is weten we niet. (*De eerder genoemde emergentie en zelf-organisatie zouden een rol kunnen spelen.*)

Het bewustzijn ervaart een 'nu' en daarbij een verleden en anticipatie op de toekomst. Dat is het ervaren van de tijd (*bedoeld is de biologische tijd die een mens ervaart*). Daarin zijn ook weer veel schakeringen. In een staat van mijmerij of verstrooid zijn is het 'nu' niet goed aanwijsbaar. In die zin is ook de tijd een element van het tunnel effect. Het is tegelijk bijzonder dat een mens zeer veel uit het verleden in zijn geheugen heeft opgeslagen en ook uitkijkt naar de toekomst die van alles kan brengen. Deze ervaringen die met de tijd zijn verbonden geven een beperkt deel van de totale werkelijkheid waar we ons in bevinden. Metzinger gaat dan dieper in op ons waarnemingsvermogen met als voorbeeld twee kleuren groen op een stuk papier, zeg nr. 24 en nr.25, die dicht bij elkaar liggen. Veel mensen zullen de kleuren als licht verschillend zien. Indien men echter één van de kleuren toont zullen de meeste mensen niet weten of het nr. 24 of nr 25 is. Metzinger verbindt hier aan de conclusie dat we de werkelijkheid globaal waarnemen en representeren maar dat het ons aan subtiele en verfijnde representatie ontbreekt. Er zijn nog talloze voorbeelden te geven waarbij herkennen mogelijk is maar subtiel onderscheiden zeer moeilijk is. Dat geeft ook de beperkingen aan van communicatie tussen mensen. (*Misschien stelt de relatieve grofheid van ons waarnemingsvermogen ons juist in staat om de ontelbare indrukken uit onze omgeving te hanteren.*) Het is een probleem om bewustzijns onderzoek te doen gekoppeld aan hersenonderzoek als het moeilijk is om verschillende subtiel verschillende toestanden verbaal te benoemen.

Zoals eerder is aangegeven: we zijn representatie systemen, de werkelijkheid die we beleven is uniek voor onze eigen persoon en alleen gebrekking te delen. Volgens Metzinger is 'denken' echter iets nieuws, door denken wordt iets gemaakt dat er niet was. (*Hier lijkt de auteur met Descartes overeen te stemmen.*)

Het is mogelijk om met allerlei experimenten het beeld dat men van zijn eigen lichaam heeft kunstmatig te verstoren. (*Iets dergelijks gebeurde bij Oliver Sacks die zijn linkerbeen*

*‘kwijt’ was en door aandacht te geven zijn been weer kon gebruiken; zie [14] en ook zijn fascinerende boek Hallucinations [15] waarin vele gevallen van verstoord bewustzijn worden beschreven.)* Er zijn hier veel empirische resultaten over en ook veel varianten van zulke proefnemingen die op andere soorten bewustzijnstoestanden wijzen.

### 3.2 Ons handelen en de vrije wil

Alleen door een fenomenaal zelf-model (PSM) kunnen we het bezit van handen en voeten beleven. Bewuste beheersing van onze ledematen heeft echter meer nodig. Om dingen uit te voeren hebben we voor onszelf een voorstelling van ‘doelen’ nodig, in de filosofie ‘intentionaliteit’ genoemd. Om over een eigen geest te beschikken is niet alleen nodig dat we denken en weten, we moeten ook een handelende geest zijn met een eigen wil.

Er zijn dan twee mogelijkheden. Er bestaat geen wilsvrijheid, alle gebeurtenissen hebben fysische oorzaken, we zijn automaten. De tweede mogelijkheid is dan dat alles wat gebeurt een doel heeft, niets gebeurt toevallig. (*Beide mogelijkheden klinken onbevredigend, we voelen ons geen automaten en de tweede mogelijkheid klinkt als metafysica.*) Er zijn tussenoplossingen, zoals het meeste wat gebeurt heeft een deterministisch fysische oorsprong maar er bestaat een klein deelgebied in onszelf waar wij vrij onze doelen kiezen. Metzinger merkt op dat ons Zelf in wezen een rekenmachine is die bestaat uit en werkt met neurologische componenten. Deze machine beheerst de doelen, de fysische en de virtuele. (*Hier vervalt Metzinger weer in de primitieve metafoor van de hersenen beheerst door algoritmische en misschien zelfs uitsluitend naïef deterministische processen. Dit sluit niet aan bij de neurodynamische systeemtheorie*). De vraag hoe vrij we zijn wordt door Metzinger niet opgelost. Het naïef deterministische wereldbeeld is in tegenstelling met ons fenomenale zelf-model waarbij we ons als wezens beschouwen die ten allen tijde een nieuwe reeks van opvolgende beslissingen kunnen nemen. Een verklaring zou kunnen zijn dat bij beslissingen de eerst opkomende gedachte voor ons onbewust is en dat de bewuste gedachte die hier op volgt ons het gevoel van spontane en vrije beslissing geeft. (*Het is een goede gedachte, natuurlijk zullen persoonlijk aanleg, levenservaringen en opvoeding onbewust al een beslissing aandragen. Echter, in moeilijke gevallen zal de chaotiek of moeilijke beslisbaarheid die zich voordoet, denk aan NP problemen, leiden tot een zeer verzwakte vorm van determinisme.*) De discussie over de vrije wil heeft niet alleen filosofische en neurologische aspecten, het heeft ook veel sociale kanten, denk aan ons systeem van rechtspreken. Ook die wetsregels zijn sterk verbonden met ons fenomenale zelfbeeld. Een consequentie kan zijn dat mensen voor een misdaad niet bestraft moeten worden omdat ze hun foute handelen gewoon niet kunnen laten, ze zijn niet verantwoordelijk. Geen straf maar alleen rehabilitatie zou dan op zijn plaats zijn. Metzinger ziet ook grote gevaren in de discussie opduiken, vooral als relativisering van de vrije wil gebruikt wordt om asociaal gedrag te rechtvaardigen. Primair is dat de hersenen geen handelingen uitkiest en geen doelen kent. Zelforganisatie is een hoofdkenmerk. De hersenen volgen de wetten van de natuur (eind hoofdstuk 4 van [13]):

De hersenen laten zich het beste beschrijven als een complex systeem dat er voortdurend naar streeft om in een stabiele toestand te geraken en daarbij orde

uit chaos te weeg brengt.

(*Met deze uitspraken doet Metzinger verder niets.*) Waarschijnlijk is het wel zo dat wezens alleen kunnen overleven indien ze zich voorzien van ‘doel-representaties. Er is echter geen statische entiteit in ons aanwezig die doelen heeft en handelingen uitkiest.

### 3.3 Bewustzijnsrevolutie

Elk systeem dat een bewust zelf voortbrengt noemt Metzinger een ‘ego-machine’. Dat kan alles zijn wat een bewust zelf-model bezit. (*Wat is dit? In deze zeer algemene formulering hoeft zo’n systeem niet veel gemeen te hebben met het fenomenaal zelf-model dat eerste-persoon-ervaring heeft.*) Er wordt gesteld dat de opvatting dat er slechts twee soorten informatie verwerkende systemen zijn, kunstmatige en natuurlijke, onjuist is. Men kan immers levende wezens voorzien van sturende apparatuur en zelfs proberen een robot te ontwerpen die biologische neurale netwerken bezit. In deze tijd is het ontwerpen van kunstmatige neurale netwerken en bijbehorende hardware nog niet ver genoeg gevorderd. Eenvoudige robots hebben een zelf-model en bepaalde vermogens om zichzelf te herstellen of om te gaan met defecten. Onder welke voorwaarden kunnen we concluderen dat zulke kunstmatige of hybride systemen bewust leven en een bewust zelf-model hebben? In eerdere hoofdstukken werd gesteld dat het systeem een geïntegreerd beeld van de werkelijkheid moet hebben en een moraal moet hebben die gaat over zijn wereld en zijn medebewoners. Er dient bij het systeem ook een beleving van de tijd te zijn, van het ‘nu’. Een kunstmatig systeem dat een dergelijk bewustzijn en een persoonlijke beleving bezit zal ook kunnen lijden en zal daartegen beschermd moeten worden.

Overigens speelt dit in de evolutie geen primaire rol, het gaat daarbij om het zich handhaven in de werkelijkheid en het zich kunnen voortplanten. Optimaal gelukkig zijn en bescherming tegen lijden zijn geen doelen van de evolutie.

### 3.4 Metzinger: commentaar en discussie

In het voorgaande zijn al (cursief) een aantal opmerkingen gemaakt. We vatten een aantal punten samen.

1. Bij een aantal onderwerpen volgt het taalgebruik en de begripsvorming terecht Antonio Damasio, [3]-[4]. Dat geldt vooral Damasio’s kijk op de eenheid van geest en lichaam, ook het optreden van representaties in de hersenen als gevolg van ervaringen van de werkelijkheid. Dat zijn geschikte metaforen.
2. Het is soms lastig om Metzinger te volgen als een omschrijving van een begrip ontbreekt of als een term slecht gekozen is. Zo gebruikt hij ‘model’ niet in wetenschappelijke zin, er is verwarring met het begrip ‘metafoor’ en soms, zoals bij het begrip ‘zelf-model’, is de betekenis geheel onduidelijk. Het woord ‘tunnel’ in ego-tunnel is ongelukkig gekozen. Zoals Dennett voert hij begrippen in die verwijzen naar diepere lagen die niet omschreven zijn.

3. De experimenten zijn interessant maar ze tonen geen systematische opzet. Er ontbreken vooropgezette hypothesen en opvolgende conclusies. Veel meer experimenteel materiaal is te vinden bij Sacks [15] en Swaab [16].
4. De hersenen en het bewustzijn dat de mens bezit integreert de representaties die door waarnemingen zijn gevormd. Wanneer bijvoorbeeld het oog iets ziet wordt dit op verschillende manieren verwerkt in verschillende delen van de hersenen. Als ik bijvoorbeeld op de fiets voor een verkeerslicht wacht en het springt op groen, dan activeert mijn gezichts-zenuw een speciaal deel van de hersenen maar vervolgens ‘weten’ mijn hersenen ook dat ik mijn spieren kan gebruiken om te gaan fietsen. In alle gevallen werken de hersenen als een geïntegreerd orgaan.
5. Het bestaan van een ego-machine die bewust leeft, een moraal heeft en ook kan lijden lijkt iets voor SF romans.
6. Zowel Dennet als Metzinger benadrukken de eerste-persoon-ervaring die mensen een gevoel van identiteit geeft. Dennet noemt het zelfs een illusie, maar ook Metzinger lijkt het fundamentele belang van deze ervaring in de vorm van PSM te onderschatten. Een mens kan worden opgevat als een uiterst complex dynamisch systeem. Uit de theorie van dynamische systemen volgt dan echter dat elk mens een uniek subject is en inderdaad aanspraak maakt op een unieke identiteit.
7. In feite lijkt Metzinger terug te vallen op een soort reductionisme waarbij ondanks alle complexiteit die beschreven wordt het ZELF niet bestaat. Vanzelfsprekend niet als ondeelbare eenheid, dan zouden we met onze concepten terug gaan naar de middeleeuwen, maar ook niet als een fluïde, dynamisch ZELF zoals dat volgens de theorie van dynamische systemen zou kunnen ontstaan.  
Zover zijn we nog lang niet met onze theorie en ontkenning van het bestaan van het ZELF in dit stadium is voorbarig.

## Referenties

- [1] Sean Carroll, *The big picture*, Dutton (imprint of Penguin Random House, 2017).
- [2] Stephen Coombes and Kyle C.A. Wedgwood, *Neurodynamics, an applied mathematics perspective*, Texts in Applied Mathematics 75, Springer (2023)
- [3] Antonio Damasio, *Descartes' Error*, Vintage Books (1994, rev. 2006)
- [4] Antonio Damasio, *Self comes to Mind, constructing the conscious brain*, Vintage Books, (2012)
- [5] Daniel C. Dennett, *Consciousness Explained*, Little, Brown and Company (1991).
- [6] Daniel C. Dennett, *From Bacteria to Bach and back*, Penguin Random House (2018).

- [7] René Descartes, *Discours de la Méthode*, publ. J. Maire, Leiden (1637), Nederlandse vert. Helena C. Pos Wereldbibliotheek (1937)
- [8] René Descartes, *Méditations*, (originele ed.Latijn, 1641, Duc des Luynes Franse vert. 1647), Trilingual Edition, Manley, D.B. and Taylor, C.S (eds. 1966), <https://corescolar.libraries.wright.edu/philosophy/8>  
Engelse vertaling met noten door Michael Morirarty, Oxford World's Classics, Oxford University Press (2008)
- [9] René Descartes, *Les Passions de l'Ame*, Ed. Legras, Parijs (1649), Samen met de Discours de la Méthode in 'Génie de France', ed. REé Hilsumm Paris. Nederlandse vertaling "Passies van de Ziel" door Theo Verbeek, Historische Uitgeverij (2008)
- [10] Y.N. Harari, *Sapiens, a brief history of humankind*, Random House (2014).
- [11] Sabine Hossenfelder, *Existential Physics*, Viking (2022).
- [12] Eugene M. Izhikevich, *Dynamical Systems in Neuroscience, the geometry of excitability and bursting*. The MIT Press (2007)
- [13] Thomas Metzinger, *Der Ego Tunnel*, orig. Engelstalige ed. Basic Books, New York; Duitse vert. met toevoegingen Piper Taschenbücher (2009).
- [14] Oliver Sacks, *A leg to stand on*, Picador (1984)
- [15] Oliver Sacks, *Hallucinations*, Picador (2012)
- [16] D. Swaab, *We zijn ons brein*, Atlas Contact (2012)