

Wat is de rol van het concept van informatie in de dubbel-aspect theorie van informatie?

Patrick De Deyne

Masterproef voorgelegd tot het behalen van de graad van Master in de Wijsbegeerte door De Deyne
Patrick

Promotor: Karim Zahidi

Academiejaar 2022-2023

Aantal woorden hoofdtekst: 22646

Aantal woorden volledige tekst: 23227

De auteur en de promotor geven de toelating deze studie als geheel voor consultatie beschikbaar te stellen voor persoonlijk gebruik. Elk ander gebruik valt onder de beperkingen van het auteursrecht, in het bijzonder met betrekking tot de verplichting de bron uitdrukkelijk te vermelden bij het aanhalen van gegevens uit deze studie.

Dankwoord

Ik moet toegeven dat ik bij het werken aan deze thesis meerdere malen door de bomen het bos niet meer zag, niet in het minst omdat de theorieën die aan de onderzoeksvraag gekoppeld zijn zo complex en rijk zijn aan ideeën en concepten, dat het moeilijk is om een zekere mate van overzicht te krijgen. Er zijn maar weinig dingen die niet op één of andere manier te maken hebben met de dubbel-aspect theorie van informatie. Het viel me op dat ik veelal aan mijn kinderen, die ondanks hun jonge leeftijd echte filosofen in de dop zijn en voortdurend de meest interessante vragen op mij afvuren, mijn poging tot het formuleren van een antwoord op hun vragen bijna steeds kon beginnen met “eigenlijk heeft dat ook met mijn thesis te maken...”

Ik dank mijn promotor Karim Zahidi voor de interessante gesprekken, vele tips en feedback bij mijn pogingen om iets over dit moeilijke onderwerp te schrijven. Zijn suggesties waren van onschatbare waarde. Ik hoop dat het voorliggende resultaat enigszins aan zijn verwachtingen voldoet.

Ik dank ook Matteo Grasso, één van de onderzoekers die betrokken is bij de ontwikkeling van de *Integrated Information Theory*. Toen ik hem een vraag stelde in verband met één van zijn lezingen kreeg ik een uitgebreid antwoord met de link naar een aantal zeer interessante teksten.

Tenslotte is het gepast om ook mijn echtgenote Lore hier te bedanken voor haar ononderbroken steun gedurende de afgelopen jaren van studie. Als men de liefde die men voor iemand voelt onder meer kan meten aan de mate waarin je die ander steunt om zijn of haar dromen na te jagen, kan ik alleen al daaruit concluderen dat ik de liefde van haar leven ben (en dat geldt *vice versa!*).

Inhoudsopgave

Dankwoord	IV
Inhoudsopgave	V
Lijst met afkortingen.....	VII
Lijst van figuren	VII
Wat is de rol van het concept van informatie in de dubbel-aspect theorie van informatie?.....	1
1 Inleiding	1
1.1 Situering van het onderwerp.....	1
1.2 Onderzoeksvraag	2
1.3 Synopsis	3
2 Achtergrond van de dubbel-aspect theorie van informatie	3
2.1 <i>The hard problem of consciousness</i>	3
2.2 Bewustzijn en gewaar zijn	4
2.3 Fenomenale, niet-fysische aspecten van de werkelijkheid	6
2.4 Enkele argumenten voor ervaring als niet-fysisch aspect van de werkelijkheid	8
2.4.1 Het argument vanuit verklaring.....	8
2.4.2 Het voorstelbaarheid argument	9
2.4.3 Het kennis argument	10
2.5 De ‘Grote Kloof’ in de studie van bewustzijn	11
3 De dubbel-aspect theorie van informatie	11
3.1 Eigenschapsdualisme.....	11
3.2 Psychofysische wetten.....	13
3.2.1 Het coherentie principe.....	14
3.2.2 Het structurele coherentie principe	15
3.2.3 Het <i>organisationele invariantie principe</i>	16
3.3 Een <i>theory of everything</i> ? Of nog verder?.....	19
3.4 Het concept van informatie.....	23
3.4.1 Informatie	23
3.4.2 Shannon informatie	24

3.4.3	Informatie volgens Chalmers.....	25
3.4.4	Informatieruimtes.....	25
3.4.5	Fysisch gerealiseerde informatie	27
3.4.6	Fenomenaal gerealiseerde informatie	31
3.5	Het dubbel-aspect principe van informatie.....	32
3.6	<i>It from bit</i>	35
3.7	Russelliaans monisme	37
3.8	De rol van informatie bij DAT.....	39
3.9	Gelijkenissen en verschillen tussen DAT en monisme	41
3.10	Ontisch structureel realisme	43
3.11	Is dit epifenomenalisme?	45
3.12	Panpsychisme	47
3.13	Het combinatieprobleem	51
4	<i>Integrated Information Theory</i>	53
4.1	De methodologie van IIT	53
4.2	Geïntegreerde informatie: een kwantitatieve benadering van bewustzijn.....	54
4.3	Qualia in IIT.....	57
4.4	Het concept van informatie bij IIT	58
4.5	De metafysica achter IIT	59
5	Conclusie	62
	Bibliografie.....	64

Lijst met afkortingen

DAT dubbel-aspect theorie van informatie

IIT *integrated information theory*

ISR informationeel structureel realisme

MUH mathematische universum hypothese

OSR ontisch structureel realisme

TOE *theory of everything*

Lijst van figuren

Figuur 1: eigenschapsdualisme	12
Figuur 2: communicatiesysteem volgens Shannon	24

Wat is de rol van het concept van informatie in de dubbel-aspect theorie van informatie?

1 Inleiding

1.1 Situering van het onderwerp

Het *mind-body* probleem gaat over de vraag wat de relatie is tussen het mentale en het fysische, of de relatie tussen mentale en fysische eigenschappen. Materialistische visies stellen dat mentale toestanden niets anders zijn dan fysische toestanden. Volgens idealisten geldt eerder het omgekeerde: fysische toestanden zijn mentaal, en de wereld is het intersubjectieve product van onze collectieve ervaring. Volgens het dualisme zijn zowel het mentale als het fysische niet te reduceren tot één van beide. Er bestaan verschillende versies van dualisme. Volgens Plato behoorde het intellect niet tot de materiële wereld maar tot de wereld van de eeuwige vormen, omdat het in staat was die vormen te begrijpen. Het dualisme van Plato was niet enkel een theorie van de geest maar tevens een metafysische theorie. Volgens het Cartesiaans dualisme is de wereld opgebouwd uit de *res cogitans*, de substantie die denkt, en de *res extensa*, de substantie die uitgebreid is in de ruimte.¹ Sinds de twintigste eeuw is het materialisme de dominerende visie op het mentale, waarbij er vanuit wordt gegaan dat het bewustzijn niets ander is dan een toestand of een functie van het brein (fysicalisme of functionalisme).²

David Chalmers stelde eind jaren 90' in het artikel "Facing up to the problem of consciousness" en in het daaropvolgende boek *The Conscious Mind* zijn dubbel-aspecttheorie van informatie voor, een vorm van eigenschapsdualisme.^{3 4} Een vroege versie van eigenschapsdualisme werd verdedigd door Spinoza, met zijn visie dat materie en geest verschillende aspecten waren van iets diepers in de natuurlijke werkelijkheid.⁵ De dubbel-aspect theorie van Chalmers is een zeer tentatieve en speculatieve theorie die voortbouwt op de al even speculatieve "it from bit" theorie van John Wheeler en Edward Fredkin,

¹ Howard Robinson, "Dualism," *The Stanford Encyclopedia of Philosophy*, 2023, 1.2 History of dualism.

² Max Velmans, "The relation of consciousness to the material world," *Journal of Consciousness Studies* 2, No.3 (1995): 256.

³ David Chalmers, "Facing up to the problem of consciousness," *Journal of Consciousness Studies* 2, No.3 (1995).

⁴ David Chalmers, *The Conscious Mind* (Oxford: Oxford University Press, 1996).

⁵ Baruch Spinoza, (1677), *The Ethics*. Reprinted in *The Ethics of Benedict Spinoza* (New York: Van Nostrand, 1876).

waarin informatie als het fundament wordt gezien waaruit onze fysische wereld is opgebouwd.^{6 7} Volgens Chalmers zijn de fenomenale aspecten in ons leven fundamentele, niet-fysische eigenschappen van onze wereld. Fysische en niet-fysische eigenschappen zijn terug te voeren tot verschillende ‘aspecten’ van informatie. In zijn meest radicale vorm kan men de dubbel-aspect theorie zien als de visie dat de wereld uit niets anders bestaat dan de realisering van elementaire verschillen. In die vorm is deze theorie niet enkel een theorie over bewustzijn, maar, net als bij Plato, een metafysische theorie over de aard van de wereld, een *theory of everything*.

1.2 Onderzoeksvraag

De dubbel-aspect theorie van informatie (verder aangeduid als ‘DAT’) heeft heel wat reacties gekregen en blijft tot op heden zeer controversieel. De meeste kritieken zijn echter gericht op de argumenten die Chalmers gebruikt om zijn claim te onderbouwen dat bewustzijn niet-fysisch is. Het doel van deze scriptie is dan ook eerder het concept van informatie, dat de kern is van de dubbel-aspect theorie, kritisch te onderzoeken. Chalmers beschouwde zijn theorie eerder als een proto-theorie, waardoor verschillende belangrijke aspecten eerder schetsmatig werden uitgewerkt. Daardoor roept de theorie al onmiddellijk een aantal vragen en problemen op, onder meer of het hierin gebruikte concept van informatie überhaupt wel coherent is. Chalmers kwam in latere werken herhaaldelijk terug op DAT, onder meer in zijn meer recente boek *The Character of Consciousness* (2010), waar hij opnieuw het zeer tentatieve karakter van zijn eerder geformuleerde theorie benadrukt. Daarin stelt hij zelf dat “(...) reflectie op wat er juist plausibel en implausibel aan [de dubbel-aspect theorie] is en waar het werkt en waar het faalt kan alleen maar leiden tot een betere theorie”.⁸ DAT dateert reeds van de jaren ‘90 van de vorige eeuw, maar wordt beschouwd als een voorloper van de recentere *Integrated Information Theory* (IIT) van Giulio Tononi, die de brug slaat tussen een filosofische theorie zoals DAT en de neurowetenschappen. Als deelvraag van de onderzoeksvraag zal ik dan ook onderzoeken in hoeverre het concept van informatie dat gebruikt wordt in de verschillende versies van IIT al dan niet overeenkomt met het concept van informatie in DAT en eventueel kan gezien worden als een verdere ontwikkeling en invulling ervan. We zullen zien dat we vertrekkende van deze onderzoeksvraag heel wat verschillende concepten en theorieën zullen ontmoeten, waarbij er hier in deze tekst geen ruimte is om die allemaal in detail te bespreken. Daarom zal ik mij steeds trachten te beperken tot de meest relevante aspecten in verband met het concept van informatie en bewustzijn.

⁶ John Wheeler, “Information, physics, quantum: The search for links,” in *Complexity, Entropy, and the Physics of Information*, ed. W. H. Zurek (Redwood City, Calif.: Addison-Wesley, 1990).

⁷ Edward Fredkin, “Digital mechanics,” *Physica D*45(1990):254-70.

⁸ David Chalmers, *The Character of Consciousness* (Oxford: Oxford University Press, 2010), 28.

1.3 Synopsis

In hoofdstuk 2 geef ik een beknopte schets van de achtergrond waartegen DAT geformuleerd werd, zoals onder meer *the hard problem of consciousness* en een opsomming van de voornaamste argumenten die Chalmers gaf voor de noodzaak van een ruimere theorie voor bewustzijn zoals DAT, samen met enkele belangrijke kritieken die op die argumenten werden geformuleerd. In hoofdstuk 3 bespreek ik vervolgens de verschillende onderdelen van DAT, waarin ik focus op het concept van informatie en de rol die informatie speelt in DAT, samen met de vragen, problemen en kritieken die deze oproepen. DAT is een vrij complexe theorie, waarbij elk onderdeel steeds verder bouwt op het vorige onderdeel. In mijn bespreking zal ik dan ook min of meer dezelfde volgorde aanhouden waarin Chalmers zijn theorie uit de doeken doet in *The Conscious Mind*, aangevuld met andere, ook meer recente, bronnen. Tenslotte bespreek ik in hoofdstuk 4 de recentere *Integrated Information Theory* (IIT) en ga ik na in hoeverre deze theorie en het concept van informatie dat ze hanteert kan gezien worden als een verdere ontwikkeling van DAT en het concept van informatie bij Chalmers.

2 Achtergrond van de dubbel-aspect theorie van informatie

2.1 *The hard problem of consciousness*

Chalmers is bekend om zijn introductie van het ‘moeilijke probleem van bewustzijn’ in het artikel “Facing up to the problem of consciousness”.^{9 10} Met zijn formulering van het moeilijke probleem van bewustzijn verdeelt Chalmers het mind-body probleem in twee: een makkelijk en een moeilijk probleem. Het *makkelijke* probleem gaat over hoe fysische systemen functies (zoals bestudeerd in de psychologie) kunnen uitoefenen en zo in levende organismen passende gedragingen veroorzaken als reactie op een omgeving. Het makkelijke probleem bestaat in feite uit vele makkelijke problemen die te maken hebben met cognitieve vaardigheden zoals het onderscheiden, categoriseren en reageren op stimulansen uit de omgeving. Volgens Chalmers roept dit wel enkele filosofische vragen op maar gaat dit vooral over een grote hoeveelheid technische vragen die - ten minste in principe - door de wetenschappen die fysische systemen bestuderen beantwoord kunnen worden. Het makkelijke probleem van bewustzijn geeft geen aanleiding tot diepe metafysische vragen. In tegenstelling daarmee is er het moeilijke probleem van *hoe* en *waarom* het verwerken van informatie door ons brein gepaard gaat met innerlijke ervaringen, met fenomenale aspecten zoals het ervaren van jaloezie of de

⁹ David Chalmers, “Facing up to the problem of consciousness,”.

¹⁰ Chalmers spreekt van *the hard problem of consciousness*. Ik gebruik het woord ‘moeilijk’ als vertaling van ‘hard’ omdat de alternatieve vertaling als ‘het harde probleem’ in het Nederlands wat vreemd klinkt.

reuk van een bloesem, die ook qualia worden genoemd.¹¹ Het moeilijke probleem van bewustzijn gaat over dit subjectieve aspect waarmee het verwerken van informatie door ons brein gepaard gaat.^{12 13} Het werkelijk bestaan van het moeilijke probleem is echter omstreden en wordt betwist door onder meer Daniel Dennett, volgens wie het moeilijke probleem van bewustzijn in werkelijkheid niet bestaat, of eerder bestaat uit een groot aantal makkelijke problemen. Volgens deze visie, die illusionisme of eliminativisme wordt genoemd, worden wij door de werking van ons brein verleid om te denken dat er ‘qualia’ bestaan, in de zin van niet-relatieve intrinsieke kwaliteiten van ervaring, terwijl deze in werkelijkheid net zo min bestaan als de Kerstman. De “speciale kwaliteiten” die we toekennen aan onze ervaringen (beschouwd als processen in ons brein) moeten we volgens Dennett eerder zien als een *representatie* van kwaliteiten van dingen in de buitenwereld.¹⁴ Dennett ontkent het bestaan van fenomenaal bewustzijn niet, maar voor hem wordt het bewustzijn volledig uitgeput door een functionele beschrijving ervan.

2.2 Bewustzijn en gewaar zijn

De verschillende concepten die Chalmers gebruikt voor het mentale in zijn beschrijving van DAT volgen rechtstreeks uit zijn standpunt dat het moeilijke probleem van bewustzijn echt bestaat en zijn dus niet theoretisch neutraal. Zo gebruikt hij in verband met het moeilijke probleem de term ‘ervaring’, die voor hem gelijk staat aan ‘quale’, ‘bewuste ervaring’, ‘sensatie’, ‘fenomenaal bewustzijn’ of kortweg ‘bewustzijn’.¹⁵ Dit is volgens hem te onderscheiden van het concept ‘gewaar zijn’, die hij wil reserveren voor de ‘makkelijke’ problemen.¹⁶ Het gewaar zijn omschrijft Chalmers als een functionele notie, namelijk de toestand waarin we toegang hebben tot bepaalde informatie, die we kunnen gebruiken om ons gedrag te sturen.^{17 18} Chalmers gebruikt als synoniem voor gewaar zijn ook nog ‘globale

¹¹ De term qualia (enkelvoud quale) is afkomstig van C.S. Peirce, zie Charles Sanders Peirce, “Lowell Lecture,” (ix) in *Writings of Charles S. Peirce: A Chronological Edition*, ed. M. H. Fisch (Bloomington, 1866) (Indiana: Indiana University Press, 1982), para 223.

¹² David Chalmers, “Facing up to the problem of consciousness,” 201.

¹³ Chalmers, *The Conscious Mind*, 24-5.

¹⁴ Daniel Dennett, “A History of Qualia,” *Topoi* 39:5–12 (2020): 7,11.

¹⁵ Dit gebruik van de term ‘bewustzijn’ is niet geheel correct: bewustzijn is eerder de aanwezigheid van een fenomenale inhoud, men is bewust van iets, een ervaring. Ik zal in deze tekst dan ook onderscheid maken tussen bewustzijn en ervaringen. Zie Max Velmans, “How to define Consciousness and how not to define Consciousness,” *Journal of Consciousness Studies* 16 (5) (2009):141-142.

¹⁶ In het artikel “Facing up to the problem of consciousness” stelt Chalmers dat als een dergelijk gebruik van terminologie algemener zou worden, dit minder aanleiding zou geven tot verwarring. We moeten dit uiteraard opnieuw zien in het licht van de theoretische keuzes die Chalmers hier reeds maakt. Zie David Chalmers, “Facing up to the problem of consciousness,” *Journal of Consciousness Studies* 2, No.3 (1995): 201-202.

¹⁷ Chalmers, *The Conscious Mind*, 28.

¹⁸ Ik gebruik de benaming ‘gewaar zijn’ als vertaling van ‘awareness’ bij Chalmers, om deze te onderscheiden van het meer voor de hand liggende ‘gewaarwording’, die in de psychologie echter veelal gebruikt wordt als synoniem

beschikbaarheid', omdat de informatie waar we toegang tot hebben onmiddellijk beschikbaar is voor het sturen van een groot deel van ons scala aan gedragspatronen.^{19 20} Gewaar zijn is nauw verbonden met het concept perceptie, dat het proces aanduidt waarbij cognitieve systemen, zoals ons brein, gevoelig zijn voor stimulansen uit onze omgeving en de informatie die deze stimulansen bevat cognitief verwerkt. Perceptie moet hierbij onderscheiden worden van haar fenomenale tegenhanger sensatie. Het onderscheid tussen deze concepten is echter niet absoluut: de psychologische aspecten van het mentale zijn doorgaans vergezeld van een zeker fenomenaal aspect, terwijl de fenomenale aspecten van het mentale ook gepaard gaan met een zeker gewaar zijn.²¹ Als ik bijvoorbeeld de ervaring heb van het zien van een groene glijbaan in de tuin, is er in mijn brein ook sprake van het verwerken van informatie, van perceptie, en ben ik ook gewaar van het aanwezig zijn van een groene glijbaan in de tuin. Gewaar zijn en ervaring zijn dus intiem met elkaar verbonden (deze correlatie zal hieronder worden behandeld als één van de 'psychofysische wetten' die Chalmers meent af te leiden). Mentale concepten kunnen hierdoor steeds in een zekere mate zowel fenomenaal als psychologisch geanalyseerd worden. Zo kan men in verband met perceptie ook een zekere fenomenale component betrekken, waarbij het dan gaat over de bewuste ervaring van wat er gepercipieerd wordt. Omgekeerd kan men bij het concept sensatie ook een psychologische component betrekken. Niettemin is perceptie eerder een psychologisch concept en sensatie een meer fenomenaal concept, aangezien we ons iets kunnen voorstellen bij onbewuste perceptie, maar niet bij onbewuste sensatie. Zowel in de gewone omgangstaal als in de filosofische en wetenschappelijke literatuur worden de fenomenale en cognitieve aspecten van het mentale echter veelal met elkaar verward en worden termen zoals bewustzijn, gewaar zijn en perceptie door elkaar gebruikt.²² De oorzaak voor deze verwarring moet deels gezocht worden in het feit dat we geen aparte taal hebben voor de fenomenale aspecten van ons leven, die nu eenmaal moeilijk onder woorden te brengen zijn. Het is bijvoorbeeld heel moeilijk of onmogelijk om aan iemand anders of aan onszelf te beschrijven wat onze ervaring van het proeven van iets zoets inhoudt. De taal die we gebruiken voor het beschrijven van onze ervaringen is dan ook afgeleid van de

voor sensatie. Zie voor dat laatste gebruik bijvoorbeeld Marc Brysbaert, *Psychologie* (Gent: Owl Press, Borgerhoff & Lamberigts nv, 2018), 94.

¹⁹ Chalmers, *The Conscious Mind*, 225.

²⁰ David Chalmers, "Moving Forward on the Problem of Consciousness," *Journal of Consciousness Studies* 4 (1) (1997), 3.5 My psychophysical laws.

²¹ Chalmers bespreekt enkele gevallen waarbij het samen gaan van gewaar zijn en ervaring kan worden worden betwist, zoals bij subliminale perceptie, waar er geen sprake is van ervaring. Hij beschouwt deze mogelijke uitzonderingen echter eerder als randgevallen en merkt op dat er bij subliminale perceptie misschien maar in minimale mate sprake is van perceptie. Zie Chalmers, *The Conscious Mind*, 18.

²² Hierbij moet opgemerkt worden dat ook Chalmers hier niet aan ontsnapt. Zo maakt hij in *The Conscious Mind* zelf ook — weliswaar in beperkte mate — gebruik van de termen "psychologisch bewustzijn" en "fenomenaal bewustzijn", een terminologie die ik hier niet volg. Zie Chalmers, *The Conscious Mind*, 26, 31.

taal waarmee we de niet-fenomenale aspecten van ons leven beschrijven.²³ De ambiguïteit van het concept bewustzijn heeft volgens Chalmers als gevolg dat de titels in de wetenschappelijke en filosofische literatuur veelal doen vermoeden dat bewustzijn het onderwerp is dat behandeld zal worden, terwijl uit de tekst vervolgens blijkt dat deze eerder handelt over de psychologische aspecten van het mentale. Op die manier wordt het moeilijke probleem van bewustzijn niet aangepakt, maar veelal omzeild.²⁴ Nog volgens Chalmers komt dit neer op een assimilatie van het fenomenale door het psychologische.²⁵

Iemand zoals Dennett, die ontkent dat het moeilijke probleem van bewustzijn bestaat, zou echter niet akkoord zijn met de hierboven besproken terminologie en concepten. Er zijn nog andere auteurs die kritiek hebben op het onderscheid dat Chalmers maakt tussen bewustzijn en gewaar zijn. Zo is volgens Benjamin Libet gewaar zijn wel degelijk een subjectief fenomeen. Verder merkt hij op dat heel wat van de zaken die Chalmers makkelijke problemen noemt, kunnen doorgaan zonder gewaar zijn. Volgens Libet moet men eerder een onderscheid maken tussen detectie en gewaar zijn.²⁶ Max Velmans is dan weer akkoord met het van elkaar onderscheiden van bewustzijn in de zin van ervaring en het functionele verwerken van informatie (wat Chalmers aanduidt met gewaar zijn), maar merkt op dat in het gewone taalgebruik de term gewaar zijn veelal als uitwisselbaar wordt gezien met de termen 'bewuste gewaarwording' en 'bewuste ervaring'. Daardoor kan volgens Velmans impliciet de indruk gewekt worden dat het verwerken van informatie *noodzakelijk* gepaard gaat met een zekere mate van bewuste ervaring, ook als dit verwerken van informatie in andere systemen dan het brein gebeurt.²⁷ Ik zal hieronder grotendeels dezelfde terminologie gebruiken als Chalmers, maar het is hiermee duidelijk dat er een zekere controverse is rond deze termen en dat er door het gebruik van bepaalde termen reeds een keuze tussen alternatieve visies wordt gemaakt die niet vanzelfsprekend is.

2.3 Fenomenale, niet-fysische aspecten van de werkelijkheid

Het oplossen van het moeilijke probleem van bewustzijn ligt volgens Chalmers in het accepteren van bewustzijn als een niet-fysisch en fundamenteel aspect van de natuurlijke werkelijkheid. Hierdoor gaat Chalmers in tegen het fysicalisme. Hoewel de termen 'fysicalisme' en 'materialisme' een verschillende historische oorsprong hebben en gebruikt worden om verschillende doctrines aan te duiden, gebruikt

²³ Chalmers, *The Conscious Mind*, 22.

²⁴ Chalmers, *The Conscious Mind*, 30-31.

²⁵ Chalmers, *The Conscious Mind*, 15.

²⁶ Benjamin Libet, "Solutions to the hard problem of consciousness," *Journal of Consciousness Studies* 3, No.1 (1996): 33.

²⁷ Max Velmans, "The relation of consciousness to the material world," *Journal of Consciousness Studies* 2, No.3 (1995): 258-259.

Chalmers beide termen als verwisselbaar voor de metafysische theorie die stelt dat alle feiten over de wereld worden uitgeput door fysische feiten over de wereld, of dat ieder feit over de wereld volgt uit fysische feiten. Volgens materialisten is ieder object of eigenschap fysisch of reduceerbaar tot fysische objecten en feiten.^{28 29 30}

Binnen het materialisme of fysicalisme kan men nog verschillende posities onderscheiden ten opzichte van bewustzijn. Chalmers onderscheidt twee versies van materialisme: “type A” en “type B” materialisme. Type A materialisten ontkennen het bestaan van het moeilijke probleem van bewustzijn en stellen dat alleen de functies, die een rol spelen bij het mentale, moeten verklaard worden. Materialisme neemt soms de vorm aan van het logisch behaviorisme of het analytisch functionalisme, die beide stellen dat bewustzijn bestaat, maar dat dit laatste volledig beschreven kan worden via gedragingen, respectievelijk functies. Zij maken nog steeds gebruik van de term bewustzijn, maar dan slechts in de functionele zin. Daarnaast is er het eliminativisme (of illusionisme), waar Daniel Dennett een vertegenwoordiger van is, dat stelt dat de term ‘bewustzijn’ enkel gebruikt dient te worden in de niet-functionele zin. Maar aangezien het mentale in die laatste zin niet bestaat, stellen zij dat het bewustzijn niet bestaat en de term dus best vermeden wordt. Volgens type B materialisten, zoals Valerie Gray Hardcastle, is er geen sprake van een ontologische ‘kloof’ tussen het fysische en het fenomenale, maar is er wel degelijk een epistemologische ‘kloof’ tussen beide domeinen. Voor hen zijn bewustzijn en de uitoefening van functies conceptueel verschillend. Dit *a priori* verschil wordt opgelost door een *a posteriori* identiteit tussen bewustzijn en bepaalde fysische of functionele eigenschappen, zoals ook water en H₂O verschillende concepten zijn, maar naar hetzelfde ding in de natuur verwijzen.³¹

32

Chalmers legt noch in “Facing up to the problem of consciousness”, noch in *The Conscious Mind* uit wat hij zelf precies als fysisch beschouwt. We zullen verder zien dat Chalmers wel een uitgesproken heel eigen visie heeft op wat de aard van het fysische is.

Bij dit alles moet opgemerkt worden dat, niettegenstaande bewustzijn volgens Chalmers niet-fysisch is, hij bewustzijn wel beschouwt als een natuurlijk fenomeen, waarbij natuurwetten geldig zijn.³³

Chalmers is van mening dat hij met zijn voorstel het wetenschappelijk wereldbeeld niet omvergooid,

²⁸ Daniel Stoljar, “Physicalism,” *The Stanford Encyclopedia of Philosophy*, 2023, 1.1 Terminology.

²⁹ Chalmers, *The Conscious Mind*, 41, 124.

³⁰ Dit kan nog preciezer geformuleerd worden door te zeggen dat materialisme stelt dat alle *positieve* feiten volgen uit fysische feiten. Negatieve existentiële feiten zoals “eenhoorns bestaan niet” volgen strikt logisch gezien niet uit het fysische, maar dit is compatibel met het materialisme. Zie Chalmers, *The Conscious Mind*, 41.

³¹ David Chalmers, “Moving Forward on the Problem of Consciousness,” 2.1 Deflationary analogies.

³² David Chalmers, “Consciousness and Its Place in Nature,” in *Philosophy of Mind. Classical and Contemporary Readings*, ed. David Chalmers (Oxford/New York: Oxford University Press, 2002): 251-5.

³³ Chalmers, *The Conscious Mind*, xiii.

maar eerder verruimd.³⁴ Hij kwam naar eigen zeggen zelf met tegenzin tot de conclusie dat bewustzijn niet-fysisch is:

Temperamentally, I am strongly inclined toward materialist reductive explanation, and I have no strong spiritual or religious inclinations. For a number of years, I hoped for a materialist theory; when I gave up on this hope, it was quite reluctantly.³⁵

2.4 Enkele argumenten voor ervaring als niet-fysisch aspect van de werkelijkheid

Chalmers bespreekt in *The Conscious Mind* uitvoerig enkele argumenten voor zijn stelling dat bewustzijn een niet-fysisch aspect van de werkelijkheid is. Hieronder bespreek ik bondig een aantal van die argumenten tegen het materialisme en enkele kritieken hierop.

2.4.1 Het argument vanuit verklaring

Het *argument vanuit verklaring* stelt dat fysische verklaringen enkel betrekking hebben op structuren en functies, waarbij die eerste spatiotemporele structuren en de tweede causale rollen zijn in het tot stand brengen van het gedrag van een systeem. Dit hangt nauw samen met de wetenschappelijke werkwijze. Wetenschappelijke theorieën steunen op het verzamelen van objectieve, derde persoon gegevens: functies en structuren zijn alles wat we kunnen observeren bij een extern object. Maar structuren en functies zijn volgens Chalmers onvoldoende om bewustzijn te verklaren.³⁶ Iets wat niet fysisch verklaard kan worden is zelf niet fysisch, waardoor volgt dat bewustzijn iets niet-fysisch is.³⁷ Materialisten kunnen hiertegen inbrengen dat we nog niet alles weten over bewustzijn maar dat de situatie vergelijkbaar is met andere problemen die uiteindelijk toch volledig via functies en structuren konden verklaard worden. Zo schetst Paul Churchland een denkbeeldige filosoof uit 19^e eeuw, die stelt dat er een moeilijk probleem van 'licht' bestaat, terwijl later bleek dat functies en structuren wel degelijk volstonden om het fenomeen licht te begrijpen.³⁸

³⁴ Chalmers, *The Conscious Mind*, xiv.

³⁵ Chalmers, *The Conscious Mind*, xiv.

³⁶ Dit wil echter niet zeggen dat ervaringen geen functie kunnen hebben, zo denkt Chalmers dat het goed mogelijk is dat ervaringen een belangrijke functie hebben met betrekking tot onze cognitie, zie Chalmers, "Facing up to the problem of consciousness," 203.

³⁷ David Chalmers, "Consciousness and Its Place in Nature," in *Philosophy of Mind. Classical and Contemporary Readings*, ed. David Chalmers (Oxford/New York: Oxford University Press, 2002): 248-9.

³⁸ Paul Churchland, "The rediscovery of light," *Journal of Philosophy* 93 (1996): 219.

2.4.2 Het voorstelbaarheid argument

Het *voorstelbaarheid argument* vertrekt vanuit het feit dat we ons een wereld kunnen voorstellen waarin ‘filosofische zombies’ bestaan, die fysisch identiek zijn aan mensen uit onze wereld, maar geen fenomenaal bewustzijn hebben. Uit die voorstelbaarheid volgt volgens voorstanders van het argument dat een dergelijke wereld, die fysisch identiek is aan de onze maar waarin geen bewustzijn voorkomt, ook metafysisch mogelijk is. Daaruit wordt afgeleid dat bewustzijn een verder, niet-fysisch bestanddeel is van onze wereld.³⁹ Dit argument werd in verschillende versies door andere auteurs geformuleerd. Een voorloper ervan werd geformuleerd door Descartes, die zich voorstelde dat onze geest zonder ons lichaam kan bestaan.⁴⁰ De zwakheid van dit argument is dat uit het voorstelbaar zijn van iets misschien niet volgt dat het ook metafysisch mogelijk is. Wat we ons voorstellen kan bij nader inzien toch incoherent blijken te zijn. Daniel Dennett denkt dat de mogelijkheid van filosofische zombies een voorbeeld is van zich iets verkeerd inbeelden.⁴¹ Jaegwon Kim dacht dat filosofische zombies conceptueel niet mogelijk zijn, omdat die leiden tot een incoherentie tussen de fenomenale oordelen die dergelijke zombies zouden uitspreken en het veronderstelde onbewust zijn van dergelijke wezens. Volgens Kim zouden dergelijke wezens beschikken over bewustzijn, maar zou wel het kwalitatief karakter daarvan onbekend zijn.⁴² In de meer gedetailleerde uitwerking van dit argument speelt het concept superveniëntie een grote rol. Superveniëntie is de relatie tussen twee sets van eigenschappen: B-eigenschappen die zich op een hoger niveau bevinden en A-eigenschappen die zich op een lager, meer fundamenteel niveau bevinden. B-eigenschappen superveniëren op A-eigenschappen als er geen mogelijke situaties zijn waarin de A-eigenschappen dezelfde zijn terwijl de B-eigenschappen verschillend zijn. Er wordt onderscheid gemaakt tussen logische, natuurlijke en metafysische superveniëntie. B-eigenschappen zijn logisch superveniënt als er geen twee logisch mogelijke situaties zijn waarin de A-eigenschappen dezelfde zijn, terwijl de B-eigenschappen verschillend zijn. Logische mogelijkheid is mogelijkheid in de meest ruime zin, waarbij men geen rekening hoeft te houden met de natuurwetten in onze wereld, maar wel met logische beperkingen, zoals het niet logisch mogelijk zijn van een gehuwde vrijgezel. Natuurlijke superveniëntie is een zwakkere vorm van superveniëntie, wanneer twee sets van eigenschappen in de actuele, natuurlijke wereld wetmatig met elkaar gecorreleerd zijn. Er is echter ook de mogelijkheid dat er een extra modaliteit bestaat tussen natuurlijk en logisch mogelijk: metafysisch mogelijk. Dit laatste zou betekenen dat er een sterke metafysische

³⁹ Chalmers “Consciousness and Its Place in Nature,” 249.

⁴⁰ Chalmers “Consciousness and Its Place in Nature,” 268-9, voetnoot 6.

⁴¹ Daniel Dennett, “The Unimagined Preposterousness of Zombies,” *Journal of Consciousness Studies* 2 (1995): 325.

⁴² Jaegwon Kim, *Physicalism, or something near enough*, Princeton Monographs in Philosophy (Princeton: Princeton University Press, 2005), 169, voetnoot 16.

noodzakelijkheid is waardoor de set van de ‘werkelijk’ mogelijke werelden toch een echte subset zou zijn van de logisch mogelijke werelden. Op die manier zouden bijvoorbeeld zombiewerelden logisch mogelijk kunnen zijn maar niet metafysisch mogelijk.^{43 44} De superveniëntie van het mentale op het fysische is onderwerp van hevige discussie.⁴⁵

Chalmers denkt dat bewustzijn slechts natuurlijk superveniënt is op het fysische. Als mijn fysische structuur zou gekopieerd worden in een ander wezen in deze wereld, zou dit wezen dus dezelfde ervaring ondervinden als mijzelf. Hierdoor zouden filosofische zombies wel logisch mogelijk zijn. Verder denkt Chalmers dat er geen enkele reden is om te geloven in het bestaan van een metafysische noodzakelijkheid omdat dit volgens hem een arbitraire en onverklaarbare beperkende voorwaarde zou opleggen aan de ruimte van mogelijke werelden.⁴⁶

2.4.3 Het kennis argument

Het *kennis argument* vertrekt vanuit een bekend gedachtenexperiment, onder meer geformuleerd door Frank Jackson, waarin Mary een experte in kleurenzicht is, die alle kennis bezit die er te weten is van de fysische processen die betrokken zijn bij kleurwaarneming. Ze heeft echter zelf nooit kleur waargenomen, en als ze uiteindelijk voor de eerste keer de kleur rood ziet lijkt het, ondanks al haar kennis van de betrokken fysische processen, alsof ze een belangrijk nieuw feit leert: *hoe het is* om de kleur rood te zien. Aangezien dit feit over bewustzijn niet afleidbaar is uit feiten over fysische processen, volgt uit het gedachtenexperiment dat materialisme onjuist is.^{47 48} Kritieken op dit argument werden geformuleerd door Lewis en Nemirow, die ontkennen dat het hier over een nieuw feit gaat. Voor hen ontbreekt het Mary eerder aan een *vaardigheid*, namelijk het kunnen herkennen van dingen die rood zijn.^{49 50}

⁴³ Chalmers, *The Conscious mind*, 33.

⁴⁴ Het concept superveniëntie is relatief recent van origine. De introductie van de term in de filosofische literatuur wordt toegekend aan R.M. Hare in de jaren '50.

⁴⁵ Brian McLaughlin & Karen Bennett, "Supervenience," *The Stanford Encyclopedia of Philosophy*, 2021, 3.1 The Modal Force of the Supervenience Relation.

⁴⁶ Chalmers, *The Character of Consciousness*, 185.

⁴⁷ Chalmers, "Consciousness and Its Place in Nature," 269, voetnoot 7.

⁴⁸ Frank Jackson, "Epiphenomenal qualia," *Philosophical Quarterly* 32 (1982): 130.

⁴⁹ David Lewis, "What experience teaches," in *Mind and Cognition*, ed. W. Lycan (Oxford: Blackwell, 1990), 77-103.

⁵⁰ Laurence Nemirow, "Physicalism and the cognitive role of acquaintance," in *Mind and Cognition*, ed. W. Lycan (Oxford: Blackwell, 1990).

2.5 De ‘Grote Kloof’ in de studie van bewustzijn

Alle drie bovenstaande argumenten zijn dus controversieel, en de discussie erover is nog steeds gaande. Chalmers spreekt in dit verband van de “Grote Kloof” in de studie van bewustzijn. Het formuleren van argumenten over deze kloof heen is moeilijk, omdat beide partijen zich volgens hem voor een groot deel beroepen op verschillende intuïties en de discussie verzand is tot een patstelling tussen twee zijden die krampachtig vasthouden aan die intuïties.⁵¹ Chalmers vindt het weinig zinvol om steeds maar opnieuw verdere argumenten te bedenken die de tegenstelling tussen materialisme en zijn eigen positie verder uitvechten. Hij denkt dat de beste manier om vooruitgang te boeken ligt in het verder uitwerken van beide onderzoeksprogramma’s. Ook Valerie Gray Hardcastle vergelijkt de discussie tussen “toegewijde” materialisten (waar ze zichzelf tot rekent) en “skeptici” zoals Chalmers met iets zoals een loopgravenoorlog tussen twee kampen, die elk moeilijk te “bekeren” zijn tot het andere kamp.⁵² Hoewel ik persoonlijk eerder overtuigd ben door het argument vanuit verklaring, spreek ik mij hier verder niet uit over de plausibiliteit van de bovenstaande argumenten en tegenargumenten. Ik denk overigens dat het goed mogelijk is dat er nog nieuwe argumenten en tegenargumenten zullen worden geformuleerd die de strijd misschien ooit zullen beslechten in het voordeel van één van beide kampen. De argumenten hierboven zijn echter vrij prominent in de *philosophy of mind* en in die zin onontbeerlijk als achtergrond van DAT.

3 De dubbel-aspect theorie van informatie

3.1 Eigenschapsdualisme

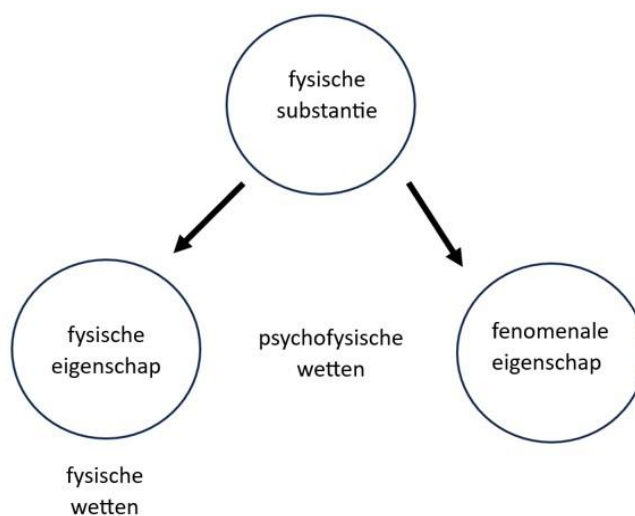
Bewustzijn is dus volgens Chalmers iets dat niet fysisch is. In de wereld zijn er fenomenale eigenschappen die verschillend zijn van fysische eigenschappen.⁵³ Verderop zullen we de precieze ontologie die Chalmers voorstelt meer in detail zien. Maar we kunnen nu reeds zeggen dat wat Chalmers verdedigt niet het cartesiaans substantiedualisme is, volgens welk er een aparte mentale substantie bestaat, verschillend van fysische substantie, die een invloed uitoefent op fysische processen. Hij is eerder voorstander van een *eigenschapsdualisme*, waarbij fenomenologische eigenschappen ontologisch onafhankelijk zijn van fysische eigenschappen, alhoewel fenomenologische eigenschappen wel wetmatig verbonden zijn met fysische eigenschappen. In *The Conscious Mind* stelt Chalmers dat we zijn visie niet mogen verwarren met een zwakkere vorm van eigenschapsdualisme,

⁵¹ Chalmers, *The Conscious Mind*, xiii.

⁵² Valerie Gray Hardcastle, “The why of consciousness: a non-issue for materialists,” *Journal of Consciousness Studies* 3, No.1 (1996): 13.

⁵³ Chalmers, *The Conscious Mind*, 123.

waarbij eigenschappen zoals bijvoorbeeld biologische fitness te reduceren zijn tot microfysische eigenschappen. Het verband tussen fysische en fenomenale eigenschappen is zeer subtiel: bewustzijn *ontstaat* weliswaar uit een fysische basis, maar *volgt* niet uit die fysische basis. Fenomenale eigenschappen zijn niet logisch superveniënt op fysische eigenschappen. Bewustzijn ontstaat uit een fysische basis volgens contingente, extra natuurwetten: *psychofysische* wetten, die niet volgen uit fysische natuurwetten, die worden verondersteld zelf een gesloten systeem te vormen.^{54 55} De figuur hieronder geeft het eigenschapsdualisme dat Chalmers voorstelt schematisch weer.



Figuur 1: eigenschapsdualisme

Fundamentele eigenschappen zijn dus nog steeds gerelateerd via fundamentele wetten, wat resulteert in een visie die naturalistisch is en compatibel is met het hedendaagse wetenschappelijk wereldbeeld. Er is wel sprake van een uitbreiding van het wetenschappelijk wereldbeeld, met nieuwe fundamentele eigenschappen en wetten. Daarom noemt Chalmers zijn eigenschapsdualisme ook wel een naturalistisch dualisme.⁵⁶ Chalmers wil zich met zijn theorie volgens eigen zeggen niet afkeren van het wetenschappelijk wereldbeeld:

(...) to embrace dualism is not necessarily to embrace mystery.⁵⁷

⁵⁴ Chalmers, *The Conscious Mind*, 125, 127.

⁵⁵ Volgens Chalmers is de fysische wereld *min of meer* causaal gesloten, door het mogelijks bestaan van kwantum onbepaaldheden. In *The Conscious Mind* bespreekt hij de mogelijkheid dat bewustzijn een rol speelt in de ineenstorting van de golffunctie, maar hij is geen voorstander van die visie, onder meer omdat het voor hem niet duidelijk is hoe probabilistische processen, zoals de ineenstorting van de golffunctie, een rol zouden kunnen spelen in het sturen van gedrag. Zie Chalmers, *The Conscious Mind*, 157.

⁵⁶ Chalmers, *The Conscious Mind*, 127-8.

⁵⁷ Chalmers, *The Conscious Mind*, 128.

Fenomenale eigenschappen zijn dus sterk wetmatig afhankelijk van fysische eigenschappen, terwijl het fysische domein van de werkelijkheid autonoom blijft. Men kan zich dan ook de vraag stellen in hoeverre dit naturalistisch dualisme verschillend is van het materialisme. Ik zal dieper ingaan op deze vraag bij mijn bespreking van de overeenkomst tussen DAT en het monisme van Russell.

3.2 Psychofysische wetten

De strategie die Chalmers volgt om bewustzijn te verklaren gaat uit van het leggen van de link tussen het fysische en het fenomenale via het psychologische. Het fysische is hier het substraat en de daarin doorgaande fysische processen die we linken aan cognitieve processen en bewustzijn, zoals onze hersenen en de fysische processen die er plaats vinden tussen de neuronen. Het psychologische is voor Chalmers het geheel van cognitieve functies die betrekking hebben tot het ‘makkelijke’ probleem van bewustzijn. Het psychologische onderscheidt hij van het fenomenale, de manier waarop iets voelt, de intrinsieke natuur van onze ervaringen. Opnieuw zien we hier de vooronderstelling van Chalmers dat er een werkelijk onderscheid bestaat tussen het fenomenale en het psychologische. Chalmers merkt echter op dat men zich kan voorstellen dat de link tussen het fysische en het fenomenale onafhankelijk is van de link tussen het fysische en het psychologische. Hij denkt echter dat er een diepe link is tussen de drie.⁵⁸

Het doel van een fundamentele theorie voor bewustzijn is het formuleren van fundamentele psychofysische wetten, die beschrijven hoe fenomenale eigenschappen uit fysische processen ontstaan. Die wetten zouden volgens Chalmers een echte verklaring voor het ontstaan van ervaringen moeten geven, en dus niet louter mogen bestaan uit een enorme verzameling van uitspraken zoals “hersentoestand X geeft aanleiding tot ervaring Y”.⁵⁹ Chalmers verwacht dat deze wetten een eenvoudige en elegante vorm hebben:

The case of physics tells us that fundamental laws are typically simple and elegant; we should expect the same of the fundamental laws in a theory of consciousness.⁶⁰

Ik weet niet of dit gerechtvaardigd is: de huidige fysica geeft weliswaar de meeste, momenteel gekende fysische wetten en elementaire deeltjes weer in een eenvoudig schema zoals het standaardmodel, maar niets sluit uit dat dit beeld in de toekomst veel complexer wordt. Ook als de ‘uiteindelijke’ fysische

⁵⁸ Chalmers, *The Conscious Mind*, 24-5.

⁵⁹ Chalmers, *The Conscious Mind*, 214.

⁶⁰ Chalmers, *The Conscious Mind*, 127.

wetten in een eenvoudig schema zouden passen geeft Chalmers geen redenen waarom psychofysische wetten een analoog eenvoudige schema zouden moeten volgen.⁶¹

De methode die Chalmers bij de zoektocht naar psychofysische wetten voorstelt is een inferentie naar de beste verklaring. Naast empirische gegevens zouden we bij het vormen van een dergelijke theorie de verschillende alternatieven moeten vergelijken op basis van criteria zoals plausibiliteit, elegantie en eenvoud. Vooral dat laatste criterium is volgens Chalmers belangrijk. De data die we te weten komen over ervaringen vanuit onze eigen persoonlijke fenomenologie, of uit experimenten waarbij proefpersonen rapporteren over hun ervaringen, zouden we via zo eenvoudig mogelijke wetten moeten linken met de onderliggende fysische processen.⁶² Chalmers hoopt dat op die manier een combinatie van experimenten, fenomenologisch onderzoek en filosofische analyse zal leiden tot de onderliggende fundamentele wetten.⁶³ In het onderzoeksprogramma dat Chalmers voorstelt zijn de resultaten, die voortkomen uit de cognitieve en de neurowetenschappen, geen ‘concurrenten’ voor de filosofische visie van het eigenschapsdualisme. Deze wetenschappen krijgen eerder een essentiële rol in het zoeken naar de fundamentele psychofysische wetten. Volgens Chalmers volgen uit de onderzoeksresultaten die reeds werden bekomen in de verschillende wetenschappen een aantal psychofysische wetten. Deze wetten situeren zich echter nog op een te hoog niveau om fundamentele psychofysische wetten te kunnen zijn. Niettemin zullen die wetten, die ik hieronder opsom, een centrale rol spelen in het ontwikkelen van de fundamentele principes die Chalmers voorstelt in DAT.

3.2.1 Het coherentie principe

In het volgende citaat uit *The Conscious Mind* stelt Chalmers een opvallende link vast tussen ervaringen en het cognitief verwerken van informatie, waarbij cognitieve toestanden een functionele, causale rol lijken te spelen:

Conscious experience does not occur in a vacuum. It is always tied to cognitive processing, and it is likely that in some sense it arises from that processing. Whenever one has a sensation, for example, there is some information processing going on: a corresponding perception, if you like. Similarly, whenever one has the conscious experience of happiness, the functional role associated with happiness is generally being played by some internal state. Perhaps it is

⁶¹ Deze ‘uiteindelijke’ fysische wetten bedoel ik hier eerder als gedachtenexperiment, aangezien we natuurlijk geen enkele garantie hebben dat we ooit deze wetten zullen kennen, los van het feit dat we, zelfs al zouden we deze ooit kennen, niet zouden kunnen weten dat deze ook de correcte, uiteindelijke wetten zijn.

⁶² Chalmers, *The Conscious Mind*, 216.

⁶³ Chalmers, “Moving Forward on the Problem of Consciousness,” 5 Conclusion.

logically possible that one could have the experience without the causation, but it seems to be an empirical fact that they go together.⁶⁴

Het coherentie principe geeft uitdrukking aan deze opmerkelijke coherentie tussen ervaringen en gewaar zijn. Fenomenologie en psychologie zijn systematisch gerelateerd met elkaar.⁶⁵

3.2.2 Het structurele coherentie principe

Volgens het meer specifieke structurele coherentie principe is er een gedetailleerde overeenkomst tussen de structurele eigenschappen van de informatie die verwerkt wordt in het brein wanneer we gewaar zijn van iets en de structurele eigenschappen van de ermee gepaard gaande bewuste ervaring. De informatie waarvan we gewaar zijn is de informatie die voor ons beschikbaar is bij het reguleren van ons gedrag. Zoals eerder besproken is gewaar zijn volgens Chalmers een puur functionele notie die te maken heeft met cognitie, te onderscheiden van bewuste ervaring. De twee zijn echter sterk aan elkaar gelinkt. Wanneer we een bewuste ervaring hebben, is er ook sprake van gewaar zijn. Omgekeerd is er ook steeds een bewuste ervaring als we ons gewaar zijn van iets. Dit tweede principe geeft uitdrukking aan de observatie dat die correlatie tussen gewaar zijn en bewuste ervaring nog verder gaat en ook een structureel aspect heeft. Ervaringen hebben een complexe structuur. Zo heeft bijvoorbeeld ons gezichtsveld een complexe geometrische structuur. Wij kunnen een verslag geven van elk van die geometrische relaties, die dus elk cognitief worden gerepresenteerd. De fijnkorrelige structuur aanwezig in onze visuele ervaringen correspondeert dus met een fijnkorrelige structuur die aanwezig is in het verwerken van informatie in ons gewaar zijn. Ook onze interne mentale beelden hebben een geometrische structuur die wordt gerepresenteerd tijdens het cognitief verwerken van informatie. Daarnaast is er tussen verschillende ervaringen ook een relatie van gelijkenis of verschil, alsook een relatie van relatieve intensiteit. De ervaring van het proeven van iets zoets is verschillend van het proeven van iets zouts, en iets kan meer of minder zoet zijn. De verschillende ervaringen bij het proeven van verschillende dingen zijn onderling minder verschillend dan het verschil tussen de ervaring van iets te proeven en de ervaring van bijvoorbeeld het horen van muziek. Voor ieder van die structurele kenmerken is er een corresponderende structuur aanwezig in ons gewaar zijn. Dit is minder duidelijk voor ervaringen zoals bijvoorbeeld emoties, maar ook daar is er sprake van relatieve intensiteit, een structureel kenmerk dat we ook terugvinden bij het verwerken van informatie in ons gewaar zijn. Meer intense emoties hebben immers een grotere invloed op latere processen. Het structurele coherentie principe is niets anders dan het isomorfisme tussen de structuren die aanwezig zijn in onze ervaring en de structuren waarvan we ons gewaar zijn. Chalmers merkt op dat dit geen

⁶⁴ Chalmers, *The Conscious Mind*, 22.

⁶⁵ Chalmers, *The Conscious Mind*, 218.

logisch noodzakelijk principe is, we kunnen ons inbeelden dat het cognitief verwerken van informatie gebeurt zonder enige bewuste ervaring. We zijn zo vertrouwd met het principe dat het door de meeste mensen als vanzelfsprekend wordt beschouwd.⁶⁶

3.2.3 Het *organisationele invariantie principe*

Waar de eerste twee principes eerder voor de hand lijken te liggen, is het derde principe dat Chalmers formuleert veel minder vanzelfsprekend. Volgens het *organisationele invariantie principe* heeft een systeem dat functioneel isomorf is met een bewust systeem dezelfde soort ervaring. Dat een systeem functioneel isomorf is met een ander systeem wil zeggen dat het dezelfde functionele organisatie heeft. Functionele organisatie is voor Chalmers gelijk aan een abstract patroon van causale interactie:

Functional organization is best understood as the *abstract pattern of causal interaction* between various parts of a system, and perhaps between these parts and external inputs and outputs.⁶⁷

De functionele organisatie, en dus niet de specifieke fysische bouw van een systeem, determineert volledig het fenomenaal bewustzijn.⁶⁸ Volgens Chalmers zouden onze ervaringen exact dezelfde blijven als we al onze neuronen vervangen door silicium chips, zolang het causaal patroon van ons oorspronkelijke brein behouden blijft.⁶⁹ ⁷⁰ Maar belangrijker dan de stelling dat functioneel isomorfe systemen dezelfde ervaringen hebben, is de claim van Chalmers dat het de abstracte causale structuur is die aanleiding geeft tot bewustzijn, eerder dan andere, meer ‘concrete’ fysische eigenschappen van het brein zoals biochemische, elektromagnetische of kwantummechanische eigenschappen. Een abstracte causale structuur bestaat uit een aantal abstracte componenten, die in verschillende toestanden kunnen verkeren, samen met causale relaties tussen die toestanden. Belangrijk is dat een dergelijke structuur op verschillende manieren fysisch kan *gerealiseerd* worden, waarbij een of ander fysisch systeem deze abstracte organisatie overneemt:

A physical system realizes a given functional organization when the system can be divided into an appropriate number of physical components each with the appropriate number of states, such that the causal dependency relations among the components of the system, inputs, and

⁶⁶ Chalmers, “Facing up to the problem of consciousness,” 212-13.

⁶⁷ Chalmers, *The Conscious Mind*, 247.

⁶⁸ Chalmers, “Moving Forward on the Problem of Consciousness,” 3.5 My psychophysical laws.

⁶⁹ Chalmers merkt op dat dit in de praktijk technisch onmogelijk zou kunnen zijn. Of dit in de praktijk mogelijk is beschouwt hij als een open vraag. Zie Chalmers, “Facing up to the problem of consciousness,” 215, voetnoot 3.

⁷⁰ Chalmers, “Facing up to the problem of consciousness,” 214.

outputs precisely reflect the dependency relations given in the specification of the functional organization.⁷¹

Abstracte causale patronen komen overal voor in fysische systemen, en op verschillende niveaus. Chalmers overloopt verschillende niveaus in het brein waar dergelijke causale patronen zouden kunnen aanleiding geven tot bewustzijn. Hij hanteert daarbij het aannemelijke criterium dat het niveau “fijnkorrelig” genoeg moet zijn om de complexe functies van onze cognitie te ondersteunen. Hij besluit dat dit niveau zich waarschijnlijk situeert op het niveau van de verbindingen tussen individuele neuronen, al sluit hij niet uit dat nog een lager niveau nodig is om de relevante causale structuur te bekomen.⁷²

Chalmers beschouwt het organisationele invariantie principe als heel aannemelijk, maar het is allerm minst vanzelfsprekend.⁷³ Dit principe is dan ook meer omstreden dan de vorige twee principes. De discussie of andere systemen, zoals computers, aanleiding zouden kunnen geven tot bewustzijn werd al een hele tijd vóór de formulering van DAT door Chalmers gevoerd en is één van de centrale discussiepunten in de *philosophy of mind*. Zo beschouwde bijvoorbeeld Searle ons bewustzijn als een biologisch fenomeen dat causaal afhankelijk is van haar specifieke biochemische oorsprong, net als andere biologische fenomenen als lactatie of fotosynthese.⁷⁴ Reductieve functionalisten, die bewustzijn beschouwen als identiek met bepaalde functionele toestanden, hebben rond deze kwestie een aantal bekende argumenten geformuleerd, waarvan Chalmers enkele voorbeelden bespreekt in *The Conscious Mind*.⁷⁵ Die argumenten werden door tegenstanders van het reductief functionalisme dan weer beantwoord met tegenargumenten, zoals een gedachtenargument van Ned Block, waarbij de neuronen van een menselijk brein gesimuleerd worden door de menselijke populatie van China, die elk zijn uitgerust met radioverbindingen om de synapsen die doorgaan in de hersenen na te bootsen. Dit zou, indien het principe van organisationele invariantie correct is, leiden tot de zeer contra-intuïtieve conclusie dat een dergelijk systeem bewust zou zijn.⁷⁶

Chalmers is echter geen reductief functionalist. Hij denkt dat ervaringen fundamenteel en niet-fysisch zijn. Bewustzijn ontstaat dus uit een bepaalde functionele toestand, maar is niet *gelijk* aan die

⁷¹ Chalmers, *The Conscious Mind*, 247-248.

⁷² Chalmers, *The Conscious Mind*, 248.

⁷³ Chalmers, *The Conscious Mind*, 247.

⁷⁴ John Searle, “Minds, brains and programs,” *Behavioral and Brain Sciences* 3 (1980): 424.

⁷⁵ Chalmers, *The Conscious Mind*, 249-253.

⁷⁶ Ned Block, “Troubles with functionalism,” in *Perception and Cognition: Issues in the Foundation of Psychology*, ed. C. W. Savage (Minneapolis: University of Minnesota Press, 1978). [Reprinted in N. Block, ed., *Readings in the Philosophy of Psychology*, Vol. 1 (Cambridge, Mass.: Harvard University Press, 1980)], 238-41.

functionele toestand. Zijn eigen visie noemt hij dan ook niet-reductief functionalisme.⁷⁷ Chalmers verdedigt zelf het organisationele invariantie principe aan de hand van enkele gedachtenexperimenten, waarbij hij uitgaat van de volgens hem “extreem plausibele aanname” dat een verandering in ervaring overeenkomt met een verandering in de manier waarop informatie functioneel wordt verwerkt. Hij bespreekt zijn eigen versie van een “wegkwijnende qualia” gedachtenexperiment, waarbij neuronen bij een proefpersoon één voor één worden vervangen door silicium computerchips, terwijl de functionele organisatie van het brein wordt behouden. Aangezien een verandering in ervaring volgens Chalmers ook moet overeen komen met een verandering in de functionele verwerking van informatie, moeten de twee functioneel isomorfe systemen dezelfde ervaringen hebben.⁷⁸ Chalmers bespreekt ook een variant op dit gedachtenexperiment, dat handelt over “dansende qualia”. Daarbij worden twee bewuste systemen met verschillende kleurervaringen beschouwd, waarbij intermediaire stadia worden gecreëerd tussen beide kleurervaringen.⁷⁹ Ik ga hier echter niet verder in op de uitgebreide bespreking van deze argumenten en hun mogelijk tegenargumenten.

Chalmers merkt op dat reductieve functionalisten zijn gedachtenexperimenten voor het principe van organisationele invariantie zouden kunnen zien als steun voor hun visie dat specifieke fysische structuren, in plaats van de onderliggende abstracte causale structuur van een systeem, gelijk zijn aan bewustzijn.⁸⁰ Dit is inderdaad ook een kritiek die terug te vinden is in de literatuur. Zo beschouwt Hardcastle structurele coherentie en organisationele invariantie juist als redenen voor de identiteit van ervaring met een materieel substraat.⁸¹ Deze kritiek zou terecht kunnen zijn. Chalmers geeft weinig argumenten voor zijn claim dat bewustzijn bepaald wordt door het abstracte causale patroon in systemen. Misschien formuleerde hij dit principe eerder vanuit een veronderstelde analogie tussen de werking van computers en de werking van hersenen. Computers verwerken informatie door de causale/functionele organisatie van hun elektronische onderdelen. Velen denken dat hersenen ook een soort ‘biologische’ computer zijn. Op die manier komen we tot de veronderstelling dat het verwerken van informatie door onze hersenen gebeurt via de causale/functionele organisatie van de onderdelen van het brein. Voor die relevante onderdelen denken we dan in de eerste plaats aan de neuronen en hun onderlinge verbindingen. Op zich is de conclusie die we trekken uit die analogie heel plausibel. Maar dit wil nog niet zeggen dat deze abstracte/causale organisatie, los van de manier waarop dit fysisch wordt gerealiseerd, het bewustzijn determineert. Zowel computers als hersenen produceren bijvoorbeeld elektromagnetische velden. Het zou kunnen dat ervaringen (al dan niet fysisch van aard)

⁷⁷ Chalmers, *The Conscious Mind*, 249.

⁷⁸ Chalmers, *The Conscious Mind*, 253-259.

⁷⁹ Chalmers, *The Conscious Mind*, 266-271.

⁸⁰ Chalmers, *The Conscious Mind*, 249.

⁸¹ Hardcastle, “The why of consciousness: a non-issue for materialists,” 9.

via psychofysische wetten gelinkt zijn aan de exacte karakteristieken van die elektromagnetische velden. De elektromagnetische velden die uitgaan van mijn hersenen en die van mijn structureel isomorfe kopie uit silicium zouden misschien wel een abstracte, causale structuur delen. Maar de resulterende elektromagnetische velden zouden in beide gevallen verschillend kunnen zijn, bijvoorbeeld in veldsterkte. Het verschil in veldsterktes zou op die manier tot een verschil in ervaringen kunnen leiden. Eenzelfde opmerking geldt voor het elektromagnetisch veld dat zou uitgaan van mijn structureel isomorfe kopie, bestaande uit de populatie van China, zoals in het gedachtenexperiment van Block. Hoewel het niet onmogelijk is dat het causale abstracte patroon een ervaring bepaalt, lijkt dit niet de enige mogelijkheid.

Het zal verder duidelijk worden dat het organisationele invariantie principe een cruciale rol heeft in DAT. Tegelijk kunnen we uit die cruciale rol, samen met het feit dat het principe controversieel is, concluderen dat dit één van de zwakke punten is van DAT. Ik zal hier niet verder ingaan op de discussie over het al dan niet correct zijn van het principe. De onderzoeksvraag is immers wat de rol is van het concept van informatie bij DAT. We zullen zien dat het linken van ervaringen aan abstracte causale structuren en het fysisch realiseren van die structuren beiden een zeer belangrijke rol spelen in de manier waarop Chalmers invulling zal geven aan het concept informatie.

De drie hierboven vermelde principes drukken regelmatigigheden uit op een vrij hoog niveau, omdat ze gaan over hoger niveau concepten zoals gewaar zijn of organisatie. Daarom zijn ze volgens Chalmers geen plausible kandidaten om te fungeren als fundamentele wetten in een theorie van bewustzijn. Ze leggen weliswaar beperkende voorwaarden op aan een ultieme theorie over bewustzijn, maar laten de vraag hoe bewustzijn precies ontstaat en wat er precies nodig is voor fenomenaal bewustzijn ondergedetermineerd. Om tot een echte theorie over bewustzijn te komen zijn er volgens Chalmers fundamentele psychofysische wetten nodig.⁸²

3.3 Een *theory of everything*? Of nog verder?

Het uiteindelijke doel van de fysica is te komen tot een '*theory of everything*' (TOE) die de volledige fysische werkelijkheid beschrijft.⁸³ Aangezien Chalmers denkt dat bewustzijn niet-fysisch is, zou een dergelijke TOE niet de volledige werkelijkheid beschrijven. Het uiteindelijke doel dat Chalmers voor ogen heeft is dan ook te komen tot een *theory of everything* die ook bewustzijn bevat.⁸⁴ Volgens Chalmers zou het, tenminste in principe, mogelijk moeten zijn om tot een fundamentele theorie te

⁸² Chalmers, "Facing up to the problem of consciousness," 215-216.

⁸³ Hiermee beweer ik natuurlijk niet noodzakelijk dat zo'n TOE ooit bereikt kan worden. En opnieuw: mochten we ooit beschikken over een TOE zullen we nooit helemaal zeker zijn dat dit de definitieve, correcte theorie is.

⁸⁴ Chalmers, *The Conscious Mind*, 126, 128-9.

komen die zowel het fysische als het fenomenale domein omvat, met zowel fysische als psychofysische wetten. Hogere psychofysische wetten zouden zich daarbij verhouden tot de lagere, fundamentele psychofysische wetten zoals de principes van de thermodynamica zich verhouden tot de onderliggende principes van de statistische mechanica in de fysica.⁸⁵ De fundamentele psychofysische wetten zouden de basiseigenschappen van ervaring verbinden met basiseigenschappen van de fysische wereld. Samen met de verdeling van de fundamentele eigenschappen zouden hieruit, aldus Chalmers, alle feiten over de wereld volgen.⁸⁶ De fundamentele psychofysische wetten zouden ook precies moeten zijn en geen ruimte mogen laten voor onderdeterminatie. Hij verwacht dan ook dat de fundamentele psychofysische wetten, wanneer gecombineerd met de fysische feiten van een systeem, ons in staat zouden moeten stellen om “de fenomenale feiten van een systeem perfect te voorspellen”.⁸⁷ Tenslotte zouden we uit de fundamentele psychofysische wetten ook de hierboven beschreven principes van coherentie en organisationele invariantie moeten kunnen afleiden en verklaren. Dit is natuurlijk zeer hoog gegrepen. Zelfs als het onderzoeksprogramma dat Chalmers voorstelt zou slagen, zou het bijvoorbeeld niet uit te sluiten zijn dat er verschillende theorieën gevonden worden, waarbij de basiseigenschappen van ervaring worden verbonden met verschillende fysische basiseigenschappen. Chalmers denkt wel dat het nog lang kan duren voor dit alles bereikt zal worden, en ziet het voorstel dat hij zelf doet in *The Conscious Mind* eerder als de start van een dergelijk onderzoeksprogramma.⁸⁸ Merk op dat het moeilijke probleem van bewustzijn werd omschreven als de vraag *hoe* en *waarom* het verwerken van informatie door ons brein gepaard gaat met innerlijke ervaringen. De theorie die Chalmers voor ogen heeft zou eerder verklaren *hoe* dat gebeurt, en niet *waarom*. In dit verband wordt er vanuit materialistische hoek een interessante kritiek geformuleerd door Valerie Gray Hardcastle. De materialistische positie komt er volgens Hardcastle op neer dat er gesteld wordt dat bewustzijn enkel ‘C’ is (waarbij C kan staan voor het neurale correlaat van bewustzijn of iets anders, naargelang de specifieke theorie).⁸⁹ Deze identiteit wordt hierbij gezien als een ‘bruut feit’ over de werkelijkheid. Zij situeert het probleem dat Chalmers heeft met dit soort verklaring in het feit dat dit voor Chalmers niet volstaat. Chalmers ziet deze identiteit niet als een echte verklaring. Maar wat materialisten doen is exact wat Chalmers doet met DAT: ervaring is (zoals we verderop meer in detail zullen zien) simpelweg een aspect van informatie, zonder dat hiervoor een verdere verklaring wordt vereist. Dit is een terechte kritiek, en dit argument lijkt sterk op de mening van Velmans wanneer hij stelt dat het linken van verschillende vormen van bewustzijn aan neurale of andere fysische systemen via de notie van

⁸⁵ Chalmers, *The Conscious Mind*, 213-15.

⁸⁶ Chalmers, *The Conscious Mind*, 127.

⁸⁷ Chalmers, *The Conscious Mind*, 277 (mijn vertaling).

⁸⁸ Chalmers, *The Conscious Mind*, 276-77.

⁸⁹ Hardcastle, “The why of consciousness: a non-issue for materialists,” 8.

informatie nog niet uitlegt hoe bewustzijn *ontstaat*.⁹⁰ Chalmers geeft dit ook toe in zijn reactie op de kritiek van Hardcastle:

Of course, by taking experience as fundamental, there is a sense in which this approach *does not tell us why there is experience in the first place*. But this is the same for any fundamental theory. Nothing in physics tells us why there is matter in the first place, but we do not count this against theories of matter. Certain features of the world need to be taken as fundamental by any scientific theory.⁹¹

Hardcastle verdedigt nu de materialistische positie door te stellen dat in een goede verklaring slechts een beperkt aantal brute feiten worden aangenomen en dat deze zich op een fundamenteel niveau situeren. Volgens haar maakt een verklaring zoals die van Chalmers gebruik van een “perverse metafysica”, omdat in DAT bewustzijn als fundamenteel en als een bruto feit over de werkelijkheid wordt gezien, terwijl bewustzijn volgens haar als een hoger niveau biologisch feit moet gezien worden, dat op zich gebaseerd is op werkelijk fundamentele onderdelen die zich op een dieper niveau van de werkelijkheid situeren.^{92 93} We kunnen ons echter de vraag stellen of het wel klopt dat in een materialistische visie bewustzijn enkel maar een biologisch feit op hoog niveau is. We kunnen immers de evolutionaire ladder afdalen en de vraag stellen in hoeverre er bij minder complexe organismen ook sprake is van fenomenale aspecten. Weinig materialisten zullen ontkennen dat er bij mensapen of zoogdieren eveneens sprake zal zijn van ervaring. Maar als pijn simpelweg het vuren is van C-zenuwvezels, is het dan niet plausibel dat het activeren van een eenvoudig zenuwstelsel zoals dat van de rondworm *C. elegans*, dat slechts bestaat uit enkele honderden neuronen, ook niet gepaard gaat met een fenomenaal aspect, ook al is dit waarschijnlijk sterk verschillend van dat van pijn bij mensen? En hoever is dit verwijderd van een complex informatieverwerkend systeem gebaseerd op silicium? Men moet zelfs niet uitgaan van het principe van organisationele invariantie van Chalmers, volgens welke een dergelijk systeem, mits functioneel dezelfde structuur, dezelfde fenomenale aspecten zou ervaren. Als het vuren van C-zenuwvezels gepaard gaat met fenomenale aspecten, waarom zou de werking van om het even welk fysisch systeem dan niet gepaard gaan met een zeker fenomenaal aspect, hoe verschillend ook? Wie het materialisme verdedigt, zal misschien moeten toegeven dat het goed mogelijk is dat fenomenale aspecten zich niet beperken tot een hoger, biologisch niveau van de

⁹⁰ Velmans, “The relation of consciousness to the material world,” 262.

⁹¹ Chalmers, “Facing up to the problem of consciousness,” 210 (mijn cursivering).

⁹² Hardcastle, “The why of consciousness: a non-issue for materialists,” 9.

⁹³ Hardcastle spreekt in dit verband ook van een zeker biologisch chauvinisme. In haar ogen is het feit dat Chalmers bewustzijn niet noodzakelijk als biologisch ziet enkel een poging om zijn theorie hiervan te redden. Zie Hardcastle, “The why of consciousness: a non-issue for materialists,” 9.

werkelijkheid. Tenzij er beperkende voorwaarden gelden, waarbij bijvoorbeeld een bepaalde mate van complexiteit of een specifiek substraat aanwezig moet zijn, maar een dergelijke beperking lijkt arbitrair en *ad hoc*. Anderzijds wijst Hardcastle op de noodzaak voor dergelijke beperkingen, zij wijst daarbij op het gegeven uit de psychologie dat veel informatieverwerking in ons brein onbewust gebeurt.⁹⁴

Chalmers geeft dus inderdaad toe dat zijn theorie niet zou verklaren *waarom* bewustzijn bestaat in onze wereld. Hij stelt dat de fundamentele psychofysische wetten in een dergelijke theorie 'brute' feiten zouden zijn over de werkelijkheid, die niet verder verklaard worden door de theorie. Hij stelt echter dat er in dit aspect geen verschil is met fysische theorieën, die fenomenen in de wereld verklaren vanuit wetten die zaken zoals massa, lading, ruimte en tijd met elkaar relateren tot op een steeds lager, fundamenteeler niveau van de fysische werkelijkheid. Fenomenen op macroscopische schaal worden verklaard aan de hand van natuurwetten en eigenschappen op een lager, fundamenteeler niveau. Op een bepaald moment komt er echter een eind aan dit proces. Er worden entiteiten met eigenschappen geponeerd die als fundamenteel worden verondersteld. We worden dan verondersteld de fundamentele wetten op dit basisniveau voor waar aan te nemen, zonder verdere verklaring, als een soort contingent maar 'bruut' feit over de wereld. Fysische theorieën verklaren evenmin wat de zaken, die door die fundamentele natuurwetten met elkaar gerelateerd worden, uiteindelijk zijn. Diezelfde theorieën geven evenmin een verklaring waarom die zaken bestaan, wat de oorsprong ervan is en waarom ze zo zijn en niet anders. Psychofysische wetten die een verband leggen tussen bewustzijn en haar fysische basis zouden in die zin niet anders zijn.

Chalmers koestert hoge ambities voor een uiteindelijke theorie, zoveel wordt duidelijk uit het volgende citaat:

Of course, it may be that in the quest for such theories, there will be developments that change our conception of an ultimate theory. It may be, for example, that we will find overarching laws that subsume the phenomena of both physics and consciousness into a grander theory, just as we found a theory that subsumed electricity and magnetism, and as physicists are now searching for a theory that unifies all the basic physical forces. *Perhaps there will be developments that are more surprising still.*⁹⁵

Vooraf die laatste zin is belangrijk. In zijn artikel "Facing up to the problem of consciousness" uit 1995 en het vlak daarna verschenen *The Conscious Mind* doet hij een zeer speculatief voorstel voor een overkoepelende theorie met fundamentele psychofysische wetten en principes. Zijn voorstel, dat hij

⁹⁴ Hardcastle, "The why of consciousness: a non-issue for materialists," 9.

⁹⁵ Chalmers, *The Conscious Mind*, 215 (mijn cursivering).

naar eigen zeggen beschouwt als een proto-theorie of een verzameling losse ideeën, gaat verder dan wetenschappelijke theorieën en is eerder een metafysische theorie.⁹⁶ In zijn meest radicale vorm poogt de dubbel-aspect theorie van informatie zowel bewustzijn als het fysische domein van de wereld te verklaren als verschillende aspecten van informatie. Daarmee bouwt hij voort op reeds eerder bestaande theorieën die informatie als fundamenteel beschouwen. We zullen verder zien dat hij daarmee in de buurt komt van theorieën die niet enkel een *theory of everything* zijn, maar ook proberen om de *waarom* vraag op te lossen, waarom er zowel een fysische als een fenomenaal domein is in de werkelijkheid en wat de oorsprong is van dit alles. Maar vooraleer we het voorstel van Chalmers kunnen bespreken moeten we eerst onderzoeken wat Chalmers precies bedoelt met informatie.

3.4 Het concept van informatie

3.4.1 Informatie

Volgens de *Stanford Encyclopedia of Philosophy* wordt met de term 'informatie' in het hedendaagse alledaagse taalgebruik een hoeveelheid gegevens, code of tekst aangeduid die bewaard, verzonden, ontvangen of gemanipuleerd wordt in om het even welk medium. Deze omschrijving van informatie als 'data, gegevens of tekst' lijkt circulair, en toont hoe moeilijk het is om een grip te krijgen op wat informatie precies is. Informatie is in onze cultuur een veelgebruikt en praktisch woord om het 'spul' mee aan te duiden dat door computers en andere instrumenten 'stroomt', maar de term wordt op een losse, onnauwkeurige en inconsistente manier gebruikt door de meeste mensen. Zo is het niet onmiddellijk duidelijk of informatie waar moet zijn om nog van informatie te kunnen spreken, getuige hiervan het gebruik van de aanduiding 'valse informatie'. Hoewel de notie van informatie al vanaf het begin een rol heeft gespeeld in de Westerse filosofie, dateert een expliciete analyse van het filosofische concept van informatie pas van de tweede helft van de twintigste eeuw.⁹⁷

Er zijn veel verschillende concepten van informatie, die geleid hebben tot een eigen subdiscipline van de filosofie, namelijk de filosofie van de informatie. Theorieën over informatie worden ingedeeld in kwalitatieve theorieën, waarin gewerkt wordt rond de betekenis die aan informatie wordt gegeven, en kwantitatieve theorieën, waarin eerder de hoeveelheid informatie centraal staat. De laatste twee decennia worden pogingen gedaan om tot een meer geünificeerd concept van informatie te komen.⁹⁸

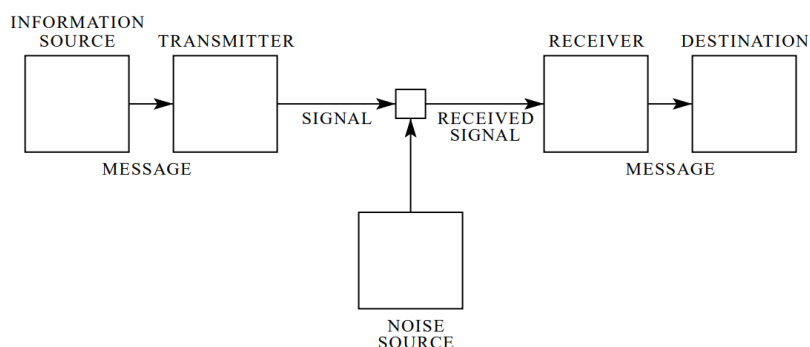
⁹⁶ Chalmers, *The Conscious Mind*, 277.

⁹⁷ Pieter Adriaans, "Information," *The Stanford Encyclopedia of Philosophy*, Fall 2020, 1. Information in Colloquial Speech.

⁹⁸ Voor een overzicht zie Pieter Adriaans, "Information," *The Stanford Encyclopedia of Philosophy*, Fall 2020.

3.4.2 Shannon informatie

Chalmers neemt als uitgangspunt voor de ontwikkeling van DAT het concept van informatie zoals ontwikkeld door Shannon.⁹⁹ De theorie van Shannon over informatie is eerder een kwantitatieve theorie.¹⁰⁰ Shannon definieerde informatie in het kader van een mathematische theorie over communicatie. Daarbij maakte hij gebruik van het onderstaande schema, waar een zender een informatiebron is die berichten stuurt over een kanaal naar een ontvanger. Een zender vertaalt hierbij een boodschap (die kan bestaan uit letters, geluiden of beelden) naar een signaal, terwijl de ontvanger het omgekeerde doet. Het schema omvat ook factoren die kunnen zorgen voor een verstoring of vervorming van het signaal. Communicatiesystemen kunnen discreet zijn, waarbij zowel de boodschap als het signaal een serie discrete symbolen zijn (zoals bij het vroegere telegramsysteem waarbij dit respectievelijk letters en een serie punten, strepen en spaties zijn), continu (waarbij zowel het signaal als de boodschap continu zijn zoals het geluid en de proportionele elektrische stroom bij een klassieke telefoonlijn), of gemengd (waarbij zowel discrete als continue variabelen voorkomen).¹⁰¹



102

Figuur 2: communicatiesysteem volgens Shannon

De hoeveelheid informatie I die een bericht x bevat wordt gegeven volgens de onderstaande formule:

$$I(x) = -\log p_x$$

De waarschijnlijkheid van een boodschap x is ergens tussen 0 en 1. Als de waarschijnlijkheid $p_x = 1$ is, dan bevat de boodschap geen informatie of 'nieuws'. Hoe onwaarschijnlijker de boodschap is, hoe groter de hoeveelheid informatie die de boodschap bevat. Belangrijk is dat het concept van informatie bij Shannon niets te maken heeft met de betekenis van een boodschap, het gaat hier eerder over een

⁹⁹ Chalmers, "Facing up to the problem of consciousness," 216.

¹⁰⁰ Claude Shannon, "A mathematical theory of communication," *Bell Systems Technical Journal* 27 (1948):379-423. [Reprinted in C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication* (Urbana: University of Illinois Press, 1949)].

¹⁰¹ Claude Shannon, "A mathematical theory of communication," 382.

¹⁰² Claude Shannon, "A mathematical theory of communication," 381.

toestand die deel uitmaakt van een geheel van mogelijke toestanden. Het is dus geen semantische visie op informatie maar eerder een kwantitatieve: als twee boodschappen even (on)waarschijnlijk zijn dan is hun hoeveelheid informatie dezelfde.¹⁰³ Semantische informatie maakt daarentegen een claim over iets in de wereld, zoals bijvoorbeeld dat de Eiffeltoren in Parijs staat. Shannon is expliciet over het ontbreken van semantische inhoud in zijn theorie over informatie:

Frequently the messages have meaning; that is, they refer to or are correlated according to some system with certain physical or conceptual entities. These semantic aspects of communication are irrelevant to the engineering problem.¹⁰⁴

3.4.3 Informatie volgens Chalmers

Chalmers gebruikt als uitgangspunt, net zoals Shannon, een meer abstracte notie van informatie, in tegenstelling tot een semantische visie, volgens dewelke informatie steeds *over iets* gaat. Chalmers zal nu echter een heel eigen visie op informatie gaan ontwikkelen. Informatie zal hij daarbij opvatten als niets anders dan abstracte of primitieve verschillen tussen mogelijke toestanden, zonder dat die toestanden verder worden omschreven. Chalmers gaat nu de kleinste hoeveelheid informatie definiëren aan de hand van de *bit*, die ook door Shannon werd gebruikt als eenheid om een hoeveelheid informatie weer te geven.¹⁰⁵ Die bit vertegenwoordigt voor Chalmers de keuze of een primitief verschil tussen twee mogelijke toestanden (voorgesteld door 0 of 1).¹⁰⁶ Vervolgens gaat Chalmers die informatie een syntactische invulling geven: hij zal informatie zien als primitieve verschillen tussen informatietoestanden die gestructureerd kunnen worden in een abstracte structuur, een informatieruimte. In zijn latere boek *Reality+* noemt hij dit 'structurele informatie'.¹⁰⁷

3.4.4 Informatieruimtes

Volgens Chalmers moeten we ons informatietoestanden voorstellen als ingebed in informatieruimtes.¹⁰⁸ Een specifieke informatietoestand is de selectie van één welbepaalde informatietoestand uit een abstracte ruimte van verschillende mogelijkheden, een informatieruimte. Een informatieruimte bestaat uit de mogelijke informatietoestanden en een basisstructuur van differentierelaties tussen die informatietoestanden. De differentierelaties, die zeer complex kunnen zijn, drukken de manier uit waarop informatietoestanden gelijkend of verschillend kunnen zijn. De

¹⁰³ Adriaans, "Information." 4.2 Shannon: Information Defined in Terms of Probability.

¹⁰⁴ Shannon, "A mathematical theory of communication," 379.

¹⁰⁵ Shannon, "A mathematical theory of communication," 379.

¹⁰⁶ Chalmers, *The Conscious Mind*, 278.

¹⁰⁷ Chalmers, *Reality+, virtual worlds and the problems of philosophy*, 155.

¹⁰⁸ Chalmers, "Facing up to the problem of consciousness," 216.

meest eenvoudige, niet-triviale informatieruimte is de ruimte die slechts twee mogelijke toestanden of keuzes bevat, met een primitieve differentierelatie, zoals de informatieruimte die slechts twee bits bevat, 0 en 1.

Informatieruimtes kunnen op twee manieren meer complex gemaakt worden. Ten eerste door een complexe differentiestructuur tussen de informatietoestanden (bijvoorbeeld niet alleen toestanden '0' of '1' maar '0', '1', '2' en '3'). Voorlopig gaat het hier enkel over de structuur van informatieruimtes. De aard van de labels '0' of '1' is hier nog niet van belang. Het is mogelijk om dit kader uit te breiden zodat het een semantisch aspect krijgt, maar voorlopig concentreren we ons hier enkel op de structuur. De differentierelatie kan ook complexer zijn, met een oneindig aantal informatietoestanden, analoog aan het continuüm van reële getallen tussen 0 en 1 (Chalmers merkt op dat Shannons notie van informatie ook deze mogelijkheid behandelt). Sommige staten liggen daarbij dichter bij elkaar dan andere. De structuur van een informatieruimte kan ook die zijn van een 2-dimensionaal continuüm, of n-dimensionaal, of nog algemener ieder mogelijke topologische ruimte, met een 'nabijheidsrelatie'.¹⁰⁹ Een andere manier om informatieruimtes meer complex te maken bestaat eruit de informatietoestanden zelf een interne structuur te geven (door bijvoorbeeld een gestructureerde ruimte met informatietoestanden of een 'bericht' zoals '1101110100'). Deze laatste informatietoestand bestaat uit tien elementen, waarbij ieder element kan gezien worden als de keuze tussen twee toestanden van een *subruimte*. De informatieruimte is het product van tien subruimtes. Opnieuw zijn er hier continue interne structuren mogelijk, met informatietoestanden die een oneindig aantal elementen hebben. Als we beide manieren combineren bekomen we informatietoestanden zoals bijvoorbeeld '25122469'. Het meest algemene geval is wanneer een informatieruimte beide niveaus van complexiteit vertoont, waarbij iedere informatietoestand een interne structuur heeft, waarbij ieder element van de informatietoestand behoort tot een subruimte met een eigen topologische differentierelatie. Hier spreekt Chalmers respectievelijk van de *combinatorische* structuur van de ruimte en de *relationele structuur* van de subruimtes.¹¹⁰

Informatieruimtes zijn abstracte ruimtes, en informatietoestanden zijn abstracte toestanden. Beiden maken — op het eerste zicht — geen deel uit van het fysische of fenomenale domein van onze wereld. In *The Conscious Mind* vraagt Chalmers de lezer echter om mee te gaan in het leggen van de link tussen deze abstracte informatieruimte en de concrete fysische wereld:

¹⁰⁹ Chalmers, *The Conscious Mind*, 278-79.

¹¹⁰ Chalmers, *The Conscious Mind*, 279.

But we can find information in both the physical and the phenomenal world, if we look at things the right way. To do this, we need to discuss the various ways in which information spaces and states can be *realized* in the world.¹¹¹

Informatie wordt volgens Chalmers dus gerealiseerd in de wereld, op dezelfde manier als de hierboven besproken abstracte causale patronen gerealiseerd worden in fysische systemen. Elders spreekt Chalmers ook van informatie die fysisch *embodied* is, belichaamd.¹¹² Het is duidelijk dat zowel de fysische als fenomenale wereld op één of andere manier informatie ‘bevat’. Wat Chalmers echter precies bedoelt met het *realiseren* van informatieruimtes en -toestanden wordt niet onmiddellijk duidelijk in *The Conscious Mind* en vereist een diepgaande lezing en interpretatie, een *close reading* van dit werk en een vergelijking met andere werken van Chalmers waarin hij verwijst naar het dubbel-aspect principe van informatie. Een moeilijkheid hierbij is dat Chalmers frequent verwijst naar de visies van andere auteurs, die echter steeds in een zekere mate verschillend zijn van de visie die Chalmers zelf uiteindelijk verdedigt. We zullen zien dat na een diepgaande lezing van de werken van Chalmers een aantal impliciete vooronderstellingen van Chalmers duidelijk worden, maar ook dan nog blijven heel wat elementen van de dubbel-aspect theorie eerder vaag. Dit kan te wijten zijn aan het feit dat Chalmers volgens eigen zeggen de dubbel-aspect theorie eerder tentatief bedoelde, als een allesbehalve definitieve poging om een ruimere theorie rond bewustzijn te construeren, waarbij hij een aantal mogelijke interpretaties van het principe aangeeft zonder een duidelijke keuze te maken tussen die verschillende mogelijkheden.¹¹³ Een alternatieve verklaring voor de onduidelijkheden in de dubbel-aspect theorie is dat deze inconsistenties en/of niet aangekaarte vragen en problemen bevat. We zullen dit hieronder proberen uit te spitten.

3.4.5 Fysisch gerealiseerde informatie

Ook al geeft Chalmers toe dat er bij Shannon geen expliciete behandeling is van de relatie tussen fysische toestanden en informatietoestanden, ziet Chalmers een analogie tussen de manier waarop volgens hem informatie fysisch gerealiseerd wordt en het beeld dat Shannon schetst van een communicatiesysteem, waarbij een boodschap die informatie bevat wordt doorgegeven via een fysisch kanaal zoals bijvoorbeeld een telefoonlijn. Voor Chalmers is een communicatiekanaal in het hierboven weergegeven diagram van Shannon te interpreteren als een causale keten, waarbij verschillende door de zender verzonden boodschappen verschillende effecten kunnen hebben op de ontvanger.

¹¹¹ Chalmers, *The Conscious Mind*, 280.

¹¹² Chalmers, “Facing up to the problem of consciousness,” 216.

¹¹³ Chalmers, *The Conscious Mind*, 277.

Chalmers beschrijft enkele voorbeelden van hoe hij informatietoestanden en informatieruimtes ‘gerealiseerd’ ziet in de fysieke wereld. Zo realiseert een klassieke lichtsakelaar een informatieruimte van twee mogelijke toestanden (‘omhoog’ en ‘omlaag’), terwijl de specifieke stand van de lichtsakelaar één van deze twee abstracte informatietoestanden realiseert. Een voorwerp dat zich in verschillende toestanden kan bevinden realiseert dus volgens Chalmers een informatieruimte, terwijl een specifieke toestand van dat voorwerp een specifieke informatietoestand realiseert.¹¹⁴ Chalmers refereert hier dus naar abstracte informatieruimtes, maar het is voorlopig niet duidelijk wat de ontologische status is van die abstracte ruimtes. Het zou kunnen dat Chalmers die abstracte ruimtes ziet als werkelijk bestaand in een soort platoonse wereld, die al dan niet fysisch gerealiseerd kunnen worden. Maar het zou ook kunnen dat Chalmers deze ruimtes ziet als een structuur die men kan herkennen in de fysieke wereld, maar die niet los kan bestaan van die fysieke wereld. Ik zal hierop terugkomen bij de bespreking van de “*it from bit*” theorie. Chalmers gaat vervolgens informatietoestanden verbinden met causaliteit. De informatieruimte geassocieerd met een fysisch voorwerp gaat hij betrekken bij een causale keten en een ruimte van mogelijke effecten op het eind van die causale keten, zoals het al of niet branden van een lamp na het aan- of uitschakelen van een sakelaar. Merk op dat hierdoor het concept van informatie bij Chalmers een semantische invulling krijgt, en hier dus afwijkt van een puur syntactische of formele invulling van het concept informatie, de notie van informatie wordt betrokken op de wereld.

Chalmers stelt dat we informatie kunnen zien als gerealiseerd of fysisch belichaamd wanneer er een ruimte is van verschillende fysieke toestanden, waarbij die verschillen kunnen doorgegeven worden via een causale keten.¹¹⁵ Een fysieke toestand zal dus corresponderen met een informatietoestand die een bepaald effect heeft op de causale keten. Een fysisch voorwerp bevindt zich volgens Chalmers dus steeds in een welbepaalde informatietoestand, terwijl de ‘causale flux’ die door de wereld loopt die toestanden voortdurend verandert. Als Chalmers een *theory of everything* voor de fysieke werkelijkheid wil formuleren kan hij in feite niets anders dan een verband leggen met causaliteit, aangezien het voor een fysisch object juist essentieel is dat het deel kan uitmaken van causale ketens, in tegenstelling tot een abstract iets zoals een getal of het concept schoonheid. Chalmers maakt in dit verband nog een verwijzing naar de denker Gregory Bateson, die informatie omschreef als een “verschil dat een verschil uitmaakt”. Bateson was onder meer werkzaam in de cybernetica.¹¹⁶ Cybernetica is de wetenschap van controle en communicatie in dier en machine. Chalmers geeft hier echter mogelijks

¹¹⁴ Chalmers, *The Conscious Mind*, 280-281.

¹¹⁵ Chalmers, “Facing up to the problem of consciousness,” 216.

¹¹⁶ Gregory Bateson, *Steps to an Ecology of Mind. Collected essays in anthropology, psychiatry, evolution, and epistemology* (Jason Aronson Inc., 1987). Reprint. Originally published: (San Francisco: Chandler Pub. Co., 1972), 460.

opnieuw een eigen invulling aan een concept van informatie dat hij leent bij een andere denker, in dit geval Bateson. Verschillen kunnen volgens Bateson juist een verschil maken omdat, bijvoorbeeld in het geval van het brein, de neurale paden of zenuwbanen langs waar de informatie zich verplaatst voorzien zijn van energie. De neurale paden zijn als het ware reeds ‘getriggerd’ om te reageren op informatie. De visie die Chalmers ontwikkelt rond informatie is mogelijk heel verschillend en veel abstracter dan de visie van Bateson. De stroom van informatie bij Bateson vindt niet plaats in een abstracte ruimte maar in de fysische realiteit. Voor het doorgeven van informatie is energie en een materiële ‘drager’ van informatie nodig. We zullen verder zien dat Chalmers een zeer abstracte kijk op de werkelijkheid verdedigt.

De structuur van een fysisch gerealiseerde informatieruimte kan discreet zijn, zoals in het voorbeeld van de klassieke lichtsakelaar, met slechts twee mogelijke effecten op de causale keten (de lamp aan of uit). Dit kan men veralgemenen tot continue informatieruimtes, zoals bij een dimmer schakelaar die verschillende lichtintensiteiten over een continue bereik als effect heeft.¹¹⁷ De dimmer schakelaar realiseert in die zin een continue informatieruimte.

Een belangrijke vraag die men zich bij dit alles kan stellen is of de fundamentele onderdelen van de microwereld en de wetten die er gelden wel voorgesteld kunnen worden door deze eenvoudige voorbeelden met schakelaars. Met name is het de vraag in hoeverre dit overeen kan komen met de verschijnselen uit de kwantumwereld. Chalmers maakt bij de beschrijving van zijn theorie natuurlijk alleen maar gebruik van een uiterst eenvoudig ‘speelgoed’ model. Hoe fysische systemen, ook zeer eenvoudige zoals één enkel elektron dat interageert met een ander elektron, exact een causale invloed op elkaar uitoefenen is nog grotendeels onbekend. Het is nog niet gelukt om de fysische wetmatigheden die we tot nu toe hebben geformuleerd te verenigen in één consistente overkoepelende theorie. Chalmers is op dit vlak misschien wat te optimistisch en lijkt te denken dat we reeds ver gevorderd zijn in onze kennis van de fysische werkelijkheid, wanneer hij het heeft over onze kennis van fundamentele, onderliggende fysische wetten.¹¹⁸ Chalmers lijkt bij zijn beschrijving van causale relaties ook eerder uit te gaan van het determinisme, waarin fysische objecten verondersteld worden te verkeren in welbepaalde toestanden.¹¹⁹ Het is echter niet zeker dat de fysische wereld op

¹¹⁷ Chalmers merkt hierbij op dat het voorbeeld dat hij hier heeft eerder een idealisering is van de werkelijkheid, waarmee hij waarschijnlijk bedoelt dat in de werkelijkheid, volgens de huidige fysica, licht in discrete pakketjes, fotonen, wordt geproduceerd. Zie Chalmers, *The Conscious Mind*, 281.

¹¹⁸ Chalmers, *The Conscious Mind*, 216.

¹¹⁹ Chalmers houdt misschien wel degelijk rekening met kwantummechanische fenomenen in zijn beschrijving van DAT, alhoewel dit niet expliciet naar voor komt. Zo zegt Chalmers dat de informatietoestanden die hij bespreekt in *The Conscious Mind* ook overeen kunnen komen met een functie over een continu domein en bereik zoals een golfvorm. Ik heb weinig kennis over kwantummechanica, maar dit laat misschien de mogelijkheid toe om golfvergelijkingen te vertalen naar deze informatietoestanden. Zie Chalmers, *The Conscious Mind*, 279.

een deterministische manier werkt. De kwantummechanica lijkt vooralsnog eerder een probabilistisch beeld van de wereld te schetsen, ook al zijn er reeds deterministische ‘interpretaties’ geformuleerd. Een andere vraag die men kan stellen is welke precieze causale ketens Chalmers bedoelt, aangezien een voorwerp een rol kan spelen in oneindig veel verschillende causale ketens. Uiteindelijk heeft een bepaalde toestand van een fysisch object een invloed op de wereld, door bijvoorbeeld een kleine verplaatsing van massa of lading. Men moet daarbij zelfs niet onmiddellijk aan het vlindereffect denken, aangezien ook een bijzonder kleine verandering van de omgeving van een fysisch object ook een verandering is. Misschien is wat Chalmers hier bedoelt de causale rol die een specifiek fysisch voorwerp speelt in de werkelijkheid, in de ontplooiing van die werkelijkheid zoals deze gebeurt in de tijd. Iedere alternatieve fysische toestand waarin een voorwerp zich op een bepaald moment zou bevinden geeft aanleiding tot een alternatieve wereldgeschiedenis. Iedere informatietoestand van een fundamenteel deeltje heeft uiteindelijk een causaal effect op de informatietoestand van alle andere deeltjes in de fysische werkelijkheid.¹²⁰ De causale interacties en de informatieruimtes waarvan Chalmers spreekt zijn oneindig of quasi oneindig complex, ook voor slechts één fundamenteel deeltje. Deze interacties en informatieruimtes kunnen we ons, zelfs al zouden we alle fysische wetten kennen, hoogstens op een vage en intuïtieve manier inbeelden. Chalmers geeft in *The Conscious Mind* toe dat dit beeld wat vaag en informeel is. Het formuleren van meer precieze criteria voor het realiseren van een informatieruimte in een fysisch systeem maakt volgens hem deel uit van het onderzoeksproject dat hij voorstelt.¹²¹

Een deel van de kritiek op DAT bestaat er inderdaad uit dat Chalmers een te simplistisch beeld van de fysische werkelijkheid schetst, en dat het misschien te vroeg is om een theorie als DAT te formuleren, als zaken zoals causaliteit en de fysische microwereld nog zo slecht begrepen zijn. Volgens Douglas Bilodeau baseert Chalmers zich op een te simpel en onkritisch concept van het fysische. Hij bekritiseert daarbij het feit dat Chalmers in *The Conscious Mind* meerdere keren verwijst naar de “keiharde” microfysische feiten en wetten, die Chalmers als fundamenteel, toegankelijk en onproblematisch beschouwd, terwijl hij in het laatste hoofdstuk van zijn boek juist de verschillende interpretaties van de kwantummechanische wereld allemaal als “een beetje gek” beschouwt.^{122 123 124 125} Bilodeau merkt op

¹²⁰ Die causale invloeden kunnen zich volgens onze huidige kennis wel slechts met lichtsnelheid verplaatsen, met een mogelijke uitzondering voor verstrengelde toestanden.

¹²¹ Chalmers, *The Conscious Mind*, 283.

¹²² Douglas Bilodeau, “Physics, Machines, and the hard problem,” *Journal of Consciousness Studies* 3, No. 5-6 (1996): 386-401.

¹²³ Chalmers, *The Conscious Mind*, 88.

¹²⁴ Chalmers, *The Conscious Mind*, 356.

¹²⁵ Chalmers verdedigt in het laatste hoofdstuk van *The Conscious Mind* de Everettiaanse interpretatie van de kwantummechanica, hoewel hij zelf toegeeft dat ze moeilijk te geloven is en waarschijnlijk onjuist. Zie Chalmers, *The Conscious Mind*, 357 en 390, noot 9.

dat er een tegenstrijdigheid schuilt in het vertrekken van het fysische als een basis voor de discussie omtrent de ontologie van het bewustzijn, als men tegelijk meent dat de betekenis van de concepten van de microfysische wereld te controversieel zijn om overeenstemming over te bereiken.¹²⁶ Chalmers baseert zich volgens Bilodeau op wat hij het ‘geometrische paradigma’ noemt:

It is assumed that an independently existing reality must take the form of a mathematical structure which the structure of our experience reflects and from which all physical truth may be derived. If physics is not working to bring that structure to light, then it is not doing its job.¹²⁷

Bilodeau merkt op dat deze visie niet noodzakelijk juist is, en eerder een beperkt en ontologisch verarmd beeld is van de werkelijkheid. Het is natuurlijk goed mogelijk dat Chalmers deze fouten begaat en, ingekapseld in wat Bilodeau het geometrische paradigma noemt, een te simplistisch beeld hanteert van de fysische wereld bij het formuleren van zijn visie op het fysische en het fenomenale. Maar hieruit volgt niet noodzakelijk dat DAT onjuist is, alhoewel het de status van DAT als ‘theorie’ ernstig aantast.¹²⁸

3.4.6 Fenomenaal gerealiseerde informatie

Informatie kunnen we niet alleen vinden in de fysische wereld, maar ook in onze ervaringen, zoals bij het ervaren van verschillende kleuren. DAT linkt bewustzijn op een heel specifieke manier aan het concept van informatie, maar Chalmers is niet de eerste om de link te leggen tussen bewustzijn en informatie. David Lewis sprak in zijn *“What Experience Teaches”* uit 1988 reeds over fenomenale informatie, waarmee hij de hypothese aanduidde dat er naast fysische informatie een andere soort informatie zou bestaan, die niet tot fysische informatie zou te reduceren zijn, net zoals informatie over een x-coördinaat ons niets kan zeggen over de y-coördinaat. Die fenomenale informatie zou gaan over intrinsieke, niet-fysische eigenschappen van ervaringen. Lewis was echter zelf een reductionistisch materialist, en was van mening dat ervaringen niets anders waren dan toestanden van het brein, die een functionele rol spelen.¹²⁹

Volgens Chalmers zijn onze ervaringen fundamenteel verbonden met de notie van informatieruimtes. Tussen onze ervaringen zijn er gelijkenissen en verschillen, zoals onze ervaring van de kleur donkerrood, die verschillend is van de ervaring van de kleur lichtrood, terwijl beide ervaringen dan weer onderling meer gelijkend zijn dan elk vergeleken met de kleur groen. Die patronen van gelijkenissen en verschillen komen volgens Chalmers overeen met de differentiestructuur van informatieruimtes. Daardoor kunnen we volgens hem fenomenale toestanden zien als gerealiseerde informatietoestanden van die

¹²⁶ Bilodeau, “Physics, Machines, and the hard problem,” 386.

¹²⁷ Bilodeau, “Physics, Machines, and the hard problem,” 395.

¹²⁸ Bilodeau, “Physics, Machines, and the hard problem,” 395.

¹²⁹ David Lewis, “What experience teaches,” 77-103.

informatieruimtes. Chalmers bespreekt slechts enkele voorbeelden in *The Conscious Mind*. Zo stelt Chalmers dat de informatieruimte die we kunnen linken aan kleurwaarneming een eenvoudige 3-dimensionele structuur heeft (met als X,Y en Z als de primaire kleuren). De meer complexe ervaringen, zoals wanneer we een volledig visueel veld ervaren, komen overeen met informatieruimtes met een complexe combinatorische structuur. Als we bijvoorbeeld naar een schilderij kijken, heeft de informatieruimte die we kunnen linken aan onze ervaring een combinatorische structuur die overeen komt met een 2-dimensioneel continuüm, waarbij ieder element van dat continuüm de 3-dimensionele relationele structuur heeft van kleurwaarneming. Bij ervaringen spelen de effecten op causale trajecten geen rol maar eerder de intrinsieke kwaliteiten van die ervaringen en de structuur van gelijkenis- en verschilrelaties tussen die ervaringen onderling.¹³⁰ De voorbeelden die Chalmers geeft om zijn idee te illustreren zijn voornamelijk gebaseerd op eenvoudige of duidelijk gestructureerde ervaringen zoals kleurwaarneming of visuele indrukken. Het is onduidelijk of dit ook geldt voor qualia zoals pijn, geur of complexe emoties. Chalmers beschouwt dit zelf ook als één van de ‘open vragen’ die zijn proto-theorie oproept.¹³¹

3.5 Het dubbel-aspect principe van informatie

De observatie dat we ervaringen kunnen zien als gerealiseerde informatieruimtes wordt door Chalmers nu in verband gebracht met het reeds besproken principe van structurele coherentie. Dat principe stelde dat er een gedetailleerde overeenkomst is tussen de structurele eigenschappen van een ervaring en de structuur van de informatie die cognitief verwerkt wordt. Chalmers zal nu verdedigen dat deze structuur ook ingebed is in de fysische processen die doorgaan in het cognitieve apparaat van ons brein.¹³² Zo geeft hij opnieuw het voorbeeld van het ervaren van kleur en de 3-dimensionele informatieruimte die eraan verbonden kan worden. Ook al weten we niet precies hoe dit gebeurt, deze informatieruimte moet ook fysisch gerealiseerd worden, aangezien we via onze cognitie beschikken over deze informatie. Een meer complexer voorbeeld is ons visueel veld. Daarbij zijn we in staat om te rapporteren over wat we zien, dus de informationele structuur hiervan moet ook fysisch gerealiseerd worden in de processen in ons brein. Er zou dus een direct isomorfisme zijn tussen fysisch gerealiseerde informatieruimtes en de informatieruimtes die fenomenaal gerealiseerd worden. Chalmers gaat nu echter nog een sterkere claim maken. Hij stelt dat deze structurele overeenkomst tussen fysisch gerealiseerde en fenomenaal gerealiseerde informatie veroorzaakt wordt door het feit dat beide verschillende ‘aspecten’ of realisaties zijn van dezelfde abstracte informatieruimte (vandaar de

¹³⁰ Chalmers, *The Conscious Mind*, 283-84.

¹³¹ Chalmers, *The Conscious Mind*, 309.

¹³² Chalmers, “Facing up to the problem of consciousness,” 216.

benaming ‘dubbel-aspect theorie van informatie’). Informatieruimtes hebben in deze visie een soort ‘dubbellevens’ dat aanleiding geeft tot de verbinding van fysische processen en fenomenaal bewustzijn. Complexe ervaringen komen overeen met een combinatorische informatieruimte, die tevens fysisch gerealiseerd wordt in de onderliggende processen in het brein. Wanneer een fenomenale toestand een informatietoestand realiseert, wordt deze informatietoestand tevens gerealiseerd in het fysische substraat van de ervaring, die het volledige brein kan omvatten. Omgekeerd worden — tenminste sommige — fysisch gerealiseerde informatieruimtes ook fenomenaal gerealiseerd.¹³³ Of het dubbel-aspect principe geldt voor alle fysisch gerealiseerde informatieruimtes zal ik bespreken in het deel over panpsychisme.

De overige argumenten die Chalmers formuleert om zijn principe te steunen komen neer op het feit dat dit principe goed overeenstemt met zijn eerder geformuleerde principes van coherentie, structurele coherentie en organisationele invariantie. Het dubbel-aspect principe lijkt sterk op dit laatste principe, dat stelde dat de abstracte causale structuur tussen neuronen ons bewustzijn bepaalt. Herinner dat een fysisch gerealiseerde informatietoestand volgens Chalmers correspondeert met een bepaald effect op een causale keten. Als een systeem een informatieruimte realiseert, komt die ruimte overeen met de causale/functionele organisatie van het systeem. Ieder ander systeem dat causaal/functioneel isomorf is met dit eerste systeem, zal volgens het principe van organisationele invariantie aanleiding geven tot dezelfde ervaring en dezelfde informatieruimte realiseren. Chalmers geeft toe dat dit geen ontegensprekelijke argumenten zijn, en dat ze zijn principe eerder op een indirecte manier steunen.¹³⁴ Uit het volgende citaat is duidelijk dat Chalmers zijn dubbel-aspect principe voor een deel afhangt van een zekere intuïtie en giswerk:

It is natural to suppose that this double life of information spaces corresponds to a duality at a deep level. We might even suggest that this double realization is the key to the fundamental connection between physical processes and conscious experience. We need *some* sort of construct to make the link, and information seems as good a construct as any.¹³⁵

Chalmers denkt echter dat dit dubbel-aspect principe van informatie kan dienen als een kader waarin men een systeem van fundamentele wetten, die de fysische en fenomenale werelden met elkaar linken, verder in detail kan uitwerken.¹³⁶

¹³³ Chalmers, *The Conscious Mind*, 284-285.

¹³⁴ Chalmers, *The Conscious Mind*, 287-88.

¹³⁵ Chalmers, *The Conscious Mind*, 285-286.

¹³⁶ Chalmers, *The Conscious Mind*, 284-86.

Merk op dat we hier niet meer spreken over eenvoudige fysische systemen zoals de hierboven besproken ‘lichtschakelaars’, maar over extreem complexe fysische systemen zoals het brein. De betrokken informatietoestanden en de ruimtes van alle mogelijke informatietoestanden hebben waarschijnlijk opnieuw een extreem complexe structuur. Chalmers is dus van mening dat ervaringen fenomenale realisaties zijn van deze zeer complexe informatietoestanden, die dezelfde informatietoestanden zijn die ook fysisch gerealiseerd worden door bewuste systemen. Het eerder besproken principe van organisationele invariantie stelde dat ervaringen volledig bepaald worden door het abstracte patroon van causale interacties tussen de neuronen van het brein. De informatieruimtes waarvan hier sprake is, zijn dus deze abstracte patronen van causale interactie. We moeten bij een ervaring dus onderscheid maken tussen verschillende zaken. Neem als voorbeeld de ervaring van de geur van een bepaalde stof. Ten eerste is er de fysische of chemische structuur van de stof die de reuk veroorzaakt. Zoals Adam Pautz opmerkt is er geen isomorfisme tussen deze structuur en de quale waarvan sprake is, aangezien we veelal een verschillende geur ervaren bij stoffen met een zeer gelijkaardige chemische structuur.¹³⁷ In zijn latere artikel “The Combination Problem for Panpsychism” stelt Chalmers dat er geen isomorfisme is tussen de macrofysische structuur van het brein en ervaringen. Hij spreekt in dit verband van een “structurele *mismatch*”.¹³⁸ Aan de hand van recente experimentele gegevens kan men dit echter betwisten. Het patroon dat gevormd wordt door het vuren van neuronen in het brein tijdens ervaringen kan men in de neurowetenschappen bestuderen aan de hand van technieken zoals fMRI, waarbij men een opmerkelijke correlatie heeft kunnen vaststellen tussen specifieke patronen van vurende neuronen en specifieke ervaringen. Deze correlatie is zo sterk dat men bijvoorbeeld ook bij verschillende personen, ondanks de individuele verschillen in de structuur van de hersenen, gelijkende patronen kan linken aan gelijkende ervaringen. Hierdoor kan men zelfs met een computer tot op zekere hoogte voorspellen welke ervaring (bijvoorbeeld het zich inbeelden van een vliegtuig of een huis) verschillende proefpersonen ondervinden.¹³⁹ Volgens Chalmers is een ervaring echter niets anders dan de fenomenale realisering van een specifieke abstracte informatietoestand die ook fysisch gerealiseerd is in een abstracte causale structuur tussen de neuronen in ons brein. Hoewel deze informatietoestand te complex is om te modelleren, laat staan te

¹³⁷ Adam Pautz, “What is the Integrated Information Theory of Consciousness? A Catalogue of Questions,” *Journal of Consciousness Studies* 26, No. 1–2 (2019): 201, voetnoot 7.

¹³⁸ David Chalmers, “The Combination Problem for Panpsychism,” in *Panpsychism: Contemporary Perspectives*, ed. Godehard Bruntrup & Ludwig Jaskolla (Oxford: Oxford University Press, 2016), 3 The many combinations problems.

¹³⁹ Philip Kragel et al., “Representation, pattern information, and brain signatures: From neurons to neuroimaging,” *Neuron* 99 (2) (2018): 257-273 en John-Dylan Haynes & Geraint Rees, “Decoding mental states from brain activity in humans,” *Nature Reviews Neuroscience* 7 (2006): 523, 530. Zie ook Pautz, “What is the Integrated Information Theory of Consciousness? A Catalogue of Questions,” 202-203. Voor experimenten en bijhorende bespreking zie <https://www.youtube.com/watch?v=CoF1Bx-gIaw>.

testen, is die visie echter niet noodzakelijk in tegenspraak met de empirische resultaten uit de hierboven beschreven fMRI experimenten. Het lijkt immers plausibel dat het patroon van vurende neuronen dat opgepikt wordt door een fMRI scanner sterk gecorreleerd is met het abstracte patroon van causale interacties tussen neuronen. Indien fMRI scans fijnkorrelig genoeg zouden zijn om alle interacties tussen de neuronen van het brein te kunnen waarnemen, zouden beide patronen mogelijks zelfs identiek zijn.¹⁴⁰ Maar natuurlijk zou een reductionistisch functionalist opnieuw kunnen zeggen dat ervaringen niets anders zijn dan de fysische processen in ons brein. We zien dus dat het dubbel-aspect principe van informatie weliswaar een mogelijke verklaring is voor ervaringen, maar het is niet het enige mogelijke verhaal.

We kunnen ons ook afvragen wat de metafysica is achter dit verhaal. Chalmers spreekt in *The Conscious Mind* in dit verband van twee mogelijke rollen voor informatie. Hij overweegt een zwakkere versie van DAT, waarbij informatie enkel een constructie is om de link te maken tussen fysische processen en bewustzijn, waarbij de laatste twee de enige fundamentele aspecten zijn van de werkelijkheid. Het fenomenale ontstaat uit het fysische volgens fundamentele psychofysische wetten, waarbij informatie een rol speelt in die wetten. Het lijkt er echter op dat Chalmers voorstander is van een sterkere versie van DAT, waarbij informatie gereïficeerd wordt en informatie zelf wordt gezien als fundamenteel. Het fysische en het fenomenale zijn in die sterkere versie van DAT ontologisch afhankelijk van informatie, die fundamenteler is.¹⁴¹ Die theorie is gebaseerd op een reeds vroeger geformuleerde theorie over de aard van de fysische wereld: de “*it from bit*” theorie. Ik zal eerst die laatste bespreken, waarna ik later nog terug kom op de rol van informatie bij DAT.

3.6 *It from bit*

Chalmers liet zich voor DAT inspireren door de “*it from bit*” visie op de fysische werkelijkheid, zoals die onder meer door John Wheeler en Edward Fredkin werd voorgesteld.^{142 143 144} Deze visie is echter eerder een verzameling van verwante visies op de werkelijkheid, waarbij informatie wel steeds een centrale rol speelt. Opnieuw verwijst Chalmers naar een andere auteur maar geeft hij in *The Conscious Mind* een eigen invulling aan het oorspronkelijke idee van Wheeler. Fysische theorieën zijn wiskundige theorieën die de wereld trachten te beschrijven maar niet verklaren hoe de wereld die beschreven wordt werkelijkheid kon worden. Wheeler zocht naar de principes die een antwoord geven op die

¹⁴⁰ Uiteraard gaat dit ver boven onze huidige en hoogstwaarschijnlijk ook toekomstige technische mogelijkheden.

¹⁴¹ Chalmers, *The Conscious Mind*, 286-87.

¹⁴² Chalmers, *The Conscious Mind*, 302-03.

¹⁴³ Wheeler, “Information, physics, quantum: The search for links,”.

¹⁴⁴ Fredkin, “Digital mechanics”.

laatste vraag en ontwikkelde daarom zijn “*it from bit*” theorie. Bij Wheeler is er sprake van een participatorisch universum, waarin waarnemers een fundamentele rol spelen. Volgens hem was de realiteit gegrond in de observaties van waarnemers, die toestanden zijn in het bewustzijn van die waarnemers. Waarnemers stellen daarbij ja/nee-vragen, met meetinstrumenten zoals deeltjesversnellers, waarbij het antwoord gegeven wordt door de fysische wereld die op haar diepste niveau immaterieel is en gegrond is in een verdeling van bits.¹⁴⁵ Chalmers merkt op dat de visie van Wheeler meer leek op een vorm van idealisme, in tegenstelling tot zijn eigen interpretatie van “*it from bit*”.^{146 147}

Chalmers lijkt in *The Conscious Mind* aanvankelijk de meest verregaande vorm van “*it from bit*” te verdedigen, waarin fundamentele fysische eigenschappen zoals massa, lading en fysische wetten af te leiden zijn uit informationele eigenschappen en wetten. Er wordt verondersteld dat men de wetten van de fysica uiteindelijk zou kunnen uitdrukken in termen van informatie, waarbij de fysische wereld bestaat uit toestanden die aanleiding geven tot andere toestanden zonder te zeggen wat die toestanden zijn. Er wordt vervolgens voorgesteld dat die toestanden niets anders zijn dan locaties in een abstracte informatieruimte.¹⁴⁸ Chalmers verwijst zowel in *The Conscious Mind* als in zijn recente boek *Reality+* naar versies van “*it from bit*” die zowel ruimtetijd zien als een primitief kader waarin informatieruimtes zijn ingebed, als versies waarbij ook ruimtetijd samengesteld is uit de relaties tussen informatieruimtes.^{149 150} De werkelijkheid wordt in die visie beschouwd als opgebouwd uit louter informatie. Doorgaans wordt aangenomen dat er een substraat, een substantie of ‘spul’ is dat de ‘drager’ is van eigenschappen. Er is meer dan louter een verzameling van eigenschappen, er is een substantie, het *ding* dat die eigenschappen heeft. “*It from bit*” wordt volgens Chalmers dus opgevat als de visie die claimt dat er niets anders is dan informatie, en dat er uiteindelijk op het diepste niveau van de werkelijkheid geen dingen zijn die deze informatie dragen:

This would lead to a picture of the world as a world of *pure* information. (...) a picture of the world as pure informational flux, without any further substance to it.¹⁵¹

¹⁴⁵ Wheeler, “Information, physics, quantum,” 309.

¹⁴⁶ Chalmers, *The Conscious Mind*, 302, 388, voetnoot 4.

¹⁴⁷ Het idee van een participatorisch universum is een piste die door Chalmers recentelijk opnieuw verkend werd. Zie “Consciousness and the Collapse of the Wave Function”, lecture, Göttingen, May 2014, https://www.youtube.com/watch?v=DIBT6E2GtjA&ab_channel=GeorgRK.

¹⁴⁸ Chalmers, “Facing up to the problem of consciousness,” 216.

¹⁴⁹ Chalmers, *The Conscious Mind*, 302-03.

¹⁵⁰ Chalmers, *Reality+, virtual worlds and the problems of philosophy*, 160-62.

¹⁵¹ Chalmers, *The Conscious Mind*, 303.

The universe at its bottom level is a universe of pure differences.¹⁵²

Volgens die visie is de wereld niets meer dan een verzameling van primitieve verschillen, aangevuld met causale en dynamische relaties tussen die verschillen.¹⁵³

Terwijl Chalmers zich naar eigen zeggen laat inspireren door "*it from bit*", ziet hij twee belangrijke problemen met deze visie. Ten eerste merkt hij op dat een ervaring meer lijkt te zijn dan een pure informatietoestand. Ervaringen hebben een specifieke intrinsieke natuur, die niet in rekening wordt gebracht door hun locatie in een informatieruimte. Een tweede probleem is dat het beeld van de wereld, die op haar meest fundamentele niveau enkel uit primitieve verschillen bestaat, velen niet coherent zullen vinden. Een dergelijk beeld lijkt te moeten worden aangevuld met een onderliggende kwaliteit, waarin die verschillen kunnen gegrond worden. Het lijkt immers plausibel dat een verschil ook steeds een verschil *in iets* moet zijn. In de "*it from bit*" versie die Chalmers schetst zijn informatietoestanden het enige dat door de causale relaties uit de fysica met elkaar verbonden wordt. Maar abstracte informatietoestanden lijken niet iets te zijn dat causaal aan elkaar kan gebonden worden. Abstracte informatietoestanden als fundament voor de wereld zijn nog om een andere reden problematisch. Chalmers schetst een denkbeeldige situatie van twee fundamentele fysische eigenschappen, die in verschillende natuurwetten een rol spelen, die dezelfde informationele structuur hebben, bijvoorbeeld als deze beide kwantiteiten zijn die voorgesteld worden door een reëel getal in het continuüm. Als de fysische wereld enkel uit informatie bestaat, lijkt het problematisch hoe we deze beide informatieruimtes, die in verschillende natuurwetten een rol spelen, van elkaar kunnen onderscheiden. Het lijkt alsof we op één of andere manier deze informatieruimtes een intrinsieke kwaliteit moeten kunnen geven om hen van elkaar te kunnen onderscheiden, te individualiseren.¹⁵⁴

3.7 Russelliaans monisme

Chalmers overloopt verschillende mogelijke strategieën om dit te doen. Zo kan men intrinsieke eigenschappen postuleren om informatietoestanden te gronden, terwijl men als oplossing voor het eerste probleem onafhankelijk hiervan fenomenale eigenschappen kan introduceren die wetmatig gelinkt zijn aan die informatietoestanden. De beste oplossing lijkt voor Chalmers echter iets dat sterk doet denken aan het monisme van Bertrand Russell. Chalmers interpreteert het monisme van Russell als de visie waarbij het fysische en het fenomenale twee verschillende aspecten zijn van één omvattend

¹⁵² Chalmers, *Reality+, virtual worlds and the problems of philosophy*, 165.

¹⁵³ Chalmers, *The Conscious Mind*, 303.

¹⁵⁴ Chalmers, *The Conscious Mind*, 303-304.

iets, waarbij die aspecten analoog zijn aan materie en energie in de fysische wereld.¹⁵⁵ ¹⁵⁶ Lading of massa hebben in die visie een intrinsieke, essentiële natuur die verborgen blijft voor fysische theorieën, die fysische deeltjes en eigenschappen enkel extrinsiek karakteriseren. Fundamentele entiteiten worden in de fysica immers enkel relationeel gekarakteriseerd aan de hand van causale en andere relaties, zoals de manier en de mate waarin ze interageren met andere deeltjes. Er worden wel eigenschappen zoals massa en lading geponeerd, maar die drukken enkel uit in welke mate deeltjes daardoor vatbaar zijn voor krachtvelden en bepalen hoe ze bewegen als ze in dergelijke krachtvelden terecht komen. Het zijn ‘verschillen die een verschil uitmaken’, die volgens Chalmers kunnen gezien worden als een specifieke informatietoestand, die deel uitmaakt van een informatieruimte die men kan linken aan bijvoorbeeld massa. Fenomenale eigenschappen of ‘protofenomenale’ eigenschappen (die zelf niet fenomenaal zijn maar samen op één of andere manier de fenomenale eigenschappen opleveren) zijn volgens het Russelliaans monisme verborgen intrinsieke eigenschappen (*‘quiddities’*) die de essentiële natuur uitmaken van lading of massa en de ‘dragere’ zijn van de extrinsieke eigenschappen die bepalend zijn voor fysische processen.¹⁵⁷ In die visie zouden het dus in werkelijkheid de intrinsieke eigenschappen zijn die gerelateerd zijn via causale/dynamische wetten en op die manier de extrinsieke fysische eigenschappen als het ware zouden “realiseren”, en zo aanleiding zouden geven tot de fysische wetten die door de fysica worden beschreven.¹⁵⁸

Op dezelfde manier zouden bij DAT de informatietoestanden, die zorgen voor de ‘verschillen die een verschil uitmaken’, die nodig zijn voor de causale relaties in de fysica, kunnen gegrond worden in fenomenale of protofenomenale eigenschappen, zodat we geen ongegronde primitieve verschillen moeten veronderstellen. Tegelijk bekomen we zo de fenomenale eigenschappen waarmee we zo vertrouwd zijn. Een informatietoestand, die fysisch gerealiseerd is in een fysische eigenschap zoals lading of massa, is op die manier telkens ook gegrond in een fenomenale of protofenomenale eigenschap (Chalmers noemt dit kortweg een “microfenomenale” eigenschap). De primitieve verschillen die aan de grondslag liggen van de fysische realiteit zijn op die manier gelijk aan deze microfenomenale verschillen. De uiteindelijke ontologie die Chalmers voorstelt met zijn dubbel-aspect theorie is er een van de fysische werkelijkheid waarin enkel de causale relaties tussen informatietoestanden belangrijk zijn, terwijl de fenomenale aspecten van de wereld enkel betrekking hebben op de intrinsieke aard van diezelfde informatietoestanden. Ervaring wordt op die manier

¹⁵⁵ Chalmers, *The Conscious Mind*, 129.

¹⁵⁶ Er is discussie rond de vraag of het neutraal monisme van Russell kan geïnterpreteerd worden als één theorie, of eerder een opeenvolging van gerelateerde maar verschillende theorieën. Zie Leopold Stubenberg & Donovan Wishon, "Neutral Monism," *The Stanford Encyclopedia of Philosophy*, Spring 2023, 4.3 Bertrand Russell. Daarom zal ik hier de interpretatie volgen van Chalmers, zie Chalmers, *The Conscious Mind*, 153-155.

¹⁵⁷ Chalmers, *The Conscious Mind*, 135-136, 302.

¹⁵⁸ Chalmers, *The Conscious Mind*, 155.

gesitueerd *binnenin* het causale netwerk dat de fysica beschrijft. Chalmers formuleert deze visie ook nog in sloganvorm: “(...) Ervaring is informatie van binnenuit, fysica is informatie van de buitenkant”.¹⁵⁹

3.8 De rol van informatie bij DAT

We hebben eerder reeds gezien dat informatie in de dubbel-aspect theorie twee mogelijke rollen kan spelen. Enerzijds overweegt Chalmers een zwakkere versie van DAT, waarbij informatie enkel een constructie is om de link te maken tussen fysische processen en bewustzijn, waarbij de laatste twee de enige fundamentele aspecten zijn van de werkelijkheid. Het fenomenale ontstaat uit het fysische volgens fundamentele psychofysische wetten, waarbij informatie een rol speelt in die wetten. Maar we zien dat Chalmers eerder een meer fundamentele rol ziet voor informatie. In een sterkere versie van DAT is informatie meer dan enkel maar een eenvoudige en elegante link tussen het fenomenaal en het fysische domein. Uit zijn bespreking van “*it from bit*” wordt duidelijk wat voor Chalmers informatie uiteindelijk is: een puur verschil op het diepste niveau van de werkelijkheid. Maar informatie alleen is onvoldoende om een wereld te realiseren. Chalmers zegt wel dat hij zich baseert op “*it from bit*”, maar uiteindelijk is de wereld die dit oplevert te abstract en te ‘ijl’ om een concrete werkelijkheid te zijn. Daarvoor is er uiteindelijk toch iets nodig om een zekere ‘substantie’ te geven aan de werkelijkheid. De rol van fenomenale eigenschappen in het Russelliaans beeld van de wereld waarvan Chalmers gebruik maakt, lijkt te veronderstellen dat die microfenomenale eigenschappen onontbeerlijk zijn voor een wereld, een noodzakelijke voorwaarde om concreet te kunnen bestaan. Dit is gelijkaardig met de visie van Luciano Floridi, die de “*it from bit*” visie bekritiseert, maar zelf een voorstander is van een verdere ontwikkeling van “*it from bit*”, het informationeel structureel realisme (ISR), een variant van het ontisch structureel realisme (zie verder). Volgens Floridi heeft informatie steeds direct ook een betekenis, en is er bij informatie al sprake van interpretatie. Als we het semantische luik van informatie losmaken, bekomen we *data*, die Floridi ook ziet als verschillen of het ontbreken van uniformiteit in de werkelijkheid of “*fractures in the fabric of Being*”.¹⁶⁰ Het feit dat Floridi bij die beschrijving van data de frase “*fabric of Being*” nodig heeft, verraadt natuurlijk dat er toch *iets* nodig is om dat verschil in te gronden, net zoals bij Chalmers uiteindelijk het geval blijkt.

Het door Chalmers geschetste beeld van de werkelijkheid, bestaande uit primitieve verschillen die aanleiding geven tot informatieruimtes, die gegrond worden in microfenomenale eigenschappen, samen met causale relaties tussen die informatieruimtes, heeft een zekere intuïtieve aantrekkingskracht. We kunnen ons immers slechts twee soorten eigenschappen indenken:

¹⁵⁹ Chalmers, *The Conscious Mind*, 303-05.

¹⁶⁰ Luciano Floridi, *The Philosophy of Information* (New York: Oxford University Press Inc., 2011), 85.

eigenschappen die te maken hebben met externe relaties tussen dingen, en intrinsieke relaties hoe iets *voelt van binnenuit*. Uit de beschrijving die Chalmers geeft van zijn voorstel zouden we kunnen denken dat hier een ontologische prioriteit speelt, waarbij er eerst de informatie of primitieve verschillen nodig zijn en vervolgens de (proto)fenomenale eigenschappen die nodig zijn om de informatieruimtes te gronden. Maar misschien is het beter om geen ontologische prioriteit te veronderstellen, maar ons voor te stellen dat zowel de primitieve verschillen als de proto(fenomenale) eigenschappen, alsook de causale relaties, allemaal dezelfde logische prioriteit hebben. Wanneer Luciano Floridi de logische ‘rangorde’ van relaties en relata bespreekt, komt hij ook tot de conclusie dat dit een plausibele oplossing zou zijn: verschillen en hetgeen verschillend is, zijn als de twee zijden van een blad papier, als een “pakket”:

(...) basic entities and structures, relata and relations, simply co-exist as a package: they make each other possible and one cannot choose which one to have: it is all or nothing.¹⁶¹

Het is wel zo dat Chalmers het steeds enkel heeft over eigenschappen om ‘substantie’ te geven aan zijn visie van de werkelijkheid. Hij spreekt niet van een substraat als drager van die eigenschappen. Een echt substraat ontbreekt hier, waardoor er wordt verondersteld dat de bouwstenen van de wereld eerder bundels van eigenschappen zijn. Hoewel Chalmers hier niet expliciet over is, lijkt hij daarmee eerder de bundeltheorie te onderschrijven, in plaats van de conceptie van een substraat of ‘*thin particular*’. De eigenschappen worden daarbij waarschijnlijk eerder verondersteld als individuele instantiaties van eigenschappen (tropes).¹⁶² Dit is natuurlijk een stevig aanneme, niet iedereen zou deze visie van bundels van eigenschappen als bouwstenen voor de werkelijkheid plausibel of zelfs coherent vinden. Wat vooral problematisch lijkt is echter het beeld van fenomenale of protofenomenale eigenschappen die fysische eigenschappen gronden. Het lijkt een zeer vreemde veronderstelling dat iets als ‘roodheid’ als ‘drager’ zou kunnen dienen voor de fysische eigenschappen van fundamentele deeltjes. Het idee is misschien op zich niet incoherent, maar het lijkt niet erg plausibel. Misschien moeten we dit eerder zien als een noodzakelijk, wetmatig samen voorkomen van fysische en fenomenale eigenschappen, als een soort van bruut feit van de werkelijkheid. Anderzijds weet ik ook niet of het idee van een substraat als drager van eigenschappen coherent is. Bij een dergelijk substraat zou men immers één voor één de eigenschappen kunnen wegdenken. De vraag is wat een dergelijk substraat dan uiteindelijk is. Is een substraat een soort van ‘naakte’ entiteit? Of kan men niet alle eigenschappen van een substraat wegdenken, zijn eigenschappen en een substraat ook een ‘pakket’?

¹⁶¹ Floridi, *The Philosophy of Information* (New York: Oxford University Press Inc., 2011), 354-55.

¹⁶² Howard Robinson, “Substance,” *The Stanford Encyclopedia of Philosophy*, 2021, 3.2 Bundle theories versus substratum and ‘thin particulars’.

Het idee van protofenomenale eigenschappen is ook heel vreemd. Hoe moeten we ons voorstellen dat protofenomenale eigenschappen op een of andere manier combineren tot fenomenale eigenschappen? Dit lijkt op het poneren van iets dat men niet begrijpt om iets te construeren dat men niet begrijpt.

Misschien kunnen we ons proberen een beeld van dit alles te vormen door de wereld te zien als opgebouwd uit de primitieve verschillen waaruit volgens Chalmers informatie bestaat, die aanleiding geven tot eigenschappen. Dit alles doet denken aan een soort van platoonse visie, waarbij de wereld samengesteld is uit eigenschappen, of zelfs kan beschouwd worden als één gigantische samengestelde eigenschap, in de zin dat sneeuw wit en gras groen is, Antwerpen ten noorden van Brussel ligt, enz... Op die manier zou de wereld *een staat van zaken* zijn, een *manier waarop de dingen* zouden kunnen zijn.¹⁶³ We zouden ons verder kunnen voorstellen dat alle relaties, structuren, causale relaties en zelfs ook ruimte opgebouwd zijn uit dergelijke verschillen, al dan niet in tijd, waarbij tijd ook als opgebouwd uit verschillen wordt gezien. Misschien moeten we dit beschouwen in het kader van een modaal realistische ontologie. Met vele, mogelijk oneindig veel, werelden die verschillen in natuurwetten, maar bijvoorbeeld ook verschillen in aantal entiteiten die deze werelden bevatten, verschillen in manieren van zijn.¹⁶⁴ Dit alles is natuurlijk niet meer dan een intuïtief idee, een heel speculatief beeld van hoe de wereld zou kunnen zijn. Maar het is het beeld dat bij mij in mijn verbeelding ontstaat bij de beschrijving van Chalmers van de wereld als een ‘wereld van verschillen’.

3.9 Gelijkenissen en verschillen tussen DAT en monisme

Chalmers merkt op dat het verschil tussen Russelliaans monisme en eigenschapsdualisme klein zou zijn: fenomenale eigenschappen zouden nog steeds fundamenteel zijn en niet vastgelegd door de eigenschappen die de fysica blootlegt. Het verschil zou eerder semantisch dan metafysisch zijn: het dualisme tussen fysische en niet-fysische eigenschappen zou vervangen worden door het dualisme tussen extrinsieke eigenschappen van de fysische realiteit, die we te weten kunnen komen en als het

¹⁶³ Deze beelden van een gigantische, complexe eigenschap die alle eigenschappen van de wereld bevat en de wereld als een mogelijke staat van zaken ziet, leen ik van Mark Balaguer in Mark Balaguer, “Platonism in Metaphysics,” *The Stanford Encyclopedia of Philosophy*, 2016, 4.5 Possible Worlds.

¹⁶⁴ Dit geheel zou interessante eigenschappen hebben met betrekking tot informatie. In zekere zin zou het geheel van alle mogelijke manieren waarop een wereld zou kunnen bestaan alle informatie, maar ook geen enkele informatie, bevatten, want het geheel van alle mogelijke informatie bevat geen enkele specificiteit meer, dus de Shannon informatie is nul. De situatie zou analoog zijn aan de situatie waarbij iemand je vraagt waar je bent, en je als antwoord een opsomming van alle mogelijke plaatsen geeft waar je zou kunnen zijn. Of om een beeld te gebruiken van Rudy Rucker: bevat een bibliotheek met elk mogelijk boek met elke mogelijke willekeurige opeenvolging van tekens wel enige informatie? Het beeld van een oneindige bibliotheek leen ik uit Rudy Rucker, *Oneindigheid: Filosofie en wetenschap van het oneindige* (Amsterdam: Uitgeverij Contact, 1986), 124-9.

ware “verborgen” intrinsieke eigenschappen, terwijl de beide echter fundamentele eigenschappen zijn.^{165 166}

Chalmers stelt zich verder de vraag of, aangezien de fysica de wetenschap is die fundamentele eigenschappen bestudeert, we de ‘fysica’ en wat fysisch is niet beter ruimer zouden interpreteren, waarbij we de fundamentele fenomenale eigenschappen ook als fysische eigenschappen zouden beschouwen. Op die manier zouden we komen tot een soort van materieel monisme, verschillend van Cartesiaans dualisme en het eigenschapsdualisme dat Chalmers verdedigt. Er zijn echter redenen om fenomenale eigenschappen niet te beschouwen als fysische eigenschappen. De huidige wetenschap schetst een wereld die, voor zover we weten, causaal gesloten is. Dit laatste wil zeggen dat iedere fysische gebeurtenis een fysische oorzaak heeft, terwijl er geen sprake is van niet-fysische oorzaken van fysische gebeurtenissen. Tevens gaan het overgrote deel van de fysici en de meeste filosofen er (al dan niet impliciet) van uit dat een toekomstige fysica, althans in principe, in staat zou moeten zijn om een gesloten, consistente theorie over de externe wereld te formuleren.¹⁶⁷ Daardoor gaat men er in de fysica meestal van uit dat we bij het construeren van een theorie over de buitenwereld geen fenomenale eigenschappen of psychofysische wetten nodig hebben. Als men het voorstelbaarheid argument volgt en de mogelijkheid van een zombiewereld zonder fenomenale eigenschappen aanvaardt, zijn fenomenale eigenschappen in een zekere zin overbodig voor de fysica zoals ze nu wordt bedreven. Tenzij natuurlijk het voorstelbaarheid argument onjuist is en de fysica nog sterk evolueert, wat natuurlijk zeer goed mogelijk is.

Fysische eigenschappen van fysische entiteiten kan men bestuderen aan de hand van de causale relaties tussen die entiteiten. Het is niet mogelijk om fenomenale eigenschappen direct te observeren in experimenten, hoewel we wel in staat zijn om data over de relatie tussen fysische processen en ervaring te halen uit onze eigen subjectieve, eerste-persoons ervaring en uit experimenten waarbij een proefpersoon rapporteert over zijn of haar ervaringen. Daarom is het volgens Chalmers meer vanzelfsprekend om een onderscheid te blijven maken tussen fenomenale en fysische eigenschappen, alsook tussen fysische en psychofysische wetten.¹⁶⁸

¹⁶⁵ Chalmers, *The Conscious Mind*, 136.

¹⁶⁶ Chalmers, “Moving Forward on the Problem of Consciousness,” 3.4. Epiphenomenalism and interactionism.

¹⁶⁷ Chalmers gaat op dit punt enigszins wat kort door de bocht en stelt dat “de fysica een gesloten, consistente theorie vormt”. Hier ga ik er met enige welwillendheid van uit dat hij eerder de hierboven meer genuanceerde formulering bedoelt. Zie Chalmers, *The Conscious Mind*, 128.

¹⁶⁸ Chalmers, *The Conscious Mind*, 128-129.

3.10 Ontisch structureel realisme

In het recente *Reality+* bespreekt Chalmers zowel “*it from bit*” als een recentere, meer algemene versie, het ontisch structureel realisme (OSR). Ontologisch structureel realisme stelt dat de fysische wereld volledig bestaat uit wiskundige structuren. In tegenstelling tot de “*it from bit*” visie hoeven die structuren echter niet digitaal te zijn.¹⁶⁹ Digitale data kunnen enkel een aantal discrete waarden aannemen, terwijl er bij analoge data tussen twee mogelijke waarden nog een theoretisch oneindig aantal tussenwaarden zit. Analoge representaties kunnen betrekking hebben op entiteiten die discreet van aard zijn, maar kunnen, in tegenstelling tot digitale representaties, ook continue structuren representeren.¹⁷⁰ OSR kan dus betrekking hebben op zowel een wereld die fundamenteel continu van aard is als op een wereld die eerder discreet van aard is. Deze visie kan men dus ook verdedigen als de fysische realiteit continu zou zijn in plaats van digitaal. Chalmers had misschien al deze veralgemening van “*it from bit*” in gedachten toen hij *The Conscious Mind* schreef. Zoals we gezien hebben liet Chalmers immers zowel differentierelaties toe met een oneindig aantal informatietoestanden (analoog aan het continuüm van reële getallen) als informatietoestanden met een continue interne structuur die een oneindig aantal elementen hebben.¹⁷¹ Chalmers noemt OSR ook een “*pure it-from-structure*” visie.¹⁷² In zekere zin gaat OSR nog wat verder dan “*it from bit*”, want volgens Chalmers is het plausibel dat noties zoals natuurwetten, causaliteit en ‘concreet bestaan’ ook structurele noties zijn in een brede zin. Die noties gaan verder dan een puur wiskundige beschrijving van de wereld, maar kunnen volgens Chalmers toch deel uitmaken van een theorie zoals OSR. Chalmers geeft geen verdere uitleg over hoe die structurele karakterisering van deze noties zou moeten gebeuren. Hij maakt hier denk ik eerder gebruik van de intuïtie van de lezer om deze te zien als structureel “in een brede zin”. Maar hoe dan ook, de noties zoals natuurwetten, causaliteit en ‘concreet bestaan’ die Chalmers vermeldt zijn niet volledig wiskundige noties en zijn dus een ‘extra’ bovenop een karakterisering van de wereld via OSR.¹⁷³ Chalmers merkt bij OSR echter hetzelfde probleem op als bij “*it from bit*” in *The Conscious Mind*. Hij stelt met name de vraag of een wereld wel volledig wiskundig van aard kan zijn. Hij verwijst daarbij naar de aanhangers van het pythagorisme, die dachten dat alles gemaakt is uit getallen, en de meer recente “*Mathematische Universum Hypothese*” (MUH) van Max Tegmark, die stelt dat onze fysische werkelijkheid een wiskundige structuur is.¹⁷⁴ Volgens Tegmark bestaan er oneindig veel mogelijke werelden, die telkens een (andere) wiskundige structuur zijn. Een volledige beschrijving van die

¹⁶⁹ Chalmers, *Reality+, virtual worlds and the problems of philosophy*, 404-406.

¹⁷⁰ Jack Copeland, “The Modern History of Computing,” *The Stanford Encyclopedia of Philosophy*, 2020, Analog Computers.

¹⁷¹ Chalmers, *The Conscious Mind*, 278-279.

¹⁷² Chalmers, *Reality+, virtual worlds and the problems of philosophy*, 406.

¹⁷³ Chalmers, *Reality+, virtual worlds and the problems of philosophy*, 411.

¹⁷⁴ Chalmers, *Reality+, virtual worlds and the problems of philosophy*, 407.

wiskundige structuur die onze fysische werkelijkheid is, vergt alleen maar wiskundige vergelijkingen. Concepten zoals ‘deeltjes’ of ‘observatie’ zijn slechts overtollige “menselijke bagage”.¹⁷⁵ Tegmark meent zelfs ook een — weliswaar naïef aandoende — verklaring te hebben voor het bestaan van die werelden. Hij verwijst daarvoor naar de formalisten, de wiskundige stroming die stelt dat wiskundige structuren niet gecreëerd worden of ‘ergens’ bestaan, maar eenvoudigweg bestaan. Aangezien wiskundig bestaan en fysisch bestaan volgens hem equivalent zijn, bestaan alle wiskundige structuren en dus alle mogelijke (wiskundige) werelden. Tegmark beschouwt zijn MUH als een wiskundige versie van het modaal realisme van David Lewis, of als een radicale versie van het platonisme, waarbij de ideeënwereld ook fysisch werkelijk is.¹⁷⁶

Chalmers draait de redenering van Tegmark enigszins om en stelt dat deze visie juist problematisch is omdat wiskundige entiteiten “te makkelijk” bestaan. Als wiskundige entiteiten bestaan in één wereld, bestaan ze in alle werelden, wat volgens hem als gevolg zou hebben dat bijvoorbeeld de wiskundige theorie van Newton ook correct zou zijn in een wereld waar de relativiteitstheorie geldig is, wat problematisch is. De oplossing bestaat er volgens Chalmers uit om onderscheid te maken tussen het abstracte bestaan van wiskundige objecten en het concrete bestaan van fysische objecten zoals velden en deeltjes. Opnieuw verwijst hij daarbij naar de noodzaak van het hebben van causale krachten om als fysisch object concreet te kunnen bestaan. Getallen beschikken immers, net als abstracte informatieruimtes, niet over concrete krachten.¹⁷⁷ Chalmers stelt in *Reality+* ook opnieuw dat ervaringen wel wiskundig beschreven kunnen worden (waarbij hij refereert naar het voorbeeld van de ervaring van verschillende kleuren en de onderlinge verschillen tussen die kleuren), maar hij herhaalt opnieuw zijn mening dat bewustzijn een bijzonder statuut heeft en meer is dan een wiskundige beschrijving:

But conscious experience seems to go beyond a mathematical description.¹⁷⁸

Gualtiero Piccinini en Neal Anderson (beide critici van de “it from bit” visie) geven een gelijkaardige kritiek op de visie van Tegmark, die volgens hen een categorie vergissing begaat. Als we het platonisme over wiskunde als vertrekpunt nemen, zijn wiskundige entiteiten immers abstracte entiteiten, die niet voorkomen in de ruimte of tijd, en niet deelnemen aan causale of dynamische relaties. Wiskundige entiteiten veranderen niet. Wat Tegmark “overtollige bagage” noemt is het enige wat de wiskundige beschrijvingen in onze wetenschappelijke theorieën verankert in de fysische wereld.¹⁷⁹ Tegmark begaat

¹⁷⁵ Max Tegmark, "The Mathematical Universe," *Foundations of Physics* 38 (2) (2008): 102.

¹⁷⁶ Max Tegmark, "The Mathematical Universe," *Foundations of Physics* 38 (2) (2008): 125.

¹⁷⁷ Chalmers, *Reality+, virtual worlds and the problems of philosophy*, 408-09.

¹⁷⁸ Chalmers, *Reality+, virtual worlds and the problems of philosophy*, 411.

¹⁷⁹ Piccinini & Anderson, "Ontic Pancomputationalism," 35-6.

volgens beide auteurs de fout dat hij uit het isomorf zijn van twee systemen (een fysisch systeem en de abstracte wiskundige structuur die men kan linken aan de wiskundige beschrijving van het fysisch systeem éénmaal die beschrijving ontdaan is van alle “bagage”) meent te kunnen besluiten dat die structuren volledig gelijkend zijn. Uit die volledige gelijkenis zou dan volgen dat de fysische werkelijkheid een wiskundige structuur is, waarbij geen rekening wordt gehouden met het ontbreken van zaken zoals kwalitatieve aspecten, causale krachten en intrinsieke eigenschappen.¹⁸⁰

Chalmers herhaalt in *Reality+* echter niet zijn oplossing die gebaseerd is op het Russelliaanse monisme, waarin informatie uiteindelijk gegrond wordt in intrinsieke, fenomenale eigenschappen. Maar verschillende passages uit *Reality+* doen sterk vermoeden dat hij toch nog speelt met dit idee, al lijkt hij in zijn meest recente boek toch ook te flirten met het idee dat bewustzijn toch in puur structurele termen zou kunnen verklaard worden:

But as things stand, there are strong reasons to think consciousness goes beyond structure. At least, to fully explain consciousness in terms of pure structure would take a scientific revolution or two. That said, there have been scientific revolutions before.¹⁸¹

Maar als Chalmers ervaringen toch als fundamenteel beschouwd en die opnieuw ziet als een manier om een puur abstracte (en daarom problematische) kijk op de wereld zoals OSR een zekere “substantie” te geven, denk ik dat hij daarmee uiteindelijk uitkomt op een versie van soort van wat hij in *Reality+* een “it-from-structure-from-it” visie noemt. Hij bespreekt dit inderdaad kort als een mogelijkheid, doch verbindt deze niet expliciet met de manier waarop hij in *The Conscious Mind* “it from bit” interpreteerde.¹⁸²

3.11 Is dit epifenomenalisme?

Chalmers verkent in zijn boek *The Conscious Mind* verschillende metafysische visies in verband met causaliteit, maar maakt geen definitieve keuze tussen de veelal sterk verschillende alternatieven die hij opsomt, ook al omdat hij van mening is dat de metafysica van causaliteit nog steeds niet opgehelderd is.¹⁸³ Ook in het latere “Moving Forward on the Problem of Consciousness” beschouwt hij zowel epifenomenalisme, alsook meer recente versies van dualistisch interactionisme en Russelliaans monisme als mogelijke opties.¹⁸⁴ Indien bewustzijn enkel natuurlijk superveniënt is op fysische eigenschappen, lijkt

¹⁸⁰ Piccinini & Anderson, “Ontic Pancomputationalism,” 35-6.

¹⁸¹ Chalmers, *Reality+, virtual worlds and the problems of philosophy*, 438.

¹⁸² Chalmers, *Reality+, virtual worlds and the problems of philosophy*, 406.

¹⁸³ Chalmers, *The Conscious Mind*, 153.

¹⁸⁴ Chalmers, “Moving Forward on the Problem of Consciousness,” 3.4 Epiphenomenalism and interactionism.

het erop dat qualia geen causale rol hebben. Men gaat er immers meestal van uit dat de fysische wereld causaal gesloten is. Als het nu mogelijk is om al het fenomenale uit onze wereld weg te denken terwijl de fysische wereld causaal gesloten blijft, dan volgt dat alles in onze wereld een causale verklaring heeft die onafhankelijk is van het fenomenale.¹⁸⁵ Velen vinden echter het hieruit volgend epifenomenalisme contra-intuïtief, onaantrekkelijk of zelfs afstotelijk. Chalmers vindt dat epifenomenalisme weliswaar minder plausibel is omdat het tegen onze intuïtie ingaat, maar hij vindt dit ook geen voldoende reden om epifenomenalisme te verwerpen. Het idee dat bewustzijn een causale rol speelt is heel intuïtief omdat fenomenaal bewustzijn veelal gevolgd wordt door fysische gebeurtenissen. Door dat samengaan veronderstellen we een causaal verband, maar een echt bewijs voor de causaliteit van bewustzijn wordt hierdoor niet geleverd. Epifenomenalisme levert wel een onelegant beeld van de natuur, omdat bewustzijn in deze visie lijkt te ‘bengelen’ aan fysische processen als een soort bijkomstigheid.¹⁸⁶

Chalmers overweegt verschillende pistes waarbij de causale geslotenheid van de fysische wereld wordt geschonden, maar lijkt toch eerder te willen vasthouden aan dit beginsel. Eerder lijkt Chalmers de voorkeur te geven aan een piste waar een causale rol voor qualia verzoend wordt met de causale geslotenheid van de fysische wereld. Hiervoor kijkt Chalmers opnieuw naar het Russelliaans monisme. Doordat de intrinsieke eigenschappen die de ‘dragers’ zijn van de extrinsieke eigenschappen zelf fenomenale of proto-fenomenale eigenschappen zijn, wordt bewuste ervaring *binnen* het causale netwerk gelegd dat door de fysica beschreven wordt, en dus niet als een louter bijkomstigheid die als het ware aan die fysische wereld bengelt. Het resultaat is een zwakkere vorm van epifenomenalisme (alternatief zou men kunnen spreken van een subtiele causale rol voor fenomenale eigenschappen), waarin toch de causale geslotenheid van de fysische wereld behouden blijft. Volgens Chalmers levert dit een veel intuïtiever beeld van de fysische wereld op. Zonder intrinsieke eigenschappen die een interne natuur bepalen van de basisentiteiten, waarmee fysische theorieën werken, blijven die theorieën beperkt tot het formuleren van causale verbanden tussen die basisentiteiten, die voor de rest ‘leeg’ blijven:

The picture of the physical world that this yields is that of a giant causal flux, but the picture tells us nothing about what all this causation relates.¹⁸⁷

De eigenschappen waarmee men elementaire deeltjes karakteriseert, zoals lading, massa of spin geven immers enkel informatie over hoe de deeltjes met andere deeltjes interageren of reageren op krachten en dergelijke. De keuze voor fenomenale of profenomenale eigenschappen die, eventueel samen met andere voorkomende intrinsieke eigenschappen, enige ‘substantie’ geven aan deze basisentiteiten

¹⁸⁵ Chalmers, *The Conscious Mind*, 150.

¹⁸⁶ Chalmers, “Moving Forward on the Problem of Consciousness,” 3.4 Epiphenomenalism and interactionism

¹⁸⁷ Chalmers, *The Conscious Mind*, 153.

ligt voor de hand omdat dit de enige intrinsieke, niet-relatieve eigenschappen zijn met wie we direct bekend zijn.¹⁸⁸

Voor de volledigheid vermeld ik dat Chalmers in *The Conscious Mind* op een bepaald moment ook interactief dualisme overweegt, maar hij is geen voorstander van deze visie. Met de mogelijke uitzondering van kwantum onbepaaldheden is de fysische wereld causaal gesloten en is er geen causale rol weggelegd voor een niet-fysische geest.¹⁸⁹ Chalmers kijkt meer naar de hedendaagse versies van interactionistisch dualisme, waarin bepaalde eigenschappen van de kwantummechanica een rol spelen. Zo zijn er versies waarin bewustzijn verantwoordelijk is voor het bepalen van bepaalde fysische waarden in een schijnbaar probabilistische verdeling of waar bewustzijn de ineenstorting van de golfvergelijking teweeg brengt.^{190 191} Ik zal deze theorieën, die het verband leggen met kwantumtheorieën echter niet verder behandelen aangezien ze geen sterk verband hebben met de onderzoeksvraag.

3.12 Panpsychisme

Chalmers merkt in *The Conscious Mind* zelf op dat het dubbel-aspect principe ondergedetermineerd is en een aantal vragen onbeantwoord laat. Informatie is immers alomtegenwoordig, zo somt Chalmers voorbeelden op variërend van CD-spelers tot automotoren. Hij zegt dit niet expliciet, maar hieruit volgt dat informatietoestanden en informatieruimtes overal in het fysische domein voorkomen, zowel op microscopisch niveau, op het niveau van microchips in computers of neuronen in hersenen, als op alle niveaus daartussen en daarboven. Daaruit rijst de vraag of alle informatie die fysisch gerealiseerd wordt, ook fenomenaal gerealiseerd wordt. Indien zo, dan komt bewustzijn veel meer voor dan we meestal aannemen, en hebben niet alleen eenvoudige organismen ervaringen maar zijn zelfs de meest eenvoudige informatie verwerkende systemen zoals een thermostaat gelinkt aan een zekere ervaring. Dit zou leiden tot panpsychisme, de visie dat ieder fysiek ding geassocieerd is met bewustzijn.^{192 193} De ervaring die gelinkt is aan een thermostaat zou hoogstwaarschijnlijk uiterst eenvoudig zijn, zonder de

¹⁸⁸ Chalmers, *The Conscious Mind*, 153-54.

¹⁸⁹ Chalmers, *The Conscious Mind*, 124-25.

¹⁹⁰ Chalmers, *The Conscious Mind*, 156-58.

¹⁹¹ Chalmers, "Moving Forward on the Problem of Consciousness," 3.4 Epiphenomenalism and interactionism.

¹⁹² Hedda Mørch, "Is the Integrated Information Theory of Consciousness Compatible with Russellian Panpsychism?," *Erkenntnis* 84 (5) (2019): 1065.

¹⁹³ Er zijn verschillende vormen van panpsychisme. De vorm waarvan hier sprake is staat bekend als onzuiver panpsychisme. In tegenstelling tot zuiver panpsychisme, dat veronderstelt dat de fysische wereld enkel bestaat uit fenomenale eigenschappen, stelt onzuiver panpsychisme dat de fysische wereld naast fenomenale eigenschappen ook bestaat uit fundamentele, niet-fenomenale relaties, zoals spatiotemporele of causale relaties. Zie Hedda Mørch, "Is the Integrated Information Theory of Consciousness Compatible with Russellian Panpsychism?," 1070, voetnoot 8.

complexe structuur die kenmerkend is voor onze ervaringen.¹⁹⁴ Die ervaring is waarschijnlijk zo eenvoudig dat er geen analoog is in onze ervaring. Chalmers gaat er uiteraard niet van uit dat thermostaten zelfbewust of intelligent zouden zijn, of in staat om te denken of concepten te begrijpen. De ervaringen gelinkt aan een thermostaat zouden zich beperken tot ‘flitsen’ ervaring van een heel rudimentaire aard.¹⁹⁵ Chalmers gebruikt hier bewust de term *gelinkt*, want als we zouden spreken van een thermostaat die ervaringen *heeft*, zou dit lijken te impliceren dat de thermostaat een subject is. Over deze kwestie zegt Chalmers het volgende:

The subject is the whole system, or better, is associated with the system in the way that a subject is associated with a brain. The right way to speak about this is tricky. We would not say that my brain has experiences, strictly speaking, but that *I* have experiences.¹⁹⁶

Dit is niet erg duidelijk. Mijn eigen —speculatieve— interpretatie hiervan is dat we hierover moeten denken in het kader van de ‘flitsen’ die Chalmers vernoemt. De qualia waarover we hier spreken zijn niet aanwezig in de wereld buiten ons lichaam: de kleur rood, of beter de ervaring van de kleur rood te zien zit niet in de rode voorwerpen die we zien. De ervaring van de kleur rood zien is volgens het dubbel-aspect principe een fenomenale realisatie van een informatieruimte die ook fysisch gerealiseerd wordt in ons brein. Gedachten, herinneringen en ons ‘begrip’ van concepten, net als ons gebruik van taal door onze ‘innerlijke stem’ zijn volgens DAT evengoed qualia. Nu beschouwen wij onszelf als een subject, een *zelf*, aangezien we weten wie we zijn via herinneringen aan zaken zoals onze vroegere ervaringen, onze naam, onze karaktertrekken en ons moreel kompas. Indien DAT correct is zijn al die zaken ook qualia of fenomenale realisaties van informatieruimtes, die in ons brein ook fysisch gerealiseerd worden. Deze ‘flitsen’ die een rol spelen bij het ervaren van onszelf als een subject, een *ik*, worden via ons korte- en langetermijngeheugen en misschien ook door zaken zoals taal aan elkaar verbonden. Als we bewust of zelfbewust zijn spelen op de achtergrond van onze ervaring misschien steeds dergelijke herinneringen in zekere mate een rol, en vormen die een inherent onderdeel van onze ervaring. Met andere woorden *is* het subject in die zin de ervaring. Als we deze gedachtegang volgen zou er bij iemand die bijvoorbeeld dement wordt waarschijnlijk steeds meer van

¹⁹⁴ Dit lijkt althans het meest plausibel, het valt immers niet uit te sluiten dat eventuele ervaringen gelinkt aan thermostaten heel rijk zijn, alhoewel het niet evident is om zich een manier voor te stellen die daartoe aanleiding zou geven.

¹⁹⁵ Chalmers merkt op dat de term ‘panpsychisme’ daarom wat misleidend is omdat deze de indruk kan opwekken dat volgens deze visie ieder fysisch ding over cognitieve capaciteiten beschikt. Zie Chalmers, *The Conscious Mind*, 298-99.

¹⁹⁶ Chalmers, *The Conscious Mind*, 296-97.

die achtergrond weg vallen, samen met qualia die overeen komen met ons begrip van concepten en taal, tot er enkel een opeenvolging van 'flitsen' optreedt met weinig verband of betekenis.

In verband met deze kijk op het subject van ervaring kunnen we ons ook afvragen of er vanuit het brein niet meerdere ervaringen tegelijkertijd uitgaan en dus meerdere subjecten. Chalmers merkt op dat slechts een klein deel van het verwerken van informatie door ons brein geassocieerd is met bewustzijn. Hij zegt hier het volgende over:

If the unrestricted double-aspect principle is correct, then presumably the answer is that all that "unconscious" information *is* realized in experience— it is just not realized in *my* experience.¹⁹⁷

Als het dubbel-aspect principe onbeperkt geldig is, zou dat betekenen dat ook deelsystemen van ons brein, zoals afzonderlijke neuronen, geassocieerd zouden zijn met (rudimentaire) ervaringen. Die ervaringen zouden echter niet overeen komen met 'subjecten', aangezien die deelsystemen niet beschikken over een geheugen zoals ons brein. De deelsystemen waarvan hier sprake is zouden dus enkel geassocieerd zijn met 'flitsen' van ervaring, die geen verband met elkaar hebben. Mijn eigen 'zelf' zou ook geen toegang hebben tot die ervaringen, en ze ook niet opslaan in mijn geheugen. Deze visie leidt tot een beeld van de wereld die gevuld is met ervaringen:

Still, this great proliferation of experiences, especially the proliferation within a single brain, might be cause for discomfort. This is exacerbated by noting that when given an information space, it is usually easy to find many slightly different information spaces simply by individuating a relevant causal pathway differently, or by carving up the relevant effects (the "differences" that the information makes) in a slightly different way.¹⁹⁸

Volgens dezelfde gedachtegang moeten we ons ook afvragen of de onderdelen waaruit een fysisch object, zoals een thermostaat, is opgebouwd, ook gelinkt kan zijn met een bepaalde ervaring.

Dit druist natuurlijk sterk in tegen onze intuïties, maar Chalmers lijkt een voorkeur te hebben voor deze visie die panpsychisme wordt genoemd. Als argument hiervoor geeft hij het feit dat de meeste fundamentele eigenschappen zoals massa of elektrische lading algemeen voorkomen. Als het inderdaad waar zou zijn dat ook ervaring fundamenteel is, zou het niet implausibel zijn dat ervaring alomtegenwoordig zou zijn.¹⁹⁹ Dit hoeft natuurlijk niet zo te zijn, aangezien de fenomenale eigenschappen volgens Chalmers niet te herleiden zijn tot fysische eigenschappen, en de bewuste

¹⁹⁷ Chalmers, *The Conscious Mind*, 299.

¹⁹⁸ Chalmers, *The Conscious Mind*, 300.

¹⁹⁹ Chalmers, "Facing up to the problem of consciousness," 217.

ervaringen waarvan we weet hebben enkel gelinkt zijn aan zeer complexe fysische structuren zoals ons brein. Een fundamentele theorie van bewustzijn zou volgens Chalmers ook een beperkende voorwaarde kunnen bevatten, die specificeert dat alleen informatieruimten met bepaalde eigenschappen niet alleen fysisch maar ook fenomenaal gerealiseerd worden.²⁰⁰ Verder is onze huidige kennis van de fysica beperkt, en valt het niet uit te sluiten dat we ons simpelweg niet bewust zijn van het bestaan van zeer zeldzaam aanwezige fundamentele fysische eigenschappen. Tenslotte zouden we kunnen zeggen dat het contra-intuïtief zijn van panpsychisme geen onoverkomelijk probleem is voor deze theorie, aangezien ook heel wat zaken uit de huidige fysica zeer contra-intuïtief zijn.

Panpsychisme heeft wel één belangrijk voordeel. Indien fenomenale aspecten van de werkelijkheid algemeen voorkomen, hoeven we geen arbitraire grens te stellen waar bewustzijn verdwijnt wanneer we steeds eenvoudiger organismen op de evolutionaire ladder beschouwen. Een zekere mate van ervaring zou volgens het dubbel-aspect principe zelfs nog voorkomen bij maximaal eenvoudige informatie verwerkende systemen zoals een thermostaat.²⁰¹ Als het dubbel-aspect principe in onbeperkte vorm geldig zou zijn, zou dit zelfs tot gevolg hebben dat er ook aan bijvoorbeeld een elektron een zekere ervaring zou kunnen gekoppeld worden:

If there is experience associated with thermostats, there is probably experience everywhere: wherever there is a causal interaction, there is information, and wherever there is information, there is experience.²⁰²

Het verschil is wel, dat een elektron, in tegenstelling tot een thermostaat, geen informatie-verwerkend systeem is. Niettemin kan een elektron in verschillende toestanden verkeren, en zijn er op microschaal ook causale interacties gaande.

Een verder vraag die hierbij oprijst, is of causale interacties actief moeten doorgaan om gelinkt te worden aan ervaring, of dat ook een onveranderlijke informatietoestand (zoals bijvoorbeeld een thermostaat in een onveranderlijke toestand, maar evengoed een brein in 'bevrozen' toestand aanleiding geeft tot een ervaring. Chalmers is het niet zeker, maar heeft de intuïtie dat er voor ervaring een actieve causale interactie nodig is, al merkt hij op dat het op een fundamenteel niveau moeilijk kan zijn om actieve en passieve causatie van elkaar te onderscheiden. Vermoedelijk doelt hij met dit laatste op de continue activiteit die er door gaat op een submicroscopisch niveau.²⁰³

²⁰⁰ Chalmers, *The Conscious Mind*, 293.

²⁰¹ Chalmers, *The Conscious Mind*, 294.

²⁰² Chalmers, *The Conscious Mind*, 297.

²⁰³ Chalmers, *The Conscious Mind*, 298.

Chalmers bespreekt nog een andere manier waarop men het dubbel-aspect principe zou kunnen beperken. Zo zouden eenvoudige systemen misschien niet beschikken over fenomenale eigenschappen, maar eerder over de eerder besproken ‘protofenomenale’ eigenschappen. Die laatste zouden enkel aanleiding geven tot fenomenale eigenschappen wanneer ze gecombineerd worden in meer complexe systemen zoals ons brein. De visie waarin dit zou uitmonden is het zogenaamde panprotopsychisme.²⁰⁴ Zowel het panpsychisme als het panprotopsychisme kampen echter met een probleem, dat ik in het deel hieronder behandel.

3.13 Het combinatieprobleem

De vraag die bij dit alles rijst is nu hoe we van eenvoudige systemen uit de microwereld, met rudimentaire ervaringen, de overgang maken naar de complexe ervaringen van complexe macrofysische structuren zoals ons brein. Hoe worden fenomenale of protofenomenale eigenschappen samengesteld tot macroscopische ervaringen van levende wezens? Chalmers is van mening dat dit het grootste probleem is voor de hierboven besproken Russelliaanse versie van DAT.²⁰⁵ Zijn ervaringen op macrofysisch niveau simpelweg een soort ‘som’ van de rudimentaire ervaringen op microniveau? Of gebeurt dit op een andere manier?

Het vraag hoe rudimentaire ervaringen op microniveau aanleiding kunnen geven tot complexe ervaringen op macroscopisch niveau, werd al eerder gesteld door andere auteurs en staat bekend als het ‘combinatieprobleem’. Een voorloper ervan werd reeds geformuleerd door William James, die argumenteerde dat ervaringen niet kunnen samen komen tot verdere ervaringen.²⁰⁶ Zowel panpsychisme als het combinatieprobleem en mogelijke oplossingen krijgen recent veel aandacht in de literatuur van de *philosophy of mind*. Chalmers bespreekt het combinatieprobleem het meest in detail in zijn “The Combination Problem for Panpsychism” uit 2016. Daarin legt hij uit dat het combinatieprobleem eigenlijk op te delen is in minstens drie deelproblemen. Zo is er het *subject combinatieprobleem*, dat bestaat uit de vraag hoe microsubjecten combineren tot macrosubjecten. Daarnaast is er het *kwaliteiten combinatieprobleem*, hoe microkwaliteiten combineren tot macrokwaliteiten. Tenslotte is er het *structuur combinatie probleem*, hoe de structuur van ervaringen op microniveau (en microfysische structuur) combineert tot de structuur van ervaringen op macroniveau.²⁰⁷ In het panpsychisme zijn er verschillende stromingen, die elk een andere visie hebben op het combinatieprobleem. Volgens het constitutief panpsychisme zijn ervaringen op macroniveau

²⁰⁴ Chalmers, *The Conscious Mind*, 298.

²⁰⁵ Chalmers, *The Conscious Mind*, 307.

²⁰⁶ William James, *The Principles of Psychology* (Henry Holt, 1895), 160.

²⁰⁷ David Chalmers, “The Combination Problem for Panpsychism,” 3. The many combination problems.

volledig of gedeeltelijk gegrond in ervaringen op microniveau. Niet-constitutief panpsychisme is de these dat ervaringen op macroniveau niet gegrond zijn in ervaringen op microniveau. Een belangrijke vorm van niet-constitutief panpsychisme is het emergent panpsychisme, waarbij ervaringen op macroniveau verrijzen uit ervaringen op microniveau via bijvoorbeeld fundamentele wetten die de beide niveaus met elkaar verbinden.²⁰⁸

Over het combinatieprobleem is en wordt nog steeds heel veel gepubliceerd, ik zal mij hier beperken tot het structuur combinatie probleem met betrekking tot het panpsychisme, en het niet hebben over panprotopsychisme. Het structuur combinatieprobleem is het meest relevant met het oog op informatie en informatieruimtes. Chalmers maakt een onderscheid tussen ‘brede’ macrofysische structuur en ‘nauwe’ macrofysische structuur. Deze laatste is de structuur in termen van onder meer ruimte, massa en lading, die men klassiek beschouwt als de fysische structuur. De brede macrofysische structuur kan iedere structuur zijn die wordt samengesteld uit de microfysische structuur, zoals de chemische, biologische of causale structuur op verschillende niveaus. Volgens de Russelliaanse visie zou de structuur van ervaringen op macroniveau eerder moeten afhangen van de nauwe macrofysische structuur, die immers samengesteld wordt uit de eerder besproken ‘quiddities’. Chalmers denkt echter dat de manier waarop ervaringen worden samengesteld meer zou moeten te maken hebben met de “samenstelling van informatie”.²⁰⁹ Chalmers geeft in “The Combination Problem for Panpsychism” geen verdere details hierover, maar ik denk dat Chalmers hier bedoelt dat men zou moeten kijken naar hoe informatieruimtes op microniveau kunnen samengesteld worden tot informatieruimtes, die de structuur geven van ervaringen op macroniveau. We kunnen deze informatieruimtes echter, door ons gebrek van kennis over de fysische wetten en omwille van de veronderstelde enorme complexiteit, niet modelleren of op een andere manier voorstellen. En dit is natuurlijk opnieuw een zwak punt van DAT dat niet onmiddellijk zal kunnen opgelost worden. We kunnen ons deze informatieruimtes enkel voorstellen in onze verbeelding, maar dit is natuurlijk onvoldoende om het idee meer concreet uit te werken. We zullen in het volgende deel zien dat dit zelfde probleem ook opduikt bij een recente visie die bewustzijn ook linkt aan informatie, de *Integrated Information Theory*.

²⁰⁸ David Chalmers, “The Combination Problem for Panpsychism,” 2. Terminology.

²⁰⁹ David Chalmers, “The Combination Problem for Panpsychism,” 7.8 The structure combination problem.

4 *Integrated Information Theory*

De *Integrated Information Theory* (IIT) van Giulio Tononi is een recente benadering van het moeilijke probleem van bewustzijn vanuit de neurowetenschappen waarin informatie, net zoals bij DAT, een centrale rol speelt. IIT is een kwantitatieve en kwalitatieve benadering van bewustzijn die ook een verband legt tussen bewustzijn en het fysisch substraat waaruit bewustzijn voortkomt. IIT is een complexe theorie, die wiskundig is uitgewerkt. Ik zal mij hier echter beperken tot het weergeven van de belangrijkste principes om vervolgens te concentreren op de metafysische aannames, de overeenkomsten en de verschillen met DAT en het concept van informatie bij IIT. Sinds de eerste formulering van IIT in 2004 door Tononi is deze verder ontwikkeld in verschillende versies door Tononi en andere auteurs.²¹⁰ Ik zal het hier voornamelijk hebben over de twee meest recente versies van de theorie (“IIT 3.0” en “IIT 4.0”), en het concept van informatie dat in deze versies van IIT wordt gehanteerd.²¹¹

4.1 De methodologie van IIT

De theorie werkt niet volgens de gebruikelijke methode in de neurowetenschappen, waarbij vertrokken wordt van de werking van neuronen in het brein, en gezocht wordt naar de neurale correlaten van bewustzijn (*neural correlates of consciousness* of NCC). Die neurale correlaten van bewustzijn zijn dan specifieke patronen van al dan niet synchroon vurende neuronen of bepaalde delen van het brein. In plaats daarvan worden in IIT eerst vijf “zelf-evidente” fundamentele eigenschappen van ervaring onderscheiden, die ‘axioma’s’ worden genoemd:

- *Bestaan*: bewustzijn bestaat.
- *Samenstelling*: bewustzijn is gestructureerd.
- *Informatie*: bewustzijn is informatief (iedere ervaring is in haar specifieke manier verschillend van andere mogelijke ervaringen).
- *Integratie*: bewustzijn is geïntegreerd en niet te ontleden in verschillende ervaringen.
- *Uitsluiting*: bewustzijn is exclusief: een ervaring is een geheel dat andere ervaringen uitsluit.

Uit deze axioma’s worden vervolgens vijf overeenkomstige ‘postulaten’ (in feite zijn dit hypotheses) afgeleid, die telkens een eigenschap geven die geïntanceerd moet worden door het fysisch substraat

²¹⁰ Giulio Tononi, “An information integration theory of consciousness,” *BMC Neurosci.* 5, 42 (2004).

²¹¹ Voor IIT 3.0 zie Olimpia Lombardi & Cristian López, “What does ‘information’ mean in integrated information theory?,” *Entropy* (Basel) Nov 22, 20(12):894 (2018) en Masafumi Oizumi, Larissa Albantakis & Giulio Tononi, “From the phenomenology to the mechanisms of consciousness: Integrated Information Theory 3.0,” *PLoS Comput. Biol.* 10 (2014). Voor IIT 4.0 zie Larissa Albantakis et al., “Integrated information theory (IIT) 4.0: Formulating the properties of phenomenal existence in physical terms,” (2022).

van bewustzijn (dat substraat wordt een *systeem* genoemd, bestaande uit subsystemen of elementen). Om een inzicht te krijgen in de theorie zijn vooral de volgende twee postulaten het meest relevant:

- *Bestaan*: de fysische wetenschappen bestuderen enkel zaken die causaal interageren met elkaar. Om te bestaan moet het fysisch substraat van bewustzijn *oorzaak-effect kracht* hebben. Dat wil zeggen dat de subsystemen van het substraat moeten kunnen reageren op oorzaken en zelf ook effecten moeten kunnen genereren.
- *Uitsluiting*: het substraat van bewustzijn bezit de maximum waarde voor een te berekenen variabele, namelijk de hoeveelheid *geïntegreerde informatie* Φ (zie verder). Subsystemen die deel uitmaken van het substraat of nog grotere systemen waarvan het substraat deel uitmaakt, die een lagere waarde voor Φ bezitten worden door dit postulaat uitgesloten en niet geassocieerd met bewustzijn.^{212 213}

De twee bovenstaande postulaten worden samen met de overige drie postulaten omgezet naar een wiskundige beschrijving die de eigenschappen van het substraat van bewustzijn weergeven.

4.2 Geïntegreerde informatie: een kwantitatieve benadering van bewustzijn

Volgens IIT kan men het 'niveau' of de 'hoeveelheid' bewustzijn van een fysisch systeem meten door het getal Φ (*phi*) te berekenen, de hoeveelheid geïntegreerde informatie. In sommige publicaties over IIT wordt gesteld dat de geïntegreerde informatie *gelijk* is aan bewustzijn, maar dit is waarschijnlijk eerder een onnauwkeurigheid in de beschrijving. Ervaringen, de inhoud van bewustzijn, worden in IIT gekarakteriseerd door de T-vormen in qualia ruimte Q (zie verder). In IIT wordt informatie steeds gelinkt aan causaliteit. Meer specifiek wordt informatie gezien als de maatstaf voor de mate waarin een systeem causale beperkingen oplegt aan haar eigen onmiddellijke verleden en toekomstige toestanden, of hoeveel we kunnen te weten komen over de vorige en volgende toestand van een systeem, door te kijken naar de huidige toestand, zonder verstoring door externe invloeden. Een bepaalde toestand waarin een systeem verkeert sluit immers heel veel andere mogelijke toestanden uit. Hierbij wordt ook gekeken naar het aantal mogelijke toestanden waarin een systeem kan verkeren: hoe groter dit aantal, hoe meer informatie de huidige toestand van het systeem bevat. Merk op dat het bovenstaande een sterke overeenkomst vertoont met de manier waarop Chalmers de functionele/causale organisatie van een bewust systeem definieert in *The Conscious Mind*:

²¹² Albantakis, "Integrated information theory (IIT) 4.0: Formulating the properties of phenomenal existence in physical terms," 3-5.

²¹³ Oizumi, Albantakis & Tononi, "From the phenomenology to the mechanisms of consciousness: Integrated Information Theory 3.0," 2-3.

A functional organization is determined by specifying (1) a number of abstract components, (2) for each component, a number of different possible states, and (3) a system of dependency relations, specifying how the state of each component depends on previous states of all components and on inputs to the system, and how outputs from the system depend on previous component states.²¹⁴

Een voorbeeld van een systeem is ons brein, waarbij de causale interacties van de onderdelen, de neuronen, kan vergeleken worden met die van logische poorten zoals AND en OR, die ook een centrale rol innemen in de werking van computers, behalve dat de werking van neuronen veel complexer is dan de eenvoudige logische poorten in een computer.

Het geïntegreerd zijn van die informatie is de mate waarin deze informatie afhangt van de onderlinge causale verbindingen tussen de onderdelen van het systeem. Hoe meer een systeem, terwijl het informatie verwerkt, de informatie verspreidt over het ganse systeem, hoe groter de hoeveelheid geïntegreerde informatie en de waarde van Φ .^{215 216} Samengevat kunnen we zeggen dat de waarde van Φ weergeeft hoe complex de interne causale structuur van een systeem is.

Zo zijn er bijvoorbeeld bij het menselijk brein bijzonder veel verbindingen tussen de neuronen en zijn er zeer veel mogelijke toestanden, terwijl er ook voortdurend signalen verstuurd worden naar andere delen van het brein. Hierdoor is de hoeveelheid geïntegreerde informatie bij het menselijk brein doorgaans groot. Daarentegen heeft een meer gefragmenteerd systeem, waarbij verschillende delen geïsoleerd van elkaar bezig zijn met het verwerken van informatie, een lagere Φ . Zo kan een digitale camera grote hoeveelheden informatie verwerken, maar de delen van de camera werken slechts weinig samen, waardoor de waarde van Φ bij een dergelijk systeem laag blijft. Niettemin stelt IIT dat ook geïsoleerde, eenvoudige, niet-biologische systemen, waarbij de waarde van Φ verschillend is van 0, in zekere mate bewust zijn. Het is echter niet duidelijk of uit IIT het bewustzijn volgt van bijvoorbeeld protonen en neutronen, die immers geïntegreerde systemen zijn van quarks. Hedda Mørch stelt dat deze mogelijkheid door IIT sterk gesuggereerd wordt, maar er niet strikt uit volgt. IIT in haar huidige vorm definieert Φ immers enkel voor systemen die men kan opvatten als samengesteld uit discrete elementen, terwijl momenteel in de fysica algemeen wordt aangenomen dat deeltjes excitaties zijn van continue velden. Mørch houdt het wel voor mogelijk dat IIT kan worden uitgebreid naar een versie

²¹⁴ Chalmers, *The Conscious Mind*, 247.

²¹⁵ Mørch, "Is the Integrated Information Theory of Consciousness Compatible with Russellian Panpsychism?," 1066 en 1066, voetnoot 2.

²¹⁶ De manier waarop Φ berekend wordt is vrij technisch, en is niet eenvoudig op te maken uit de diverse publicaties over IIT. Een eenvoudig maar zeer verhelderd filmpje over deze methode kan men vinden op https://www.youtube.com/watch?v=Sp9B7raW_O4&ab_channel=AstonishingHypothesis.

waaruit een volledig panpsychisme zou volgen, inclusief fundamentele deeltjes met een zekere mate van bewustzijn.²¹⁷

Aanhangers van IIT beweren dat de theorie een belangrijke voorspellende en verklarende kracht heeft. Zo kan IIT onder meer verklaren waarom de kleine hersenen (het cerebellum) weinig belangrijk zijn voor bewustzijn, terwijl de grote hersenen onontbeerlijk zijn voor het voorkomen van het bewustzijn. Het cerebellum beschikt, in tegenstelling tot de grote hersenen, over minder verbindingen die feedback of terugkoppeling van informatie mogelijk maken, maar eerder over *feed-forward* verbindingen, zoals we die ook veel terug vinden in de huidige computers. De aanwezigheid van feedback verbindingen zorgt voor een grote causale verbondenheid van de elementen van een systeem en dus voor een hoge Φ .²¹⁸ Men kan dit echter omkeren en de vraag stellen of een abstracte theorie zoals IIT wel iets fundamenteel nieuws zegt over bewustzijn in plaats van gewoon uiting te geven aan het — eerder triviale — vermoeden dat de hogere cognitieve functies van ons brein en (zelf)bewustzijn te maken hebben met de activiteit van die delen van het brein die een complexe structuur hebben en feedback van informatie mogelijk maken. Als dat het geval zou zijn, zou IIT op zich bijvoorbeeld niets zeggen over het al dan niet bewust zijn van artificiële systemen. In de literatuur is deze kritiek inderdaad terug te vinden. Zo betwijfelt Michael Cerullo de voorspellende kracht van IIT door als tegenvoorbeeld een alternatieve theorie te formuleren die evenzeer als IIT voorspellingen kan doen, zoals de verschillende rollen van het cerebellum en de grote hersenen bij het bewustzijn. Deze alternatieve theorie is gebaseerd op de “zelf-evidente eigenschap” dat bewustzijn gerelateerd is aan informatie die zich verplaatst in feedback lussen in een systeem, iets dat men weet uit de circulaire verplaatsingen van signalen tussen de thalamus en de cortex bij visuele waarneming.²¹⁹ Ari Peuhu geeft een gelijkaardige kritiek op IIT. Volgens hem geeft IIT met betrekking tot het voorspellen van de mate van bewustzijn slechts gehoor aan wat al duidelijk werd uit de neurowetenschappen, namelijk dat de integratie van informatie feedback lussen vergt tussen neuronen. Volgens Peuhu heeft IIT voor de rest veel te weinig aandacht voor neuro-architecturale aspecten van het brein en de verspreiding van neurotransmitters doorheen de cortex. Daarom is IIT volgens hem een “veel te algemeen en geïdealiseerd kader om bewustzijn te kunnen verklaren.”²²⁰

²¹⁷ Mørch, “Is the Integrated Information Theory of Consciousness Compatible with Russellian Panpsychism?,” 1067, voetnoot 4.

²¹⁸ Mørch, “Is the Integrated Information Theory of Consciousness Compatible with Russellian Panpsychism?,” 1071.

²¹⁹ Michael Cerullo, “The Problem with Phi: A Critique of Integrated Information Theory,” *PLoS Comput Biol* 11(9) (2015): 5.

²²⁰ Ari Peuhu, “Critique of the Integrated Information Theory of Consciousness Or, the Relevance of Ontological Information,” *Journal of Consciousness Studies* 28, No. 5–6 (2021), 68.

Daarbij komt dat Φ alleen voor heel eenvoudige systemen exact kan berekend worden. Op die manier wordt de door IIT voorgestelde analyse reeds vanaf systemen met meer dan ongeveer twaalf elementen praktisch ondoenbaar.²²¹ Voor meer complexe systemen kan de waarde van Φ enkel geschat worden aan de hand van heuristieken. Pautz merkt op dat hierdoor de vraag rijst of IIT wel empirisch testbaar is.²²² Een andere kritiek op IIT is dat het niet duidelijk is wat in de theorie precies bedoeld wordt met ‘niveau van bewustzijn’, aangezien dat bewustzijn kan variëren volgens verschillende dimensies (zoals onder meer intensiteit en complexiteit).²²³

4.3 Qualia in IIT

IIT heeft ook een kwalitatief luik, en zodoende kan de theorie ook worden beschouwd als een poging om een antwoord te geven op het moeilijke probleem van bewustzijn. Net als bij DAT probeert IIT qualia niet te linken aan mogelijke neuronale correlaten van bewustzijn (NCC's) zoals het vierdimensionale fysische patroon van spatiotemporeel vuren van neuronen in het brein. IIT linkt qualia eerder aan een abstracte ‘oorzaak-effect’ ruimte die gelinkt wordt aan de causale interacties die plaatsvinden in het brein. In tegenstelling tot de eerder vage beschrijving die Chalmers geeft van informatieruimtes bij DAT wordt deze abstracte ruimte, die qualia-ruimte of kortweg Q-ruimte wordt genoemd, bij IIT wel meer in detail omschreven. De Q-ruimte is een hoogdimensionele ruimte waarbij elke as overeenkomt met iedere mogelijke voorbije of toekomstige toestand van een systeem.

Specifieke qualia corresponderen nu met een bepaalde subruimte van deze Q-ruimte, die ook ‘vormen’ worden genoemd.^{224 225} De patronen van vurende neuronen in het brein bepalen voor een deel hoe een dergelijke vorm er zou uit zien, maar worden niet alleen door deze patronen bepaald. Zo zijn ook neuronen die niet vuren, maar deel uitmaken van het systeem, belangrijk voor het bepalen van deze vorm.²²⁶ De vormen in Q-ruimte worden, enigszins verwarrend, ook ‘concepten’ genoemd. Het woord concept wordt hier echter niet in de traditionele zin gebruikt, waarbij een concept niets te maken heeft met causaliteit en eerder verwijst naar een buitentalige entiteit. De betekenis van de term concept bij

²²¹ Oizumi, Albantakis & Tononi, “From the phenomenology to the mechanisms of consciousness: Integrated Information Theory 3.0,” 24.

²²² Pautz, “What is the Integrated Information Theory of Consciousness? A Catalogue of Questions,” 191, noot 3.

²²³ Pautz, “What is the Integrated Information Theory of Consciousness? A Catalogue of Questions,” 193-94.

²²⁴ Matteo Grasso, “Towards a Double-Aspect Integrated Information Theory of Consciousness?,” conference paper, “The Philosophy of David Chalmers” Tübingen Masterclass in Theoretical Philosophy, Tübingen (2015): 5.

²²⁵ Pautz, “What is the Integrated Information Theory of Consciousness?,” 204.

²²⁶ David Balduzzi & Giulio Tononi, “Qualia: The geometry of integrated information,” *PLoS Computational Biology* 5, (8) (2009): 12.

IIT is de causale rol van een mechanisme binnenin het systeem.²²⁷ Pautz noemt deze concepten of vormen ook T-vormen (T van Tononi), een benaming die ik hier zal volgen. Afhankelijk van de bron worden deze abstracte T-vormen als *identiek* aan qualia verondersteld of eerder als een *karacterisering* van die qualia opgevat, waarbij er enkel een overeenkomst in *structuur* is tussen de Q-ruimte en de structuur van ervaringen. Elk van deze beide interpretaties betekent een belangrijk verschil voor de onderliggende metafysische aannames van IIT (zie verder). In verschillende bronnen worden de beide interpretaties door elkaar gebruikt, waardoor het niet echt duidelijk is welke precieze interpretatie wordt gevolgd door de auteur.²²⁸

De T-vormen zijn dus fysische, maar abstracte structuren. T-vormen zijn echter zo complex dat ze voor echte systemen, zoals het brein, maar zelfs ook voor de meest eenvoudige zenuwstelsels zoals van de rondworm *C. elegans*, in de praktijk niet te bepalen zijn. Een systeem met n binaire elementen heeft 2^n mogelijke toestanden, wat leidt tot T-vormen met evenveel dimensies. Een menselijk brein kan zich in een gigantisch aantal mogelijke toestanden bevinden, wat leidt tot een ruimte van T-vormen die bijna oneindig dimensioneel is.²²⁹ We zien dus dat de abstracte structuren, die gelinkt worden of gelijk zijn aan qualia, in IIT al beter gedefinieerd zijn. Tegelijk zien we echter dat deze, net zoals bij DAT, omwille van hun complexiteit niet praktisch kunnen voorgesteld worden voor werkelijk voordoende ervaringen zoals we die kennen vanuit onze fenomenologie.

4.4 Het concept van informatie bij IIT

De geïntegreerde informatie waarvan sprake is bij IIT zouden we kunnen omschrijven als causale informatie, aangezien ze de maatstaf is voor de mate waarin een systeem causale beperkingen oplegt aan de toestanden van het systeem in het onmiddellijke verleden en toekomst van het systeem. Geïntegreerde informatie is duidelijk verschillend van informatie volgens Shannon, die geen semantische invulling krijgt en daardoor niets zegt over de interne causale structuur van een systeem. Het concept van informatie bij IIT vertoont wel een duidelijke overeenkomst met het concept van informatie bij Chalmers, aangezien deze bij IIT ook een causale en een (via de correspondentie of identiteit tussen T-vormen en qualia) fenomenale invulling krijgt. Als de T-vormen hierbij worden gezien als louter karakteriserend voor specifieke qualia, lijkt deze interpretatie van IIT meer op de zwakkere versie van DAT die informatie ook eerder zag als de link tussen het fysische en het mentale. Als de T-

²²⁷ Olimpia Lombardi & Cristian López, "What does 'information' mean in integrated information theory?," 3. Differences between IIT Information and Shannon Information.

²²⁸ Giulio Tononi & Christof Koch, "Consciousness: here, there and everywhere?," *Phil. Trans. R. Soc. B* 370: 20140167 (2015): 9.

²²⁹ Pautz, "What is the Integrated Information Theory of Consciousness? A Catalogue of Questions," 205.

vormen echter worden gezien als identiek met specifieke qualia, lijkt IIT in dit aspect meer op de radicalere vorm van DAT die Chalmers lijkt te verdedigen, waar informatie fundamenteel is en de basis vormt voor het mentale.

Daarnaast is geïntegreerde informatie ook een intrinsieke eigenschap van een systeem, aangezien ze informatie geeft over de interne causale structuur van een systeem. Matteo Grasso stelt wel dat deze intrinsieke eigenschap van informatie vooral bij de latere versies van IIT op de voorgrond treedt.²³⁰ Olimpia Lombardi en Christian López merken in dat verband op dat informatie volgens Shannon daarentegen extrinsiek is.²³¹ Zoals we in onze eerdere bespreking van Shannon informatie reeds zagen, gaat deze eerder over inputs en outputs van een communicatiesysteem zonder dat het exacte mechanisme van dit communicatiesysteem een rol speelt. Het concept van informatie bij Chalmers zouden we ook eerder extrinsiek van aard kunnen noemen, gezien Chalmers het heeft over fysische objecten die causale interacties aangaan in een causale keten, zonder dat er sprake is van de intrinsieke causale structuur van een systeem.

Tenslotte is het belangrijk om op te merken dat bewustzijn in IIT enkel wordt toegekend aan systemen waarvan de *geïntegreerde* informatie verschillend is van 0, dus niet aan om het even welk systeem dat informatie bevat. Daarnaast zorgt het postulaat van uitsluiting bij IIT voor een belangrijke beperking van het aantal systemen dat verondersteld wordt bewust te zijn. In die zin geeft IIT een concrete invulling voor de mogelijkheid die Chalmers overweegt om het soort informatie dat aanleiding geeft tot bewustzijn te beperken.

4.5 De metafysica achter IIT

Zoals we boven uit de bespreking van het kwalitatieve luik van IIT reeds zagen is de metafysische visie achter IIT niet eenduidig. Zowel aanhangers als critici van de theorie interpreteren daarenboven de principes van IIT op een andere manier, en de uitspraken van enkele belangrijke vertegenwoordigers van IIT worden soms verschillend geïnterpreteerd. Volgens Adam Pautz, die verschillende kritieken formuleerde in verband met IIT, is deze theorie neutraal tussen type B materialisme, het eigenschapsdualisme van Chalmers en emergent fysicalisme ten opzichte bewustzijn. Volgens het emergent fysicalisme zijn mentale eigenschappen nieuwe fysische eigenschappen die tevoorschijn komen ('emergent' zijn) uit complexe fysische systemen. Type B materialisten zouden de principes van IIT kunnen beschouwen als onderschreven door de identiteit tussen bewustzijn en de fysische toestanden die aanleiding geven tot geïntegreerde informatie. In de literatuur zijn inderdaad

²³⁰ Grasso, "Towards a Double-Aspect Integrated Information Theory of Consciousness?," 9.

²³¹ Lombardi & López, "What does 'information' mean in integrated information theory?," 3. Differences between IIT Information and Shannon Information.

voorbeelden te vinden van die visie, zoals Lawrence Ward en Ramón Guevara, die op basis van de principes van IIT tot de conclusie komen dat de elektromagnetische velden die het brein genereert, *gelijk zijn* aan ervaringen.²³² Aanhangers van het eigenschapsdualisme en emergent fysicisme kunnen de principes van IIT eerder zien als fundamentele psychofysische wetten die de emergentie van bewustzijn linkt met de fysische toestanden waar geïntegreerde informatie uit voortkomt.²³³ Tononi is volgens Pautz eerder een aanhanger van het type B materialisme, vermoedelijk omdat hij de identiteiten die Tononi onderschrijft tussen fenomenale eigenschappen en informationele/causale eigenschappen van fysische systemen interpreteert als de identiteiten die type B materialisten veronderstellen.²³⁴ Christof Koch, een andere prominente aanhanger van IIT, omschrijft zichzelf zowel als een “romantisch reductionist” als een dualist die denkt dat bewustzijn fundamenteel verschillend is van de materiële oorzaak en niet kan gereduceerd worden tot fysische eigenschappen.²³⁵ Matteo Grasso, één van de IIT-onderzoekers, onderzocht de metafysische aannames van IIT en kwam tot de conclusie dat IIT compatibel is met zowel type B materialisme als met het eigenschapsdualisme zoals van Chalmers.²³⁶ Een belangrijke overeenkomst tussen DAT en IIT is volgens Grasso dat de principes van IIT in overeenstemming zijn met de principes van structurele coherentie en organisationele invariantie bij Chalmers. Net als in DAT is bewustzijn een aspect van informatie, net zoals de fysische werkelijkheid, die een ander aspect van informatie is:

DAT and IIT define consciousness as a certain aspect pertaining information, while physical reality constitutes another manifestation of the same fundamental principle (*“it from bit” doctrine*).²³⁷

Deze verwijzing van Grasso naar *“it from bit”*, afkomstig uit een niet gepubliceerd document dat gebruikt werd tijdens een conferentie, is echter de enige verwijzing naar deze theorie die ik kon vinden in de geraadpleegde literatuur over IIT. Grasso beschouwt Tononi, in tegenstelling tot Pautz, als een aanhanger van het soort eigenschapsdualisme dat Chalmers verdedigt, omwille van de claim door

²³² Lawrence Ward & Ramón Guevara, “Qualia and Phenomenal Consciousness Arise From the Information Structure of an Electromagnetic Field in the Brain,” *Front. Hum. Neurosci.* 16: 874241 (2022).

²³³ Pautz, “What is the Integrated Information Theory of Consciousness? A Catalogue of Questions,” 188-89.

²³⁴ Oizumi, Albantakis & Tononi, “From the phenomenology to the mechanisms of consciousness: Integrated Information Theory 3.0,” 3.

²³⁵ Christof Koch, *Consciousness: Confessions of a Romantic Reductionist* (Cambridge, MA: MIT Press, 2012), 119.

²³⁶ Matteo Grasso, “Integrated information theory and the metaphysics of Consciousness,” conference paper, Online Consciousness Conference, february 2013.

²³⁷ Matteo Grasso, “Integrated information theory and the metaphysics of Consciousness,” 8.

Tononi dat bewustzijn een fundamentele eigenschap is, “net als materie, lading of energie”.²³⁸ ²³⁹ Grasso benadrukt dat als IIT een exhaustieve theorie over bewustzijn wil zijn, de metafysische aannames en haar filosofische standpunten verder moeten uitgeklaard worden.²⁴⁰ Ondanks de verschillende metafysische interpretaties bij de verschillende onderzoekers die betrokken zijn bij de ontwikkeling van IIT kunnen we stellen dat er grote overeenkomsten zijn met DAT. Het is in die zin mogelijk om IIT te interpreteren als een verdienstelijke poging om bepaalde aspecten van DAT verder uit te werken, met name de manier waarop we informatieruimtes gelinkt aan causale interacties zouden kunnen zien. IIT probeert daarnaast ook de link te maken met de neurowetenschappen, al is het niet duidelijk of ze testbaar is en of haar voorspellende kracht wel zo groot is als haar voorstanders beweren. De complexiteit van de veronderstelde abstracte structuren die aanleiding geven tot qualia is echter opnieuw een obstakel om de theorie beter uit te werken.

²³⁸ Tononi, “An information integration theory of consciousness,” 19.

²³⁹ Giulio Tononi, “Consciousness as integrated information: a provisional manifesto,” *The Biological Bulletin*, 215 (3), (2008): 233.

²⁴⁰ Grasso, “Integrated information theory and the metaphysics of Consciousness,” 13.

5 Conclusie

De rol van het concept van informatie bij DAT bestaat in feite uit twee verschillende rollen, naargelang de versie van DAT die men bedoelt. De eerste rol is enkel de link verzorgen tussen fysische processen en ervaring, waarbij die laatste twee in DAT gezien worden als fundamentele, niet tot elkaar te reduceren domeinen. De tweede rol die het concept van informatie kan spelen moet men zien in het kader van een radicalere versie van DAT, die Chalmers lijkt te verdedigen, waarbij informatie als fundamentele bouwsteen van de wereld wordt gezien. Informatie, die uiteindelijk niet meer omhelst dan primitieve verschillen, wordt volgens het dubbel-aspect principe gerealiseerd in twee ‘aspecten’ van informatie, een fysisch aspect en een fenomenaal aspect. Op die manier kan DAT gezien worden als een uitbreiding van “*it from bit*” of de huidige, meer algemene versie in de vorm van ontisch structureel realisme. Het is een uitbreiding van die visies met bewustzijn, dat in DAT als fundamenteel wordt beschouwd. Chalmers is echter uiteindelijk toch gedwongen om via fenomenale eigenschappen, die immers meer lijken te zijn dan pure informatieruimtes, enige ‘substantie’ te geven aan deze visie op de werkelijkheid.

Chalmers bedoelde DAT als proto-theorie, als een ruwe schets van hoe de informatie de link tussen het fysische en het mentale zou kunnen verklaren. DAT lijkt coherent te kunnen zijn, maar heeft een aantal belangrijke zwakke punten. Hoewel DAT heel complex is komt de dubbel-aspect ‘theorie’ op die manier uiteindelijk eerder neer op een zeer speculatief, intuïtief idee of vermoeden van een metafysisch principe. Maar het idee is op zich interessant, en verdient een meer nauwkeuriger en uitgebreidere formulering, inclusief het expliciet maken van de vele vooronderstellingen.

IIT kan geïnterpreteerd worden als een verdere uitwerking van het idee dat informatie fundamenteel is voor ervaring, die tevens de link maakt tussen DAT en de neurowetenschappen. IIT wordt echter geplaagd door gelijkaardige problemen. Er is geen overeenstemming in wat de achterliggende metafysische kijk is van de theorie, en de oorzaak-effect structuren en T-vormen die de theorie veronderstelt, zijn net als de informatieruimtes waarvan DAT spreekt te complex om exact uit te werken.

Informatie speelt een steeds belangrijker rol in de fysische wetenschappen. Het is heel gemakkelijk om, als men reeds aangetrokken is door het idee van “*it from bit*” en men tevens overtuigd is van de bijzondere status van bewustzijn, informatie te zien als datgene wat ook onderliggend is voor bewustzijn. Het vermoeden dat het fysische en het fenomenale beide aspecten zijn van informatie is dan heel natuurlijk. Ik vermoed dat dit ook voor Chalmers de inspiratie was om zijn theorie te formuleren. Men wordt gemakkelijk intuïtief aangetrokken door de hypothese van een principe dat op

simpele en elegante wijze het fysische en het fenomenale domein van onze werkelijkheid linkt aan een abstracte, onderliggende wereld van informatie, een wereld die misschien niets anders is dan een verzameling van verschillen, van manieren waarop de wereld zou kunnen bestaan. Jammer genoeg lijkt het er echter op dat dit idee nog voor lange tijd of misschien zelfs wel voor altijd een vermoeden zal blijven.

Bibliografie

Adriaans, Pieter. "Information.", *The Stanford Encyclopedia of Philosophy* (Fall 2020 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/fall2020/entries/information/>>.

Albantakis, Larissa et al. "Integrated information theory (IIT) 4.0: Formulating the properties of phenomenal existence in physical terms." (2022) arXiv:2212.14787.

Balaguer, Mark. "Platonism in Metaphysics.", *The Stanford Encyclopedia of Philosophy* (Spring 2016 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/spr2016/entries/platonism/>>.

Balduzzi, David & Giulio Tononi. "Qualia: The geometry of integrated information." *PLoS Computational Biology* 5 (8) (2009): 1–24.

Bateson, Gregory. *Steps to an Ecology of Mind. Collected essays in anthropology, psychiatry, evolution, and epistemology.* Jason Aronson Inc., 1987 Reprint. Originally published: San Francisco: Chandler Pub. Co., 1972.

Bilodeau, Douglas J. "Physics, Machines, and the hard problem." *Journal of Consciousness Studies* 3, No. 5-6 (1996): 386-401.

Block, Ned. "Troubles with functionalism." In *Perception and Cognition: Issues in the Foundation of Psychology*, edited by C. W. Savage. Minneapolis: University of Minnesota Press, 1978. [Reprinted in N. Block, editor, *Readings in the Philosophy of Psychology*. Vol. 1. Cambridge, Mass.: Harvard University Press, 1980].

Brybaert Marc. *Psychologie*. Gent: Owl Press, Borgerhoff & Lamberights nv, 2018.

Chalmers, David. "Facing up to the problem of consciousness.", *Journal of Consciousness Studies* 2, No.3 (1995): 200-219.

———. *The Conscious Mind*. Oxford: Oxford University Press, 1996.

———. "Moving Forward on the Problem of Consciousness." *Journal of Consciousness Studies* 4 (1) (1997): 3–46. <https://consc.net/papers/moving.html>.

———. "Consciousness and Its Place in Nature." In *Philosophy of Mind. Classical and Contemporary Readings*, edited by David Chalmers. Oxford/New York: Oxford University Press, 2002.

———. *The Character of Consciousness*. Oxford University Press, 2010.

———. "The Combination Problem for Panpsychism." In *Panpsychism: Contemporary Perspectives*, edited by Godehard Bruntrup & Ludwig Jaskolla, 179–214. Oxford University Press, 2016.

———. *Reality+, virtual worlds and the problems of philosophy*. New York: W. W. Norton & Company, 2022.

Churchland, Paul M. "The rediscovery of light." *Journal of Philosophy* 93 (1996): 211-28.

De Mol, Liesbeth. "Turing Machines.", *The Stanford Encyclopedia of Philosophy* (Winter 2021 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/win2021/entries/turing-machine/>>.

Dennett, Daniel C. "The Unimagined Preposterousness of Zombies." *Journal of Consciousness Studies* 2 (1995): 322–26.

———. "A History of Qualia." *Topoi* 39 (2020): 5–12. <https://doi.org/10.1007/s11245-017-9508-2>.

Floridi, Luciano. *The Philosophy of Information*. New York: Oxford University Press Inc., 2011.

Fredkin, Edward. "Digital mechanics." *Physica* D45 (1990): 254-70.

Grasso, Matteo. "Integrated information theory and the metaphysics of Consciousness." Conference paper, Online Consciousness Conference, february 2013.

https://www.researchgate.net/publication/281208686_Integrated_information_theory_and_the_metaphysics_of_consciousness/references.

Grasso, Matteo. "Towards a Double-Aspect Integrated Information Theory of Consciousness?" Conference paper, "The Philosophy of David Chalmers." Tübingen Masterclass in Theoretical Philosophy, Tübingen (2015).

Hardcastle, Valerie Gray. "The why of consciousness: a non-issue for materialists." *Journal of Consciousness Studies* 3, No.1 (1996): 7-13.

Haynes, John-Dylan & Geraint Rees. "Decoding mental states from brain activity in humans." *Nature Reviews Neuroscience* 7 (2006). DOI: 10.1038/nrn1931.

Irvine, Andrew David. "Bertrand Russell.", *The Stanford Encyclopedia of Philosophy* (Spring 2022 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/spr2022/entries/russell/>>.

Jackson, Frank. "Epiphenomenal qualia." *Philosophical Quarterly* 32 (1982): 127-136.

James, William. *The Principles of Psychology*. Henry Holt, 1895.

Kefalas, Asterios G. "Cybernetics." In *Encyclopedia of Information Systems*, edited by Hossein Bidgoli, 365-378. Academic Press Inc, 2002.

- Kim, Jaegwon. *Physicalism, or something near enough*. Princeton Monographs in Philosophy, 2005, Princeton University Press.
- Koch, Christof. *Consciousness: Confessions of a Romantic Reductionist*. Cambridge, MA: MIT Press, 2012.
- Kragel, Philip A. et al. "Representation, pattern information, and brain signatures: From neurons to neuroimaging." *Neuron* 99 (2) (2018): 257-273.
- Lewis, David. *On the Plurality of Worlds*. Oxford: Blackwell, 1986.
- . "What experience teaches." In *Mind and Cognition*, edited by W. Lycan, 77-103. Oxford: Blackwell, 1990.
- Libet, Benjamin. "Solutions to the hard problem of consciousness." *Journal of Consciousness Studies* 3, No.1 (1996): 33-35.
- Lombardi, Olimpia & López, Cristian. "What does 'information' mean in integrated information theory?" *Entropy* Vol.20 (12), Article 894 (2018). [https://doi: 10.3390/e20120894](https://doi.org/10.3390/e20120894).
- McLaughlin, Brian & Karen Bennett. "Supervenience.", *The Stanford Encyclopedia of Philosophy* (Summer 2021 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/sum2021/entries/supervenience/>.
- Mørch, Hedda Hassel. "Is the Integrated Information Theory of Consciousness Compatible with Russellian Panpsychism?" *Erkenntnis* 84 (5) (2019): 1065-1085.
- Nemirow, Laurence. "Physicalism and the cognitive role of acquaintance." In *Mind and Cognition*, edited by W. Lycan. Oxford: Blackwell, 1990.
- Oizumi, Masafumi, Larissa Albantakis & Giulio Tononi. "From the phenomenology to the mechanisms of consciousness: Integrated Information Theory 3.0." *PLoS Comput. Biol.* 10 (2014).
- Pautz, Adam. "What is the Integrated Information Theory of Consciousness? A Catalogue of Questions." *Journal of Consciousness Studies* 26, No. 1–2 (2019): 188–215.
- Peuhu, Ari. "Critique of the Integrated Information Theory of Consciousness Or, the Relevance of Ontological Information." *Journal of Consciousness Studies* 28, No. 5–6 (2021): 58–78.
- Piccinini, Gualtiero & Neal G. Anderson. "Ontic Pancomputationalism." In *Physical Perspectives on Computation, Computational Perspectives on Physics*, edited by Michael E. Cuffaro & Samuel C. Fletcher, Cambridge University Press, 2018.

Robinson, Howard. "Substance.", *The Stanford Encyclopedia of Philosophy* (Fall 2021 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/fall2021/entries/substance/>>.

———. "Dualism.", *The Stanford Encyclopedia of Philosophy* (Spring 2023 Edition), Edward N. Zalta & Uri Nodelman (eds.), URL = <<https://plato.stanford.edu/archives/spr2023/entries/dualism/>>.

Rucker, Rudy. *Oneindigheid: Filosofie en wetenschap van het oneindige*. Amsterdam: Uitgeverij Contact, 1986.

Shannon, Claude. E. "A mathematical theory of communication." *Bell Systems Technical Journal* 27 (1948): 379-423. [Reprinted in C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication*. Urbana: University of Illinois Press, 1949].

Spinoza, Baruch. *The Ethics*. (1677) Reprinted in *The Ethics of Benedict Spinoza* New York: Van Nostrand, 1876.

Stoljar, Daniel. "Physicalism.", *The Stanford Encyclopedia of Philosophy* (Summer 2023 Edition), Edward N. Zalta & Uri Nodelman (eds.), URL = <<https://plato.stanford.edu/archives/sum2023/entries/physicalism/>>.

Stubenberg, Leopold & Donovan Wishon. "Neutral Monism.", *The Stanford Encyclopedia of Philosophy* (Spring 2023 Edition), Edward N. Zalta & Uri Nodelman (eds.), URL = <<https://plato.stanford.edu/archives/spr2023/entries/neutral-monism/>>.

Tegmark, Max. "The Mathematical Universe." *Foundations of Physics* 38 (2) (2008): 101–150.

Tononi, Giulio. "An information integration theory of consciousness." *BMC Neurosci.* 5 (2004).

———. "Consciousness as integrated information: a provisional manifesto." *The Biological Bulletin* 215 (3) (2008): 216-242.

Tononi, Giulio & Christof Koch. "Consciousness: here, there and everywhere?" *Phil. Trans. R. Soc. B* 370 (2015). <http://dx.doi.org/10.1098/rstb.2014.0167>.

van Inwagen, Peter. *Material Beings*. Ithaca and London: Cornell, 1990.

Velmans, Max. "The relation of consciousness to the material world." *Journal of Consciousness Studies* 2 (3) (1995): 255-65.

———. "How to define Consciousness and how not to define Consciousness." *Journal of Consciousness Studies* 16 (5) (2009): 139-156.

Ward, Lawrence M. & Ramón Guevara. "Qualia and Phenomenal Consciousness Arise From the Information Structure of an Electromagnetic Field in the Brain." *Front. Hum. Neurosci.* 16: 874241 (2022). doi: 10.3389/fnhum.2022.874241.

Wheeler, John. A. (1982). "The computer and the universe." *International Journal of Theoretical Physics* 21 (1982): 557–572.

———. "Information, physics, quantum: The search for links." In *Complexity, Entropy, and the Physics of Information*, edited by W. H. Zurek. Redwood City, Calif.: Addison-Wesley, 1990.

