



## Kwalia en het machinetoestandsfunctionalisme: een probleem in de filosofie van het bewustzijn.

WO BACHELOR FILOSOFIE SCRIPTIE

MAX GIEZEN, 550603

BEGELEID DOOR DR. F. VAN DE PUTTE

GEADVISEERD DOOR DR. J.P. KLOEG

Aantal woorden: 10.491

## Inhoud

Lijst van afkortingen. ....	2
Inleiding. ....	3
Putnams functionalisme in de filosofie van de geest.....	3
Zijn kwalitatieve waarnemingen functionalistisch verklaarbaar? .....	3
De probleemstelling, oplossingsmethode en relevantie van deze scriptie.....	4
Hoofdstuk 1: Putnams “machinetoestandsfunctionalisme”.....	6
Wat is een Turingmachine?.....	6
De analogie tussen geesten en machines.....	8
De pijnhypothese.....	11
Randvoorwaarden van een functionele-toestandstheorie. ....	14
De meervoudige realiseerbaarheid van mentale toestanden.....	16
Putnams “machinetoestandsfunctionalisme”.....	17
Hoofdstuk 2: Een behandeling van het afwezige kwalia-argument. ....	19
Het afwezige kwalia-argument.....	19
Een functionalistisch bezwaar tegen het afwezige kwalia-argument. ....	20
Kritiek en aanvulling.....	22
Hoofdstuk 3: Een behandeling van het omgekeerde kwalia-argument. ....	25
Het omgekeerde kwalia-argument.....	25
Een functionalistisch bezwaar tegen het omgekeerde kwalia-argument. ....	26
Kritiek en aanvulling.....	28
Conclusie.....	33
Bibliografie.....	34

## Lijst van afkortingen.

C-FO & H-FO: C-FO is de FO, geconcretiseerd door Chinezen, die de FO van de hersenen nabootst - dat is de H-FO. Zie “FO/FOs” in deze lijst en “Hoofdstuk 2: een behandeling van het afwezige kwalia-argument.”

EGS/EGSn: Een “empirisch gegeven systeem” of ten minste twee “empirisch gegeven systemen” zijn fysische objecten met een structuur en samenstelling die als PA kunnen worden gekarakteriseerd. Zie “Pan/PA” in deze lijst en sectie “De meervoudige realiseerbaarheid van mentale toestanden.”

FO/FOs: De “Functionele Organisatie” of de “Functionele Organisaties” van EGSn. Zie “EGS/EGSn” in deze lijst en “Hoofdstuk 1: Putnams “machinetoestandsfunctionalisme”.”

MRH: De “meervoudige realiseerbaarheidshypothese” is de hypothese van Hilary Putnam die voorstelt dat fysische objecten van verschillende structuur en samenstelling, met verschillende fysische eigenschappen, mentaal identieke toestanden kunnen realiseren. Zie sectie “De meervoudige realiseerbaarheid van mentale toestanden.”

MT-functionalisme/MT-functionalist: “Machinetoestandsfunctionalisme” is de eerste specifieke versie van het algemene functionalisme, uitgewerkt door Hilary Putnam – een “machinetoestandsfunctionalist”. Zie sectie “Putnams “machinetoestandsfunctionalisme”.”

PAn/PA: “Probabilistische Automaten” of “Probabilistische Automaat”; ook wel waarschijnlijkheidsautomaat genoemd. Zie sectie “De aard van mentale toestanden.”

TT/TTn: De “Totale Toestand” of de “Totale Toestanden” waar EGSn zich temporeel in bevinden. Zie “EGS/EGSn” in deze lijst en sectie “De aard van mentale toestanden.”

## Inleiding.

### Putnams functionalisme in de filosofie van de geest.

Ik wil duidelijk maken dat ik met “geest” het Engelse “mind” bedoel. *Geest* betekent in deze opvatting niet “de onstoffelijke, spirituele levensvorm”, maar “het vermogen tot bewustzijn, ervaring, voelen en denken”; het intellect.

De *filosofie van de geest* is een verzameling studies die de geest, de mentale functies, de mentale eigenschappen, de mentale processen en het bewustzijn onderzoeken, en hun relatie met het fysieke lichaam. Die relatie is voor elke studie betwistbaar, en daarom wordt die relatie sinds mensenheugenis het “lichaam-geestprobleem” genoemd. Binnen dit domein zullen we in deze scriptie focussen op de bijdrage van Hilary Putnam (1926-2016).

Putnam beschouwde het *lichaam-geestprobleem* als analoog aan een *Turingmachine* in de jaren 1960. Hij betoogde dat elk gegeven argument over fysiek-mentale toestanden identificeerbaar is met een argument over structureel-logische toestanden van een bepaalde geconfigureerde computermachine.<sup>1</sup> Verder heeft Putnam bijgedragen aan de filosofie van het bewustzijn met zijn formulering van de *meervoudige realiseerbaarheidshypothese* (MRH). Die stelt dat sommige fysische objecten mentaal identieke toestanden kunnen realiseren, zonder dat ze daarom dezelfde fysische eigenschappen hebben.<sup>2</sup>

Putnams analogie en hypothese leidden tot het *functionalisme*, het eerste standpunt in de filosofie van de geest dat mentale toestanden opvat als functionele toestanden. Net zoals longen, nieren en harten biologische functies vervullen, respectievelijk zuurstof in de bloedsomloop brengen, bloed zuiveren en bloed rondpompen, vervullen de mentale toestanden pijn en honger ook functies, respectievelijk lichamelijke schade signaleren zodat het organisme iets zal doen om het te genezen, en lichamelijk gebrek aan voedsel signaleren zodat het organisme iets zal doen om voedsel tot zich te nemen. Bijgevolg kan het functionalisme worden gebruikt om functionele uitspraken te doen over mentale toestanden: deze toestanden dienen de overleving van het organisme.

### Zijn kwalitatieve waarnemingen functionalistisch verklaarbaar?

Het functionalisme heeft nog geen perfecte oplossing geboden voor het lichaam-geestprobleem. Want als alle mentale toestanden volledig samenvallen met hun functie, dan gebeurt hetzelfde met “fenomenale” mentale toestanden. Ned Block (1942-heden) beweert echter dat er “fenomenale” mentale

---

<sup>1</sup> Hilary Whitehall Putnam, “Minds and Machines”, in *Dimensions of Mind: A Symposium*, bew. Sidney Hook (New York: University Press, 2016), 148-179.

<sup>2</sup> Hilary Whitehall Putnam, *Mind, Language and Reality: Philosophical Papers, Volume 2* (Cambridge: University Press, 1975).

toestanden zijn die “kwalia” bevatten, die als zodanig niet functioneel hoeven te zijn.<sup>3</sup> – Ik leg uit wat deze begrippen betekenen.

Immanuel Kant (1724-1804) bracht in de 17e eeuw naar voor dat onze zintuiglijke ervaring wordt gevormd door de geest, waarbij Kant een onderscheid maakte tussen de “noumenale” en de “fenomenale” wereld.<sup>4</sup> De *noumenale* wereld omvat de dingen die buiten ons bewustzijn of ons kenvermogen liggen – wij hebben geen directe ervaringsrelatie met de noumenale wereld. De *fenomenale* wereld omvat de zintuiglijk doorgegeven ervaring, verwerkt door ons geestelijk vermogen en ons fysieke zenuwstelsel.

De kwalitatieve eigenschappen die door ons worden waargenomen wanneer wij fenomenale verschijningen ervaren, zijn *kwalia*. Voorbeelden zijn geuren, geluiden, smaken, aanrakingen en kleuren. Voorbeelden van individuele kwalia (kwele) zijn de sensatie van de kleur rood, de smaak van zoetheid, de aanraking van warmte, enz.

De (vermeende) onverenigbaarheid van de notie van kwalia enerzijds, en het functionalisme anderzijds, vond haar uiting in twee bekende argumenten in de eigentijdse filosofie: het “afwezige kwalia-argument” en het “omgekeerde kwalia-argument”. De precieze inhoud van het “afwezige kwalia-argument” wordt uiteengezet in hoofdstuk 2, en het “omgekeerde kwalia-argument” in hoofdstuk 3.

### De probleemstelling, oplossingsmethode en relevantie van deze scriptie.

De probleemstelling van deze scriptie luidt als volgt: Hoe zijn de twee kwalia-argumenten te rijmen met het “machinetoestandsfunctionalisme” (MT-functionalisme), of zijn ze er onverenigbaar mee? Om achter een antwoord op die vraag te komen, bekritiseer ik de functionalistische argumenten tegenover de kwalia-argumenten en voeg aan de eerstgenoemde argumenten verdedigingen toe om een sterkere ondersteuning te bieden. Het uiteindelijke doel is dus om te tonen hoe men recht kan doen aan beide kwalia-argumenten, en toch het “MT-functionalisme” kan behouden.

De methode zal als volgt worden ontvouwd. In hoofdstuk 1 zal ik het “MT-functionalisme” van Putnam beschrijven. In hoofdstuk 2 het “afwezige kwalia-argument” van Ned Block (1942-heden), aan de hand van David Chalmers (1966-heden). Daarna behandel ik een tegenargument van William Lycan (1945-heden). En tenslotte, mijn eigen kritiek op en aanvulling van het tegenargument. In hoofdstuk 3 beschrijf ik het “omgekeerde kwalia-argument” van John Locke (1632-1704). Daarna het tegenargument van Michael Tye (1950-heden). Ik eindig hoofdstuk 3 met mijn kritiek op en aanvulling van het

---

<sup>3</sup> Ned Joel Block, “Troubles with Functionalism”, *Philosophy of Science* 9, (1978): 261-325.

<sup>4</sup> Immanuel Kant, *Critique of Pure Reason*, vert. Marcus Weigelt (London: Penguin Classics, 2003).

tegenargument. Ik sluit de scriptie af met een samenvatting en enkele onderzoeksvragen voor verdere overweging.

De relevantie van deze scriptie ligt in haar vermogen om een minimale verdediging te bieden die nuttig is voor hedendaagse functionalistische denkers. Afgezien van het feit dat het functionalisme al dan niet een populaire aanhang heeft binnen de hedendaagse filosofie van de geest, functioneert de functionalistische visie als een verklarende factor in wetenschappelijke en fenomenologische zaken – en dat is niet minder dan nuttig.

## Hoofdstuk 1: Putnams “machinetoestandsfunctionalisme”.

In dit hoofdstuk zal ik Putnams “MT-functionalisme” definiëren. Met deze doctrine was Putnam een belangrijke voorloper van het functionalisme:

“[Alan] Turing’s [(1912-1954)] work was explicitly invoked by many theorists during the beginning stages of 20th century functionalism, and was the avowed inspiration for a class of theories, the “machine state” theories most firmly associated with Hilary Putnam (1960, 1967) that had an important role in the early development of the doctrine.”<sup>5</sup>

Ik heb dit hoofdstuk in zes secties verdeeld. De eerste sectie gaat over het concept van de “Turingmachine”. Het tweede over Putnams analogie tussen geesten en machines. De derde over diens pijnhypothese. De vierde over de randvoorwaarden van een functionele-toestandstheorie. De vijfde over de meervoudige realiseerbaarheid van mentale toestanden. En de zesde over mijn opvatting van het “MT-functionalisme”.

### Wat is een Turingmachine?

In 1960 verscheen het boek *Dimensions of Mind: A Symposium*, met daarin Putnams hoofdstuk “Minds and Machines”. Hierin stelde hij de analogie aan de orde tussen de toestanden van een “Turingmachine” en die van de mens. Ik zal eerst beschrijven wat hij bedoelde met een “Turingmachine”.

Turing zat in een wetenschappelijk team tijdens de Tweede Wereldoorlog. Hij was vooral verantwoordelijk voor het breken van de Enigma code die door Nazi Duitsland werd gebruikt. Het team had succes, mede dankzij Turing als grondlegger van de moderne computerwetenschap. Op 23-jarige leeftijd stelde Turing voor dat het idee van een “Turingmachine”, die elk computerapparaat of -programma kan emuleren, in de toekomst daadwerkelijk kan bestaan. Ongeveer 80 jaar later, en “Turingmachines” zijn nog steeds een relevant concept binnen de theoretische computerwetenschap.<sup>6</sup>

Turing kwam met het idee van een *Turingtest*, een experiment dat antwoord moest geven op de vraag: “Kunnen machines denken?”

---

<sup>5</sup> Janet Levin, “Functionalism”, in *The Stanford Encyclopedia of Philosophy*, bew. Edward N. Zalta (Stanford: Metaphysics Research Lab, Winter 2021), URL: <https://plato.stanford.edu/archives/win2021/entries/functionalism/>.

<sup>6</sup> “Queen Pardons Computing Giant Alan Turing 59 Years After His Suicide”, World, NBC news, laatst bew. 1 april, 2014, URL: <https://www.nbcnews.com/news/world/queen-pardons-computing-giant-alan-turing-59-years-after-his-n241>.

“Is it theoretically possible for a finite state digital computer, provided with a large but finite table of instructions, or program, to provide responses to questions that would fool an unknowing interrogator into thinking it is a human being?”<sup>7</sup>

Een machine die de Turingtest zou doorstaan, zou als kunstmatige intelligentie kunnen worden beschouwd. Want als mensen intelligente wezens zijn, of daarop lijken, terwijl een kunstmatige machine dat menselijke verschijnsel ook op ons projecteert, dan zou men kunnen concluderen dat de kunstmatige machine intelligent is of zich intelligent gedraagt. Hier begint de analogie, en Putnam breidt haar uit. Ik toon die uitbreiding in de volgende sectie nadat ik Putnams begrip van een “Turingmachine” in deze sectie heb gedemonstreerd.

Volgens Putnam is een *Turingmachine* kort gezegd het volgende:

“[A] Turing machine is a device with a finite number of internal configurations, each of which involves the machine’s being in one of a finite number of *states*, and the machine’s scanning a tape on which certain symbols appear.”<sup>8</sup>

Deze “*states*” noem ik “toestanden”. Ik illustreer deze aan de hand van een door Putnam gebruikte machinetabel:

		A	B	C	D
(s <sub>1</sub> )	1	s <sub>1</sub> RA	s <sub>1</sub> LB	s <sub>3</sub> LD	s <sub>1</sub> CD
(s <sub>2</sub> )	+	s <sub>1</sub> LB	s <sub>2</sub> CD	s <sub>2</sub> LD	s <sub>2</sub> CD
(s <sub>3</sub> )	blank space	s <sub>3</sub> CD	s <sub>3</sub> RC	s <sub>3</sub> LD	s <sub>3</sub> CD

De eenhoofdige letters “A”, “B”, “C” en “D” in de bovenste rij zijn “logische” toestanden – *logisch* omdat de “deterministische regels” van de machinetabel bepalen in welke volgorde de logische toestanden elkaar opvolgen en wat wanneer wordt uitgevoerd. De vakjes eronder zijn instructies, de *deterministische regels*, zoals “s<sub>2</sub>LD”, wat betekent: “Druk het symbool s<sub>2</sub> af op het vierkant dat u op dat moment aan het scannen bent (na het wissen van het symbool dat er op dat moment op staat), en ga

<sup>7</sup> Janet Levin, “Functionalism”, URL: <https://plato.stanford.edu/archives/win2021/entries/functionalism/>.

<sup>8</sup> Putnam, “Minds and Machines”, 151.



verder met het scannen van het vierkant onmiddellijk links van het vierkant dat u net hebt gescand; ga ook naar toestand D”.<sup>9</sup>

Een machine<sup>10</sup> in het vorige geval heeft noodzakelijkerwijs een scanner, een printer en een band met aparte vierkanten nodig. Een logische beschrijving van een Turingmachine bevat echter geen specificatie van de fysieke aard van de machine:

“[A] given “Turing machine” is an *abstract* machine which may be physically realized in an almost infinite number of different ways.”<sup>11</sup>

Zodra een Turingmachine fysiek is gerealiseerd, heeft de machine naast logische ook “structurele” toestanden:

“[A physically realized Turing machine] has from the engineer’s point of view an almost infinite number of additional “states” (though not in the same sense of “state” – we shall call these *structural states*).”<sup>12</sup>

Een machine heeft dus logische toestanden – die abstract zijn, en structurele toestanden – die op zijn minst fysisch zijn.

### De analogie tussen geesten en machines.

Het onderscheid dat wordt gemaakt tussen logische en structurele toestanden is belangrijk omdat wij deze toestanden dan kunnen vergelijken met mentale en fysieke toestanden van mensen. Putnam demonstreert dit met overeenkomsten tussen structuur-detectieprocessen van een machine en de onze:

“A machine which is able to detect at least some of its own structural states is in a position very analogous to that of a human being, who can detect some but not all of the malfunctions of his own body, and with varying degrees of reliability.”<sup>13</sup>

Stel dat de machine afdruckt “vacuümbuis 312 is defect”. Dan, volgens Putnam, komt de vraag: “How did the machine ascertain that vacuum tube 312 failed?”<sup>14</sup> Het antwoord geldt, volgens Putnam, evenzeer voor de mens als voor de machine. Ik illustreer de structuur-detectieproces analogie aan de hand van het

---

<sup>9</sup> Ibid., 151-2.

<sup>10</sup> Vanaf hier verwijst “machine” altijd naar een fysisch gerealiseerde Turingmachine.

<sup>11</sup> 159.

<sup>12</sup> Ibid.

<sup>13</sup> 160.

<sup>14</sup> Ibid.

door Putnam beschreven mankementendetectieproces van de machine (A); en het verwondingsdetectieproces van de mens via een neurologische beschrijving van pijn (B).

(A) In het geval van een machine wordt het fysische mankement in de structurele toestand gedetecteerd met elektronische zintuigen, waarna de zintuiglijke gegevens worden afgedrukt op de invoerband. Vervolgens scant de machine de band en interpreteert deze programmatisch, zodat zij gegevensrelevante instructies uitvoert die bij bepaalde logische toestanden horen, mede dankzij de machinetabel.<sup>15</sup>

(B) Als iemand zegt “mijn hand is verbrand”, dan is de persoon in kwestie zich daarvan bewust geworden door het proces van pijn gewaarwording dat begint met de fysieke beschadiging van zenuwuiteinden “A-delta vezels” en “C-vezels”, die vervolgens signalen naar de hersenen sturen via zenuwbanen. Op het moment dat het pijnreguleringsysteem van de hersenen de ontvangen signalen deelt met het bewuste deel, d.w.z. de thalamus die de signalen deelt met de hersenschors, is de persoon in kwestie zich bewust van het verbranden van zijn/haar hand. Vervolgens sturen de hersenen signalen naar de spieren zodat ze zich afwenden van de oorzaak van de pijn; zoals het wegtrekken van uw hand van kokend water of een hete plaat.<sup>16</sup>

Hoe de machine vaststelt dat vacuümbuis 312 defect is, hangt dus af van intern-fysische zintuigen die gegevens over het structurele mankement doorgeven aan de scanprocessor, die de gegevens vervolgens interpreteert als logische informatie op basis van de interne programmering van de machinetabel. Op dezelfde manier hangt de vaststelling van mijn verbrande hand af van intern-fysieke zintuigen die signalen over de fysieke verwonding doorgeven aan de hersenen, waarna de hersenen de signalen interpreteren als mentale informatie op basis van interne programmering van mentale toestanden. Dit verwondingsdetectieproces is analoog aan het mankementendetectieproces.

De analogie tussen structuur-detectieprocessen was een voorbeeld van het vaststellen van overeenkomsten tussen machinetoestanden en mensentoestanden. In de volgende vergelijking demonstreer ik wat ik de beschrijving-methodologie analogie noem, waar Putnam overeenkomsten aantoonde tussen beschrijvingen van mensen en machines.

Putnam betoogt dat de menselijke psychologie vanuit twee perspectieven beschreven kan worden: de “behavioristische houding” en de “klassieke psychologie”:

---

<sup>15</sup> 159-160.

<sup>16</sup> James C. Watson, “Overview of Pain”, laatste herziening: april, 2020, URL: <https://www.merckmanuals.com/home/brain,-spinal-cord,-and-nerve-disorders/pain/overview-of-pain?redirectid=553?ruleredirectid=30>.

“The “behavioristic” approach [aims] at eventually providing a complete physicalistic description of human behavior. [In] terms of “mental states” [and] “impressions” – [a] description which would specify the laws controlling the order in which the states succeeded one another, [was] in fact the program of [“]classical psychology[”] to provide.”<sup>17</sup>

Putnam toont vervolgens de twee mogelijke perspectieven om machines te beschrijven – namelijk dat van ingenieurs en dat van logici, en verbindt ze analoog met de twee beschrijvingen van de menselijke psychologie:

“[There] are two possible descriptions of the behavior of a Turing machine – the engineer’s structural blueprint and the logician’s “machine table” – so there are two possible descriptions of human psychology.”<sup>18</sup>

De structurele blauwdruk van de ingenieur wordt vergeleken met de behavioristische beschrijving, en de machinetabel van de logicus met het programma van de klassieke psychologie. En aangezien een machinetabel logische toestanden definieert, lijkt het mij aannemelijk dat het programma verwijst naar het concept van ons denksysteem dat mentale toestanden definieert.

Met de beschrijving van de structuur-detectieproces analogie en de beschrijving-methodologie analogie, heb ik twee illustratieve delen van Putnams mens-machine analogie laten zien:

“The analogy [between] logical states [and] mental states, [on] the one hand, and structural states [and] physical states, [on] the other, is one that I find very suggestive. [Further] exploration of this analogy may make it possible to further clarify the notion of a “mental state”[.]”<sup>19</sup>

Het zojuist gegeven citaat is waarschijnlijk voldoende om de essentie van de analogie te kennen, want het is een behandeling van het lichaam-geest probleem: beide kanten van het probleem kunnen worden geanaliseerd met een machine. “Further exploration”, met betrekking tot de analogie en het begrip van de mentale toestand, heeft Putnam in een andere tekst aangetoond. Ik behandel die in de volgende secties.

---

<sup>17</sup> Putnam, “Minds and Machines”, 161.

<sup>18</sup> Ibid., 160-1.

<sup>19</sup> Ibid.

## De pijnhypothese.

De tekst waarin Putnam de mens-machine analogie verder uitdiept met betrekking tot het begrip van mentale toestanden heet “The Nature of Mental States” (1967).<sup>20</sup> In deze sectie wordt die tekst besproken om het “MT-functionalisme” te verduidelijken. Daarbij maakt hij opnieuw gebruik van de Turingmachine, hetgeen aansluit bij de vorige secties.

Putnam legt uit wat mentale toestanden eigenlijk zijn. Hij bestudeert deze door de toestand van pijn als casestudy te gebruiken. Hij stelt de vraag: “Is pijn een hersentoestand?” Hij antwoordt:

“I [argue] that pain is not a brain state, in the sense of a physical-chemical state of the brain (or even of the whole nervous system), but another *kind* of state entirely. I propose the hypothesis that pain, or the state of being in pain, is a functional state of a whole organism.”<sup>21</sup>

Met zijn antwoord veronderstel ik drie punten:

- (1) Pijn is geen *hersentoestand*;
- (2) Pijn kan niet fysicalistisch verklaard worden omdat (1);
- (3) Pijn is een functionele toestand van een heel organisme.

Identiteitsuitdrukkingen van pijn en lichaam zouden geen recht doen aan wat pijn is als (1) waar is. Een hersentoestand-hypothese zoals “c-vezels die vuren of die gestimuleerd worden zijn identiek aan pijnsensaties”, zou de oorzaak van menselijke pijn kunnen verklaren, maar zou weinig voldoening geven in de verklaring van wat pijn überhaupt is. Waarom zou een wezen per se c-vezels nodig hebben om pijn te voelen? Als u zeker weet dat u pijn hebt, kunt u toch niet louter uit dat pijngevoel afleiden of uw c-vezels al dan niet gestimuleerd worden? U hoeft niet te weten hoe u er lichamelijk aan toe bent of wat er lichamelijk gebeurt om te weten dat u pijn voelt. U hoeft zich alleen maar te wenden tot uw innerlijke ervaring van het gevoel van pijn om dat te weten. Dus pijn kan niet alleen een fysieke toestand van de hersenen zijn.

Als gevolg van (1) zijn fysicalistische verklaringen van pijn niet volledig, volgens (2). Want stel dat we ontdekken dat levensvorm X zich lijkt te gedragen alsof hij “pijn” heeft nadat we hem geslagen hebben, en we ontdekken geen c-vezels in X, zouden we dan tot de absurde conclusie komen dat X *toch* geen pijn voelt? Zouden we niet eerder uit het gedrag van X afleiden hoe pijn van functionele waarde is in

---

<sup>20</sup> Hilary Whitehall Putnam, “The Nature of Mental States”, in *Readings in Philosophy of Psychology, Volume I*, bew. Ned Block (Cambridge: Harvard University press, 1980), 223-231.

<sup>21</sup> *Ibid.*, 226.

relatie tot X zelf? Putnam antwoordt positief met (3), wat hij onderbouwt met de pijnhypothese – wat ik nu zal aantonen.

Hij bouwt de pijnhypothese met behulp van de wiskundige “Probabilistische Automaat” (PA):

“[The] notion of a Probabilistic Automaton has been generalized to allow for “sensory inputs” and “motor outputs” – that is, the Machine Table specifies, for every possible combination of a “state” and a complete set of “sensory inputs”, an “instruction” which determines the probability of the next “state”, and also the probabilities of the motor outputs”.<sup>22</sup>

In principe moet een PA zoals een Turingmachine worden opgevat als een abstract concept. Er is geen fysisch object dat als *de* PA kan worden aangewezen. Aan de andere kant kan een fysisch object, met de juiste eigenschappen, gekarakteriseerd worden als *een* PA. Bijvoorbeeld, elke haai of elke kat kan een PA genoemd worden. Ze zijn in staat om de toestand van pijn binnen te gaan, maar dit betekent niet dat wanneer ze dat doen, ze hetzelfde proces ondergaan. De waarschijnlijkheid van toestand-overgangen zal in de meeste gevallen sterk verschillen, omdat het inwendige van een haai en een kat op verschillende omstandigheden zijn ingesteld. Daarom neem ik aan dat functionalistische wetenschappers deze wezens anders zouden “Beschrijven”. Putnam legt uit hoe een dergelijke “Beschrijving” abstract wordt opgevat:

“A *Description* of  $S$  where  $S$  is a system, is any true statement to the effect that  $S$  possesses distinct states  $S_1, S_2, \dots, S_n$  which are related to one another and to the motor outputs and sensory inputs by the transition probabilities given in such-and-such a Machine Table.”<sup>23</sup>

Met andere woorden, een Beschrijving omvat een machinetabel die laat zien hoe elke toestand kan worden bereikt als resultaat van een combinatie elementen van alle mogelijke zintuiglijke invoeringen, of gewaarwordingen, alle motorische uitvoeringen, of (verbaliserings)gedragingen – in gedachten of naar buiten toe, en alle andere toestanden. Ik illustreer als volgt.

Stel dat Jeffrey in de toestand “chagrijnig” is. Er zijn een aantal logische vervolgt toestanden gebaseerd op de combinatie van “chagrijnig zijn” en bepaalde gewaarwordingen. Bijvoorbeeld, Bilal laat koffie vallen over het witte flanel van Jeffrey (gwaarwording). En omdat Jeffrey op dat moment chagrijnig is (toestand), uit hij zijn brutale irritatie aan Bilal (gedrag). Aldus is Jeffrey overgegaan van de toestand

---

<sup>22</sup> Ibid.

<sup>23</sup> Ibid.

“chagrijnig zijn” naar “boos zijn”. Maar omdat elke toestand verschillende waarschijnlijke gevolgen met zich meebrengt, zou de volgende keer wanneer Kaan baklava op de nette schoenen van chagrijnige Jeffrey laat vallen, Jeffrey mogelijk overgaan naar de toestand van “vergevingsgezind zijn”; Jeffrey schaamde zich dat hij zich zo had uitgelaten over Bilal, dus hij zorgt ervoor dat hij de volgende keer dat hij chagrijnig is, zich niet meer zo uitdrukt. Zo hebben toestanden met alle bijbehorende relaties praktisch nut in ons vermogen om te functioneren. Ik ga terug naar het concept van de machinetabel in een Beschrijving.

De machinetabel van een Beschrijving verschilt van die van een Turingmachine omdat, bij een PA, de overgangen van toestanden meer afhangen van waarschijnlijkheden dan van deterministische regels - wat het omgekeerde geval is bij een Turingmachine. Putnam noemt de machinetabel van een Beschrijving de “Functionele Organisatie” (FO):

“The Machine Table mentioned in the Descriptions will then be called the Functional Organization of  $S$  relative to that Description, and the  $S_i$  such that  $S$  is in state  $S_i$  at a given time will be called the Total State of  $S$  (at that time) relative to that Description.”<sup>24</sup>

Stel dat u de volledige Beschrijving van Jeffrey kent, dat wil zeggen, u weet ook in welke Totale Toestand (TT) hij zich op dat moment bevindt, dan weet u precies hoe hij zich waarschijnlijk zal gedragen. U zou, met absolute kennis van de wereld, elke verandering in de TT kunnen voorspellen. Ondanks we feitelijk minder kennis hebben over Jeffrey en de wereld, kunnen we Beschrijven hoe hij zich waarschijnlijk zal gedragen tijdens welke toestanden en welke zintuiglijke invoeringen.

Ik heb beschreven wat Putnam bedoelt met een PA, Beschrijving, FO en TT. Nu kan ik aan de hand van die begrippen zijn pijnhypothese technisch plaatsen.

De pijnhypothese stelt dat elk organisme dat pijn kan voelen een PA met een bepaalde FO is. Het type FO dat organismen bevatten wordt ingekaderd door Beschrijvingen. Elke Beschrijving definieert een verzameling van mogelijke zintuiglijke invoeringen. Zodra een bepaalde combinatie van deze verzameling wordt geactiveerd met een bepaalde TT, komt het organisme in de toestand van pijn.<sup>25</sup> Ik heb het nu over pijn, maar er zal een combinatie moeten zijn per mogelijke activering naar welke mogelijke toestand dan ook.

---

<sup>24</sup> Ibid.

<sup>25</sup> 227.

### Randvoorwaarden van een functionele-toestandstheorie.

Putnam beschouwt de pijnhypothese als minder vaag dan een hersentoestand-hypothese omdat de pijnhypothese “[is] far more susceptible to investigation of both a mathematical and an empirical kind.”<sup>26</sup> Niettemin, de omzetting van de hypothese in theorie zou de volgende moeilijke stap zijn:

“The difficult step [will] be to pass from models of *specific* organisms to a *normal form* for the psychological description of organisms – for this is what is required to make [our hypothesis of pain] precise.”<sup>27</sup>

Een *functionele-toestandstheorie* moet in staat zijn om (waarschijnlijke) toestanden te onderscheiden, en hun mogelijke realisatie te verbinden met bepaalde Beschrijvingen; toestanden zijn kenmerkend voor de functionele aard van het organisme in relatie tot zijn Beschrijving. Aldus heeft een functionele-toestandstheorie tegenover een hersentoestand-theorie het voordeel dat zij realiseerbare mentaal identieke toestanden toeschrijft aan fysiek niet-identieke organismen. Dit komt omdat het mentaal identieke eigenschappen verklaart vanuit een abstract begrepen overeenkomst tussen organismen hun functionerende aard. Bijvoorbeeld, waar een hersentoestand-theorie pijn niet kan verklaren bij wezens zonder herkenbare pijnreceptoren, kan een functionele-toestandstheorie pijn verklaren door het doorgronden van verschillende systeem relaties en hun functie voor het geheel. Het maakt dus niet uit van welk materiaal een wezen is gemaakt, als ik maar weet hoe het wezen in elkaar steekt.

Ik heb laten zien hoe Putnam pijn identificeert als een functionele toestand, en hoe voor hem pijn geen hersentoestand is. Hij vergelijkt zijn methode ook met het verklaren van pijn uit puur gedragsonderzoek. Ik zal het toelichten.

Een *gedragsdispositie* is de neiging om op overeenkomstige zintuiglijke prikkels te reageren met een bepaald soort gedrag. Ik zou denken dat een gedragsdispositie-theorie zou lijken op een functionele-toestandstheorie en door Putnam zou worden beschouwd als een (gedeeltelijk) juiste kijk op organismen, maar volgens Putnam is dat niet zo:

“Not only does the behavior-disposition theory seem hopelessly vague; if the “behavior” referred to is peripheral behavior, and the relevant stimuli are peripheral stimuli (e.g., we do not say anything about what the organism will do if its brain is operated upon), then the theory seems clearly false.”<sup>28</sup>

---

<sup>26</sup> Ibid.

<sup>27</sup> Ibid.

<sup>28</sup> 230.

Een *perifeer gedragskenmerk* is een kenmerk waarvan de aan- of afwezigheid niet veel andere kenmerken impliceert. Evenzo brengt een *perifere stimulus* niet een verscheidenheid van karakteristieke gedragingen voort. Als iemand bijvoorbeeld gelukkig is, kan dat betekenen dat hij of zij de wereld goed inschat of een optimistisch gevoel voor humor heeft – maar niet veel anders.

Er zijn een aantal bezwaren tegen het soort gedragsdispositie-theorieën wanneer zij gaan over perifeer gedrag. Men kan, op grond van de bezwaren, een dergelijke theorie niet gebruiken om de complexiteit van gedrag te verklaren zonder de mentale toestanden in vraag te stellen. Ik ken twee algemene bezwaren, en één bezwaar van Putnam zelf:

Algemene bezwaren:

- (1) Stel u eens voor dat u naar een theatervoorstelling kijkt. Een gewoon uitzijende man ligt in een stoel terwijl een kaakchirurgachtig uitzijende vrouw in de open mond van de man kijkt. Ze lijkt met gereedschap een kies uit de man te trekken. De man schreeuwt, maakt spastische bewegingen, en de tranen rollen over zijn wangen. Dit alles, ter herinnering, tijdens uw observatie van een theatervoorstelling. De man acteert. Daarom kan het gelijktijdig waarnemen van de relatie tussen pijnprikkels en pijngedrag niet het argument zijn voor de mentale toestand van pijn. Immers, pijn gedrag acteren is niet alleen voor een theater acteur.
- (2) Andrea beweert angst en opluchting te voelen: opluchting dat de linkerarm weer in de kom is geduwd, wat op de lange duur medisch beter is, en angst voor wat er straks zal gebeuren met de ontwrichte rechterarm, die pijn zal doen. Andrea zou twee soorten gedrag moeten vertonen: opluchting en angst. Zou een overeenkomstige complexe mentale toestand geen tegenstrijdige verklaring zijn? Zo niet, hoe kan men dan een dergelijke complexiteit concluderen met louter gedragsobservatie? Dat lijkt onmogelijk.

Het bezwaar van Putnam:

- (3) Putnam beschrijft in een gedachte-experiment *Super-Spartanen*, die niet terugdeinzen wanneer zij pijn hebben; zelfs niet het geringste knippen van een ooglid is waarneembaar wanneer hun testikels in de staafmixer gaan. Deze mogelijkheid wijst in de richting dat de mentale toestand van pijn van meer voorwaarden afhankelijk is dan louter gedragsaspecten en er niet noodzakelijk van afhangt.<sup>29</sup>

---

<sup>29</sup> Hilary Whitehall Putnam, "Brains and behavior", in *Analytical Philosophy: Second series*, bew. Ronald J. Butler (Oxford: Basil Blackwell, 1965), 332.



De bezwaren laten zien waarom de aard van mentale toestanden niet kan worden verklaard door zuiver gedragswetenschappen – omdat deze slechts concluderen dat gedrag X lijkt op pijn lijden, terwijl de bezwaren mogelijkheden bieden waarbij (het begrip) pijn niet kan worden gecorreleerd of verklaard via gedrag. Het MT-functionalisme vervalt niet in hetzelfde probleem, omdat het pijn kan identificeren zonder het begrip pijn überhaupt te gebruiken:

“[The] functional state we have in mind is the state of receiving sensory inputs which play a certain role in the [FO] of the organism. This role is characterized, at least partially, by the fact that the sense organs responsible for the inputs in question are organs whose function is to detect damage to the body, or dangerous extremes of temperature, pressure, etc., and by the fact that the “inputs” themselves, whatever their physical realization, represent a condition that the organism assigns a high disvalue to.”

Een dergelijke uitwerking van pijn zou kunnen dienen als de wet van pijn in het (door mij verzonnen) boek der psychologische wetten – als zodanig toegepast op alles wat een psyché zou hebben. Dit boek zou impliceren dat een verzameling van “empirisch gegeven systemen” (EGSn), die elk een psyché met een passende FO bevatten, kunnen komen tot de functionele toestand van pijn. Als dit waar zou zijn, dan zou dat betekenen dat verschillende “EGSn” ten minste eenzelfde psychische toestand kunnen bereiken. Het zou dan niet moeten uitmaken of ze fysisch niet-identiek zijn. Zolang de FOs dezelfde mogelijkheden van functionele realisaties behouden, kunnen we concluderen dat de aard van mentale toestanden een systeem-relacioneel functioneel karakter heeft – een karakter dat, dankzij fysisch verschillende “EGSn”, meervoudig realiseerbaar blijkt te zijn.

Ik gebruik de term “EGSn” om duidelijk te maken dat het niet noodzakelijk om een organisme hoeft te gaan; het kan evengoed een machine zijn. – In de volgende sectie wordt dit uitgelegd.

### **De meervoudige realiseerbaarheid van mentale toestanden.**

De MRH stelt dat “fysische objecten” van verschillende “structuur en samenstelling”, met verschillende “fysische eigenschappen”, mentaal identieke toestanden kunnen realiseren. Ik leg dit uit.

Wij noemen een “fysisch object” met “structuur en samenstelling” een *empirisch gegeven systeem* (EGS) dat, in dit geval, mentale toestanden kan realiseren, zoals een schubdier, okapi, mens, walrus, wijngaardslak, en elk ander zintuiglijk waarneembaar ding – instrumenteel waarneembaar of niet – dat als PA kan worden gekarakteriseerd. Het maakt het tot een *fysisch object* omdat we het kunnen waarnemen als een feitelijk ding. En het heeft *structuur en samenstelling* omdat het meerdere delen bevat die samen maken wat het systeem vormt en kan doen; zoals het skelet voor een deel de mens

vormt en menselijke activiteiten mogelijk maakt. Twee 3D-afdrukken van hetzelfde skelet, zouden twee fysische objecten tonen met dezelfde structuur en samenstelling. Indien het eerste skelet echter gemaakt is van materiaal X, en het tweede van materiaal Z, zouden zij alleen verschillen in *fysische eigenschappen* – mits deze niet tegelijkertijd een verschil in structuur en samenstelling impliceren.

Twee fysisch niet-identieke EGSn kunnen, indien mogelijk, volledig overeenstemmen in hun FOs – hun toestanden en geassocieerde relaties. Ik illustreer dit aan de hand van het begrip “functioneel isomorfisme”, dat wordt beschreven in een artikel van Putnam: “Philosophy and our mental life”, in *Mind, Language and Reality*:<sup>30</sup>

“Two systems are functionally isomorphic if *there is a correspondence between the states of one and the states of the other that preserves functional relations.*”<sup>31</sup>

Met andere woorden, een duif geboren als een samenvoeging van biochemische componenten, en een drone gemaakt van samengestelde kunstmatige componenten, zijn functioneel isomorf als (1) voor elke toestand in de duif er een overeenkomstige toestand is in de drone, en vice versa, en (2) voor elke toestand *x* die gevolgd wordt door *y* in de duif, dezelfde overgang gebeurt in de drone. Het doet er niet toe hoe verschillend de fysische realisaties van die toestanden zijn.<sup>32</sup> De functionele toestand “omhoogvliegen” kan bijvoorbeeld fysisch worden gerealiseerd door duivenvleugels of door dronepropellers.

In essentie gaat het er bij de MRH om dat er minstens twee fysisch niet-identieke EGSn zijn die minstens één mentaal identieke toestand kunnen realiseren – met dien verstande dat een mentale toestand hier wordt opgevat als functioneel. Geen empirische vaststelling van volledig functioneel isomorfisme doet dus niets af aan de MRH.

### Putnams “machinetoestandsfunctionalisme”.

Nu ik delen van Putnams functionalisme heb beschreven, ga ik kort zijn “MT-functionalisme” beschrijven.

*MT-functionalisme* is een mechanistisch perspectief binnen de filosofie van de geest dat in het algemeen stelt dat afgebakende abstracte toestanden realiseerbaar zijn door verschillende EGSn vanwege hun gelijkenissen in functionele relaties. Dat functioneren in organische EGSn wordt verklaard door fysieke en mentale beschrijvingen door fysici en psychologen. Het functioneren in kunstmatige EGSn wordt

---

<sup>30</sup> Putnam, *Mind, Language and Reality*: [...], 291-303.

<sup>31</sup> *Ibid.*, 291.

<sup>32</sup> 292-3.

verklaard door technische en logische beschrijvingen van ingenieurs en logici. Deze verklaringen van twee soorten EGSn stellen ons in staat een analogie te trekken tussen toestanden van mensen en toestanden van machines, waaruit overeenkomsten in functionele relaties worden afgeleid. Hierdoor kan het MT-functionalisme dergelijke abstracte toestanden identificeren als functionele toestanden, wat impliceert dat onze eigen mentale toestanden dat ook zijn. Dit universalistische functionalisme overziet de gevolgtrekking van psychologische wetten uit de reeks Beschrijvingen voorgesteld door MT-functionalisten.

## Hoofdstuk 2: Een behandeling van het afwezige kwalia-argument.

Block schreef het afwezige kwalia-argument tegen het MT-functionalisme. Ik ga in op zijn bezwaar aan de hand van een tekst van Chalmers: “Absent Qualia, Fading Qualia, Dancing Qualia” in *Conscious Experience*. Vervolgens op een argument tegen het bezwaar van Block, namelijk dat van Lycan. Dit staat beschreven in “Homunctionalism and “Qualia””, in *Consciousness*. Tenslotte bekritiseer ik Lycans tegenargument, en bied dan een positieve aanvulling. Door dit te doen, hoop ik een sterkere verdediging te bieden voor het MT-functionalisme.

### Het afwezige kwalia-argument.

Chalmers refereert naar Blocks “Troubles with Functionalism” in *Readings in Philosophy of Psychology*, en legt aan de hand daarvan Blocks gedachtevoorstel uit:

“Block (1980) points out that the [FO] of the brain might be instantiated by the population of China, if they were organized appropriately, and argues that it is bizarre to suppose that this would somehow give rise to a group mind.”<sup>33</sup>

Block zegt, met andere woorden, dat men de volledige functionaliteit van onze hersenen dienovereenkomstig zou kunnen realiseren via de bevolking van China, waarvan ik uitga dat elke Chinees een neuron is en de verbindingen met andere Chinezen axonen zijn. Het is belachelijk, volgens Block, dat de Chinese bevolking als geheel mentale toestanden zou realiseren. De Chinese-FO (C-FO) zou niet in staat zijn om bewustzijn te creëren, zelfs niet met replicatie van de hersenen-FO (H-FO). Bijgevolg is het mogelijk dat er niets is dat ervaart – de kwalia zijn afwezig.

Het essentiële punt van Block is dat twee EGSn die beide vergelijkbaar zijn in hun FO niet noodzakelijkerwijs bewuste systemen zijn – zelfs als één van de twee systemen dat wel is. Bewustzijn wordt niet simpelweg gecreëerd wanneer je bottom-up alle onderlinge relaties en functies van een bewustzijn-producerend systeem nabootst.

Het afwezige kwalia-argument leidt tot een discutabel element over de aanname van het MT-functionalisme. Als bijvoorbeeld wordt aangenomen dat een mens en een androïde functioneel isomorf zijn, dan gaan MT-functionalisten er volgens mij van uit dat ze allebei mentaal identieke toestanden ervaren. Als de androïde ze echter niet ervaart, hoe kunnen MT-functionalisten dan op realistische wijze de MRH aanvaarden, laat staan functioneel isomorfisme?

---

<sup>33</sup> David J. Chalmers, “Absent Qualia, Fading Qualia, Dancing Qualia”, in *Conscious Experience*, bew. Thomas Metzinger (Kansas: Allen Press, 1995), 311.

Als een MT-functionalist antwoordt dat bewustzijn of geen bewustzijn geen verschil maakt in de overeenstemming tussen functionele relaties en toestanden, denk ik dat de MT-functionalist een fout begaat. Dit komt omdat de TT van een mens mogelijke motorische outputs omvat die deels veroorzaakt worden door bewustzijnszaken:

“[The] behavior of the Machine (in this case, an organism) will depend not merely on the sensory inputs, but also on the [TT] (i.e., on other values, beliefs, etc.)[.]”<sup>34</sup>

“On other values, beliefs, etc.” doet mij denken aan zaken als politieke overtuigingen, ideale wereldbeelden, religieuze opvattingen, moralistische gedragssystemen, enz. Het lijkt me erg ingewikkeld om uit te leggen hoe (denk)gedrag zich verhoudt tot zulke aangelegenheden, maar minder ingewikkeld als ik het mag uitleggen met het begrip bewustzijn – dat, dankzij het afwezige kwalia-argument, een androïde waarschijnlijk niet heeft.

Het argument van de afwezige kwalia heeft dus betrekking op twee belangrijke kwesties:

- (1) De FO van een EGS-1 die overeenkomt met een FO van een EGS-2 met bewustzijn, maakt EGS-1 niet noodzakelijkerwijs ook een EGS met bewustzijn.
- (2) Als (1) waar is, kunnen kunstmatige EGSn en organische EGSn niet noodzakelijk overeenkomen in FO, ook al lijkt het zo, omdat een machine zonder bewustzijn a priori andere functionele relaties heeft dan een organisme met bewustzijn.

Ik heb laten zien hoe het argument van de afwezige kwalia afbreuk doet aan het MT-functionalisme. Het doet dat door de mens-machine analogie en de inherente aanname van bewustzijn in MRH in vraag te stellen. We wenden ons in wat volgt tot Lycan om te zien hoe we het argument van Block kunnen afzwakken.

### [Een functionalistisch bezwaar tegen het afwezige kwalia-argument.](#)

Ik geef eerst een beschrijving van de methode van het “homunculaire” functionalisme. Daarna van het tegenargument van Lycan.

Lycan gebruikt de beschrijving van Daniel Dennett (1942-heden) over de methodologie van een bepaald AI-onderzoeksteam, en brengt het hiërarchische karakter van de benadering van problemen naar voren. Dat wil zeggen, u begint met een hoe-vraag, en u stelt sub-hoe-vragen die eerst moeten worden opgelost, die sub-sub-hoe-vragen kunnen ontwikkelen als u de hogere vraag wilt oplossen, enz.<sup>35</sup> Lycan vertaalt

---

<sup>34</sup> Putnam, “The Nature of Mental States”, 230.

<sup>35</sup> William G. Lycan, *Consciousness* (Massachusetts: Institute of Technology, 1995), 39-40.

dit in een psychologisch gekarakteriseerd vermogen om het menselijk subject op verschillende niveaus te verklaren, wat ik de homunculaire visie noem:

“[The] homunculi or sub-personal agencies[,] that are needed to explain the subject’s having [certain] psychological abilities[,] [have] smaller homunculi in order to explain the previously posited molar behavior of the original homunculi, etc., etc.”<sup>36</sup>

Een voorbeeld van een hiërarchische indeling voor die opeenstapeling van niveaus gaat als volgt: Het menselijk systeem heeft of is een aantal subsystemen. Elke stap na de eerste is een dieper subsysteem. Deze worden hier van links naar rechts weergegeven: organisme; orgaansystemen; organen; weefsels; cellen; en chemicaliën. Een voorbeeld van homunculaire methodologie is dat het onbetrouwbaar is om de vorm van hoofdpijn te willen verklaren door het organisme alleen op normale grootte te bestuderen. We kennen immers verschillende vormen van hoofdpijn, bijvoorbeeld door hitte, cafeïne, hoge bloeddruk, enzovoort, die waarschijnlijk gewoon als “hoofdpijn” worden ervaren. Om uit te zoeken wat voor soort hoofdpijn een subject ervaart, suggereert de homunculaire visie om op een dieper niveau te kijken, bijvoorbeeld moleculair onderzoek. Anders zouden we niet in staat zijn om hormonale hoofdpijn op betrouwbare wijze te verklaren. Zo onderscheiden we verschillende genuanceerde toestanden als “hormonale hoofdpijn” die eerst gewoon als “hoofdpijn” werden gekarakteriseerd. Als zodanig kunnen diepere niveaus worden gebruikt om mentale toestanden (genuanceerder) te verklaren. Met dit in gedachten, parafraseer ik het homunculaire tegenargument van Lycan:

Als u een klein mens was, en u bevond zich in de linker occipitale kwab van Block, dan leek het u alsof u midden in een lege ruimte stond. Soms zag u een molecuul (dat er ongeveer uitzag als een cluster basketballen) voorbij suizen. Soms botsen clusters op elkaar en kaatsen ze tegen elkaar. Stel dat iemand u op dat moment zou suggereren dat u zich bevond in het visuele systeem van een enorm bewust wezen, wiens lichaam uitsluitend bestond uit die basketbalclusters, en dat het wezen een levendige en homogene rode visuele sensatie ervoer uitsluitend op grond van het feit dat die anders zo inerte basketbalclusters zo heen en weer suidden en stuiterden. Dit zou u waarschijnlijk volkomen absurd vinden, net zoals (volgens Lycan) het voorbeeld van de bevolking van China voor Block absurd zou lijken. *En u zou het mis hebben*, als Block bij goed licht voor een gladde rode muur zou staan.<sup>37</sup>

Lycan noemt de berusting van het oordeel van de kleine mens in de absurditeit van een bewust wezen *gestaltblindheid*: u kunt Block niet als persoon zien maar als een opeenstapeling van weinigzeggende brokken – dus u beseft niet dat het hele wezen ervaart. Zodra u in een kamer staat met Block, gewoon

---

<sup>36</sup> Ibid., 40.

<sup>37</sup> 51.

op menselijke schaal, begint u automatisch te *gestaltiseren*: u begrijpt plotseling dat die opeenstapeling van brokjes eigenlijk gewoon een subjectief-bewuste persoon was die Block heette, en dat hij inderdaad de kleur rood ervoer.<sup>38</sup>

Lycan argumenteert eigenlijk dat de conclusie van Blocks gedachte-experiment nog niet zo absurd is als Block het doet lijken, zolang u maar die niveaus van verklaring en begrip in acht neemt. Met andere woorden, Lycan impliceert dat als de Chinezen werkelijk in die configuratie zijn geplaatst en op die manier met elkaar communiceren, geheel analoog aan een brein, dat u dan op een bepaald niveau zou kunnen spreken van “bewustzijn” van dat geheel – maar dat dit buiten ons voorstellingsvermogen ligt. Zelfs als u op menselijke schaal zou gestaltiseren en met behulp van een vertaalhandleiding zou kunnen ontcijferen wat het gedrag van enkele Chinezen op een abstracter niveau voor de C-FO betekent, zou u nog steeds niet het functionerende geheel waarnemen, aldus Lycan:

“Block’s attention is focused on the hectic activities of the little men, and so he is seeing the homunculi-head as if through a microscope, rather than as a whole macroscopic person whose inner mechanisms are so finely articulated as not even to appear articulated at all!”<sup>39</sup>

### Kritiek en aanvulling.

De toevoeging van Lycan valideert het MT-functionalisme door de structureel-hiërarchische setting van een organisch EGS te karakteriseren als een inter-homunculaire organiserende eenheid. Een dergelijke eenheid is op zichzelf een aaneenschakeling van verschillende sub-instellingen die verschillende doelstellingen nastreven. Samen vormen zij de persoonlijke werkzaamheid van de hoofdentiteit van het EGS.

Deze multifunctionele complexiteit vereist dat wij de mogelijke vaststelling van tegenstrijdigheid van temporele functionele toestanden tussen de afzonderlijke samenwerkende homunculussen aanvaarden, aangezien cellen mogelijk activiteiten van een ander functioneel karakter kunnen verrichten dan moleculen. En we hoeven de mogelijkheid van externe manifestatie van bewustzijn door die homunculussen niet absurd te vinden, omdat ook wij een opeenstapeling zijn van homunculussen en bewust zijn. Met deze stappen kunnen MT-functionalisten meer terrein winnen in het bevestigen van correlatie tussen functionele toestanden en bewustzijn in homunculair samengestelde EGSn, en ook de inter-structurele relaties van functionaliteit op analoge wijze karakteriseren.

---

<sup>38</sup> Ibid.

<sup>39</sup> 52.

Na het gestaltiseren werd de man zich bewust van een andere subjectief-bewuste persoon, Block, die inderdaad de kleur rood ervoer. Zou gestaltiseren daarentegen kunnen leiden tot een voldoende representatieve waarneming van de C-FO? En zou het ons dan doen beseffen dat we een ander bewust wezen ervaren?

Lycan noemde, in onze vorige sectie, de werking van de C-FO als een macroscopisch persoon waarvan de mechanismen fijn gearticuleerd zijn. Het lijkt mij dat de eigenschap van fijne articulatie meer van toepassing is op een bewegend persoon dan op een bewegende groep mensen. De hersenen alleen al tellen ruwweg honderd miljard neuronen en tien keer zoveel gliacellen,<sup>40</sup> waar ik niets van merk, terwijl er vandaag de dag nog niet eens acht miljard mensen zijn – laat staan Chinezen. Het zou daarom fysiek onmogelijk zijn om de H-FO te repliceren en te gestaltiseren via de C-FO. Zelfs de gedachte eraan vereist een extreem gebruik van mijn verbeelding.

Waarom een niet empirisch gedachtenexperiment doen als we ons ook een bewustzijn kunnen voorstellen bij het observeren van de homunculaire organiserende eenheid van een insectenkolonie?

Daniel Friedman en Eirik Søvik introduceerden in 2021 de “Ant Colony Test” (ACT) als een rigoureuze test voor bewustzijn. Daarmee tonen ze aan dat sociale insectenkolonies aan veel van de voorwaarden voor bewustzijn voldoen waaraan mensen voldoen.<sup>41</sup> Hun artikel synthetiseert de ACT en het functionalisme. Ze doen dit door de aannames en implicaties van functionalistische neurowetenschappelijke theorieën van bewustzijn te beschrijven, waarbij ze neurale hersenritmes analyseren aan oscillaties in collectief koloniegedrag. Deze analogie kijkt naar overeenkomsten tussen samenwerkingsverbanden van neuronen en taakgroepen van mierenarbeiders die kwantitatief en kwalitatief zijn. De auteurs erkennen dat zowel in de hersenen als in de mierenkolonie hogere-orde uitkomsten worden gevormd door evolutie om adaptieve reacties op ecologisch belangrijke stimuli te optimaliseren, zoals het geval is met de dynamiek van de hersenen.<sup>42</sup>

“Should we take this Functionalist turn, then ant colonies do indeed appear to exhibit multi-scale oscillatory processes that could be considered to be similar evidence for colony consciousness. In other words, if the oscillations in the brain are (a proxy for) consciousness, then similar but slower processes in ant colonies could be considered as the basis of (a proxy for) consciousness as well.”<sup>43</sup>

---

<sup>40</sup> Suzana Herculano-Houzel, “The human brain in numbers: a linearly scaled-up primate brain”, *Frontiers in human neuroscience* 3, nr. 31 (november 2009): 1-11.

<sup>41</sup> Daniel Friedman & Eirik Søvik, “The ant colony as a test for scientific theories of consciousness”, *Synthese* 198, nr. 5 (7 februari 2019): 1-24.

<sup>42</sup> *Ibid.*, 8.

<sup>43</sup> *Ibid.*



Mijn evaluatie is als volgt. Van mijn ongelof in het bewuste wezen bij het waarnemen van de bewegende basketbal-clusters, tot mijn besef van het bewuste wezen dat eigenlijk een opeenhoping is van de bewegende clustermoleculen. Dit is een denkproces dat suggereert dat kleine deeltjes bewustzijn kunnen vormen. Toch blijft het een hypothese: we kunnen onszelf niet zover inkrimpen dat we in iemands brein gaan staan, en dat doen in onwetendheid over onze locatie. Blocks C-FO is ook een hypothese: het is feitelijk onmogelijk te concretiseren. Daarentegen kan de bewustzijnsrelatie van het gedrag van de mierenkolonie met het koloniebewustzijn theoretisch benaderd worden.

Zintuiglijke studie van mierenkolonies is mogelijk: waarnemen van mierenkolonies kan per mier, maar ook als geheel; elk niveau, zoals in het menselijk lichaam, kan bestudeerd worden. Een mieren-FO, een replicatie van de H-FO, is in feite mogelijk: voor elk minuscuul hersendeel is er wel een mier. En op basis van omvangrijke empirische informatie over het gedrag van mierenkolonies kan, zoals de ACT laat zien, vanuit een functionalistisch neurowetenschappelijk perspectief worden beargumenteerd dat voldoende overeenkomst in de analogie tussen het gedrag van mierenkolonies en hersengedrag mogelijk mierenkolonie-bewustzijn impliceert – zoals hersengedrag dat voor ons doet. Bijgevolg is een afleidbaar androïde bewustzijn via empirisch onderzoek van systematisch functionele (sub)organen die lijken op die van de mens misschien ook zo gek nog niet.

### Hoofdstuk 3: Een behandeling van het omgekeerde kwalia-argument.

Locke schreef het omgekeerde kwalia-argument in *An Essay Concerning Human Understanding*.<sup>44</sup> Ik ga dit argument toelichten met de negatieve implicaties voor het MT-functionalisme. Ik verduidelijk dit met het werk van Chalmers. Daarna ga ik verder met het tegenargument van Tye, in *Ten Problems of Consciousness: A Representational Theory of the Phenomenal Mind*.<sup>45</sup> Tenslotte bekijk ik dat tegenargument kritisch en vul het aan om hopelijk tot een sterkere verdediging van het MT-functionalisme te komen.

#### Het omgekeerde kwalia-argument.

Ik beschrijf kort wat Locke's "ideeën" en "kwaliteiten" zijn. Deze zijn relevant voor het betoog. Wat "ideeën" volgens Locke zijn, is te vinden in het vijfjnde en honderden pagina's tellende *Essay*. Vanwege deze reikwijdte benadruk ik slechts gedeelten van de concepten die relevant zijn voor ons onderwerp. Wanneer ik dus benadruk wat een "idee" is, betekent dat altijd "naast andere zaken" – dit geldt ook voor de "kwaliteiten".

*Ideeën* zijn mentale objecten; we zien een fenomeen en we denken dat dit fenomeen een appel uitstraalt, we hebben dus ideeën over de verschijnselen. *Simple* ideeën zijn uniforme verschijnselen en zijn niet te onderscheiden in andere ideeën; zoals een uniforme kleur rood. *Complexe* ideeën zijn combinaties van ten minste twee ideeën – of ze nu simpel of complex zijn; zoals een appel (combinatie van de ideeën rood, groen, koel, zoet, zuur etc.). Wij komen tot onze ideeën door ervaring, die twee vormen kan aannemen: (1) *uiterlijke* ervaring via de vijf algemene zintuigen en (2) *innerlijke* ervaring via de verrichtingen van de geest.

Een *kwaliteit* is het vermogen van een object om ideeën in ons op te wekken. Dus een pizza margherita, die ideeën kan produceren over roodheid, geelheid, warmte, stevigheid, textuur, platheid, rondheid, enzovoort, moet een set van overeenkomstige kwaliteiten hebben. *Primaire kwaliteiten* zijn kwaliteiten die een pizza margherita werkelijk bezit; het zijn kwaliteiten waar de pizza niet zonder kan, zoals stevigheid, beweeglijkheid, uitgestrektheid, en vorm. Primaire kwaliteiten staan los van mijn bewuste ervaringen die ik heb met het object. *Secundaire kwaliteiten* daarentegen bezit de pizza margarita niet werkelijk; het zijn kwaliteiten waarvan het bestaan afhangt van mijn uiterlijke en innerlijke ervaringsmogelijkheden, zoals ervaringen van temperatuur, geluid, smaak of kleur.

---

<sup>44</sup> John Locke, *An Essay Concerning Human Understanding: Clarendon Edition of the Works of John Locke*, bew. Peter H. Nidditch (Oxford: University Press, 1975).

<sup>45</sup> Michael Tye, *Ten Problems of Consciousness: A Representational Theory of the Phenomenal Mind*, 2de print (Massachusetts: Institute of Technology, 1996).

Ik paraphraseer omdat Locke in een oudere vorm van het Engels schrijft. Locke begint het omgekeerde kwalia-argument door te stellen dat het mogelijk is voor een mens-1, met een bepaalde fysieke toestand, om twee “omgekeerde” ideeën over secundaire kwaliteiten te verkrijgen tijdens de perceptuele waarneming van object-1 en object-2; *omgekeerd* omdat een mens-2 met betrekking tot perceptuele waarneming van dezelfde objecten eenvoudigweg twee ideeën verkrijgt die de tegenovergestelde ideeën zijn van mens-1. Bijvoorbeeld, wanneer mens-1 het eenvoudige idee blauw verkrijgt tijdens de perceptuele waarneming van een violet en geel tijdens de perceptuele waarneming van een goudsbloem, verkrijgt mens-2 de ideeën van de respectieve kleuren omgekeerd. En als mens-1 het idee van blauw voortbrengt in alle dingen die de structuur van een violet hebben, en mens-2 het idee van blauw voortbrengt in alle dingen die de structuur van een goudsbloem hebben, zouden zij regelmatig andere verschijningen onderscheiden met de kenmerken van blauw en geel – alsof de verschijningen precies dezelfde waren als de ideeën in de geesten van andere mensen.<sup>46</sup> – dat is waarom we spreken van omgekeerde kwalia.

Tegenover het argument van de afwezige kwalia, stelt het argument van de omgekeerde kwalia niet dat een succesvolle analogie tussen een kunstmatig EGS en een organisch bewust EGS geen bewustzijn in het kunstmatige EGS impliceert. Het argument beweert in feite dat zelfs als dergelijke EGSn functioneel isomorf gekarakteriseerd zijn, dit niet de mogelijkheid uitsluit dat één ervan anders ervaart. Als dit argument slaagt, betoogt Chalmers, zouden we terugkeren naar het identificeren van mentale toestanden als hersentoestanden: “Instead [of depending on the appropriate FO], the specific nature of experiences must be dependent on non-organizational properties, such as specific neurophysiological properties.”<sup>47</sup>

Het rampzalige effect van het omgekeerde kwalia-argument op het MT-functionalisme heeft volgens mij te maken met de mogelijke ambiguïteit die EGSn ontvangen van hun zintuiglijke ingangen. Want als functionele toestanden onder andere afhangen van zintuiglijke invoeringen, en verschillende EGSn bestaan in dezelfde omstandigheden, waaruit we zouden kunnen afleiden dat ze zintuiglijke invoer krijgen van dezelfde fysieke objecten, dan is dat nog geen garantie dat ze ook de verschijnselen van die zintuiglijke invoer (ongeveer) hetzelfde ervaren. Dit verzwakt de MT-functionalistische verklaring van meervoudig realiseerbare mentaal identieke toestanden. Als gevolg hiervan kunnen MT-functionalisten niet langer psychologische wetten afleiden uit Beschrijvingen.

### Een functionalistisch bezwaar tegen het omgekeerde kwalia-argument.

Tye karakteriseert het omgekeerde kwalia-argument als “it [being] susceptible to the criticism that it commits a hasty generalization.”<sup>48</sup> Het argument gaat er immers van uit dat het MT-functionalisme stelt

---

<sup>46</sup> Locke, *Essay*, 388-90.

<sup>47</sup> Chalmers, “Absent Qualia, [...]”, 311-2.

<sup>48</sup> Tye, *Ten Problems of Consciousness: [...]*, 201.

dat bepaalde groepen EGSn een relationele volledigheid van functioneel isomorfisme bezitten, ook al ervaren die systemen verschillende fenomenale eigenschappen. Het is echter duidelijk dat fenomenale verschillen kunnen leiden tot functionele verschillen. Tye demonstreert dit met twee illustraties:

- (1) Twee computers kunnen hetzelfde wiskundeprobleem oplossen met een wiskundeprogramma dat verschilt van het wiskundeprogramma van de andere computer. Tye concludeert: “At one gross level, the machines are functionally identical, but at lower levels, the machines can be functionally different.”<sup>49</sup>
- (2) Twee wijnproevers hebben ongeveer hetzelfde onderscheidingsvermogen; door smaak onderscheiden zij rode wijn van smoothies, thee, bier, witte wijn enz. Niettemin, wanneer zij een verscheidenheid van brouwsels krijgen voorgeschoteld, sommige met een fruitige smaak, andere leerachtig, sommige ziltig, enz., zullen zij bij het vergelijken ongetwijfeld ergens niet overeenkomstig reageren – mits men over een voldoende groot aantal brouwsels beschikt. Tye concludeert: “These more subtle differences in our responses reflect the fact that there are differences in the fine-grained operation of our gustatory systems.”<sup>50</sup>

Volgens Tye sluit het wijnvoorbeeld, evenals het computervoorbeeld, functioneel isomorfisme niet volledig uit, omdat er tot op zekere hoogte functioneel identieke eigenschappen zijn. Bovendien zou een MT-functionalist de eigenschap van niet-functioneel isomorfisme kunnen ontdekken via het juiste experiment:

“If Oddball’s experiences are fully phenomenally inverted with respect to color, relative to the rest of us, then he will actually experience blue as being on the left and yellow as being on the right. When he is asked which color is brighter, he will say that the one on the right is, whereas the rest of us will pick the one on the left.”<sup>51</sup>

Een bepaalde helderheid van kleur als zintuiglijke invoer is medeverantwoordelijk voor het genereren van een functionele toestand, zoals “dat is te helder om naar te kijken”. Daarom zijn EGSn met dergelijke omgekeerde spectrums in de eerste plaats niet functioneel isomorf, terwijl het omgekeerde kwalia-argument postuleert dat het MT-functionalisme dergelijke systemen als functioneel isomorf kenmerkt wanneer ze wel of geen omgekeerde kwalia ervaren.

---

<sup>49</sup> Ibid., 202.

<sup>50</sup> Ibid.

<sup>51</sup> 203-4.

Wat Tye volgens mij bedoelt is dat MT-functionalisten de premisse van het omgekeerde kwalia-argument aanvechten. Dit zou dan mogelijk de volgende premisse kunnen zijn: “Het MT-functionalisme identificeert functioneel identieke relaties tussen twee systemen, ongeacht hoe beide systemen fenomenen betreffende die relaties ervaren.” MT-functionalisme bestrijdt dit, volgens Tye, door precies aan te tonen hoe ervaringsverschillen aantoonbaar functionele verschillen opleveren, en als zodanig deze niet als functioneel identieke relaties identificeert.

### Kritiek en aanvulling.

Tye’s verdediging verzekert dat het MT-functionalisme een meer rigoureuze methodologie zou moeten hanteren bij pogingen om een veelheid van systemen als volledig functioneel isomorf te karakteriseren. Niettemin zou een volledig functioneel isomorfisme legitiem kunnen worden gekarakteriseerd tussen twee EGSn die niet alleen functioneel identieke organisaties hebben, maar ook structureel identieke toestanden vertonen. Als dit echter zo zou zijn, zou het MT-functionalisme worden uitgeschakeld – aangezien de mogelijkheid van functioneel isomorfisme tussen EGSn met verschillende structurele toestanden essentieel is.

Tye hangt een *representationalistische* theorie van het bewustzijn aan, een richting die *intentionaliteit*, de gerichtheid van mentale toestanden over iets, tracht uit te breiden tot het denken over het bewustzijn. Toch lijkt er iets van de *identiteitsthese* in Tye’s omgekeerde kwalia-verdediging te zitten: brein en geest zijn hetzelfde, en mentale en fysieke toestanden zijn dezelfde toestanden. Tye’s illustratie van twee wijnproevers suggereert immers een fysiek verschil in smaakorganen dat leidt tot een mentaal verschil in smaakervaring.

De identiteitsthese domineert echter niet Tye’s theorie, aangezien hij stelt dat zelfs als twee fysiek en functioneel identieke organismen bestaan, zij toch verschillende verschijnselen ervaren:

“All that we need to suppose is that the evolutionary histories and natural habitats of the two creatures are different, and that the brain states that realize sensations in the first creature are causally correlated with different external features from those the same brain states are causally correlated with in the second creature. In these circumstances, their sensations will have different representational contents and different phenomenal characters.”<sup>52</sup>

Bijgevolg kunnen functionele toestanden als dubbelzinnig worden opgevat; enerzijds kunnen sommige EGSn ze realiseren, dus zijn ze meervoudig realiseerbaar, anderzijds ervaren EGSn ze anders, dus zijn

---

<sup>52</sup> 205.

ze niet meervoudig realiseerbaar. Ik geloof echter dat functionele toestanden niet dubbelzinnig behoeven te worden geïnterpreteerd, ook al worden zij fenomenaal verschillend ervaren. Ik leg dit uit.

Het MT-functionalisme omschrijft dat abstracte toestanden kunnen worden gerealiseerd door verschillende EGSn. Het gestelde predicaat “externe EGS X ervaart mentale toestand Y” blijft echter conjecturaal. Want de enige ervaringswereld die ik met zekerheid vanuit een subjectief perspectief kan ervaren, mijn eigen wereld, is niet die van iemand of iets anders – althans, ik heb tot nu toe geen herinnering aan ervaringen opgedaan vanuit iemand anders dan mijn eigen lichaam.

Mijn strikte onwetendheid van andere werelden komt voort uit een *transcendentale beperking*: de manier waarop ik dingen kan weten, beperkt mijn vermogen om andere werelden ongemedieerd te kennen. Ik neem aan dat als alle andere mensen subjectieve ervaringswerelden hebben, zij ook te maken hebben met de transcendentale beperking. Ik vermoed dat dit waar is. Als zodanig schaadt iedere geleerde het paradigma waarmee die omgekeerde kwalia tracht te verenigen: hoe kan een geleerde ervaren hoe kwalia, indien aanwezig, door X wordt ervaren als de geleerde X slechts observeert vanuit een derde perspectief?

De vraag moet niet zijn of omgekeerde kwalia verenigbaar zijn met een filosofie van de geest, maar: “Welke filosofie van de geest houdt het beste rekening met de mogelijkheid van omgekeerde kwalia?” Mijn doel is om te laten zien dat het MT-functionalisme beter rekening houdt met omgekeerde kwalia dan hersentoestand-theoretici en behavioristen doen. We beginnen met de gedragswetenschap die behaviorisme heet.

Het behaviorisme kan mentale toestanden niet met succes verklaren of identificeren, zoals in hoofdstuk 1 is gebleken, en ziet ze zelfs als elementen van een ondoordringbare zwarte doos. Dit laatste kan een juist inzicht zijn vanwege de transcendentale beperking, maar de valse correlatie tussen gedrag en plausible mentale toestanden lijkt op behaviorisme dat geen rekening houdt met omgekeerde kwalia. We gaan verder met de identiteitstheorie als hersentoestand-theorie.

Ullin Place's (1924-2000) identiteitstheorie definieert hogere mentale toestanden als lagere fysieke toestanden, waarbij hij analytisch alle hogere mentale toestanden reduceert tot lagere fysieke toestanden. Uiteindelijk is de mentale toestand “jeuk aan je rug” dezelfde als de fysieke hersentoestand die optreedt wanneer zenuwvezels in de ruggengraat geactiveerd worden. Het enige verschil tussen de twee is de

*semantische* beschrijving, de betekenisvolle symboliek, terwijl ze *ontisch*, de feitelijke fysische werkelijkheid, hetzelfde zijn.<sup>53</sup>

De identiteitstheorie klinkt aantrekkelijk omdat organismen geen mentale toestanden kunnen realiseren in afwezigheid van de fysieke toestand van het organisme. En als dat zo is, wat volgens mij zo is, waarom zou het fysieke dan niet voldoende zijn voor de aard van het mentale? Misschien is het fysieke wel voldoende, maar integendeel voor de metafysische verklaring van mentale toestanden. Het MT-functionalisme is hier gebruiksvriendelijker wanneer ik de universele vormen van mentale toestanden tussen EGSn wil verklaren, aangezien de identiteitstheorie slechts over per empirisch bestudeerd organisme gaat. Ik zal nu nagaan hoe het MT-functionalisme rekening houdt met omgekeerde kwalia.

Het MT-functionalisme moet afzien van het verklaren van analogieën tussen de abstracte functionering van twee systemen door louter empirisch onderzoek van perifeer gedrag, omdat ik meen dat bezwaren tegen het behaviorisme dit ontkrachten.<sup>54</sup> Bovendien moet het afzien van gelijksoortige realisaties van abstracte toestanden door middel van gelijksoortige zintuiglijke invoeringen: objectieve zintuiglijke gegevens, krachtens de mogelijkheid van omgekeerde kwalia, zijn niet langer primaire objecten voor mentaal identieke toestanden; EGSn die ogenschijnlijk functioneel isomorf zijn, kunnen mentaal niet-identieke toestanden realiseren door omgekeerde kwalitatieve verschijnselen te ervaren tijdens dezelfde perceptie van een onafhankelijk object X.

Toch bevatten organismen hun diepste kern die de FO bepaalt, ongeacht hoe en wat organismen fenomenaal ervaren. Die diepste kern is het *genoom*; de verzameling van informatiemoleculen, oftewel het gehele DNA. In de wetenschap van de genetica is de discipline van de *functionele genetica*, waarvan het doel is te bepalen hoe de afzonderlijke componenten van een biologisch systeem samenwerken om een bepaald “fenotype” te produceren,<sup>55</sup> ons geconcentreerde discipline. Een *fenotype* is het geheel van waarneembare eigenschappen van een organisme. Deze eigenschappen kunnen ruim worden opgevat en omvatten ook functionele eigenschappen.

Via de functionele genetica en de evolutietheorie kunnen we komen tot een synthese tussen omgekeerde kwalia en het MT-functionalisme. De evolutietheorie leert ons dat organismen bepaalde doeloriëntaties hebben die vooral te maken hebben met het zo lang mogelijk overleven om zich te kunnen reproduceren.

---

<sup>53</sup> J. J. C. Smart, “The Mind/Brain identity Theory”, in *The Stanford Encyclopedia of Philosophy*, bew. Edward N. Zalta (Stanford: Metaphysics Research Lab, Spring 2017), URL: <https://plato.stanford.edu/archives/spr2017/entries/mind-identity/>.

<sup>54</sup> Zie pagina 15 van deze scriptie.

<sup>55</sup> EMBL’s European Bioinformatics Institute, “What is functional genomics?”, Functional genomics I, geraadpleegd op 1 juni, 2022, URL: <https://www.ebi.ac.uk/training/online/courses/functional-genomics-i-introduction-and-design/what-is-functional-genomics/>.

De functionele genetica brengt ons informatie over de totstandkoming van organismen, namelijk via de functionerende inter-structurele systematiek waaruit het organisme is opgebouwd. De evolutionaire doelgerichte richtlijnen van organismen en hun functioneren daartoe volgens de studie van de functionele genetica, zijn de kernelementen die tot uitdrukking komen voorafgaand aan de bepaaldheid van de fenomenale ervaringen. Ik stel voor dat die uitdrukking de manifestatie is van de creatie en lokalisatie van de FO die het organisme bezit.

Ik neem deze wetenschappelijke wending omdat deze ons in staat stelt elk ogenschijnlijk kwalitatief verschil tussen (ervaringen van) mentale toestanden te verklaren als een gevolg van historische ontwikkelingen van organische EGSn die zich per individu anders ontwikkelden dan alle andere. Als zodanig wordt rekening gehouden met de mogelijkheid van omgekeerde kwalia: elke ontwikkeling heeft zich anders voorgedaan en verloopt anders, en brengt andere (proxies voor) ervaringen met zich mee.

Hoewel de uniekheid van de individuele ontwikkeling noodzakelijkerwijs een unieke ervaring voor elk organisch EGS binnen dezelfde externe situatie met zich mee zou brengen, denk ik dat het een te verre bewering is om te zeggen dat organische EGSn op een dergelijke manier in functionele aard verschillen. Want evolutionaire en fysische wetten bepalen de mentale toestanden van organische EGSn zoals honger, pijn, vermoeidheid, geilheid, strijdlust, angst, enz. Dit gebeurt niet willekeurig, maar ruwweg door respectievelijk de behoefte aan voedsel voor energie, de wil om te overleven, de behoefte aan intrinsieke oplading en genezing, de wil om zich voort te planten en/of te genieten, de behoefte aan verdediging of expansie, de wil om te vluchten of het besef van extern kwaad. Hoewel deze toestanden door verschillende organische EGSn in verschillende situaties verschillend kunnen worden ervaren en/of gerealiseerd, houden de functionele genetica en de evolutietheorie vast aan de causale overeenkomst van het eigenlijke type van FOs en functionele toestanden tussen deze EGSn. Daarmee lijkt de waarschijnlijkheid van een soortgelijke ervaring van eenzelfde soort toestand, zij het in het verloop van verschillende externe situaties, eerder denkbaar dan domweg omgekeerd. Ik leg uit.

Organisme-1 dat een dergelijk omgekeerd spectrum ervaart, en dientengevolge minder goed functioneert met betrekking tot de relationeel zwakkere verwezenlijking van evolutionaire doelen, is kenmerkend voor een wezen dat een genetische ontwikkeling heeft ondergaan die toevallig ongunstiger is. De opeenvolgende realisaties van mentale toestanden van organisme-1 kunnen toestanden omvatten die in dezelfde situatie niet realiseerbaar zijn voor organismen zonder het omgekeerde spectrum, maar dit hoeft niet te betekenen dat organisme-1 niet meer dezelfde mentale toestanden kan realiseren als die welke het omgekeerde spectrum niet omvatten. Ik illustreer dit als volgt.



Stel dat Soumis plezier beleeft aan de gebeurtenis waarbij Dominatrix hem fysiek beschadigt, wat niet onmogelijk is omdat de psychologie kennis bevat van seksuele masochisme stoornis,<sup>56</sup> en Estime pijn beleeft aan dezelfde gebeurtenis. Het is waarschijnlijk dat Soumis en Estime verschillend zullen functioneren. Maakt dit de MRH overbodig? Naar mijn mening, nee. Want als we Soumis in een gebeurtenis plaatsen waarin hij pijn ervaart, en Estime niet, dan betekent dit dat zij dezelfde mentale toestand kunnen realiseren – ook al is dat in tegengestelde situaties. Dit komt door de mogelijkheid van omgekeerde kwalia: de realisatie van tenminste twee mentaal identieke toestanden ligt niet meer voor de hand in de gewaarwordingen van dezelfde omstandigheden. Desondanks, welke kwalia Soumis en Estime ook omgekeerd ervaren, de kern van genetisch en evolutionair determinisme bepaalt het type toestand – die, net als elke andere identieke toestand, een vergelijkbare functionele waarde heeft in relatie tot Soumis en Estime's organische systemen.

Tenslotte moeten we bij machines niet louter kijken naar programma's of gedragsbewegingen om functionele toestanden gelijk te stellen met de onze, maar naar de basis voor het invoeren en draaien van het systeem; zoals BIOS van een PC met wiskundige fysica erbij. Want die technisch-wiskundige basis is vergelijkbaar met de evolutionair-genetische basis: beide de diep bepalende kern van de FOs van machines en organismen. MT-functionalisme houdt dan ook rekening met omgekeerde kwalia.

---

<sup>56</sup> George R. Brown, "Sexual Masochism Disorder", East Tennessee State University, Laatst gewijzigd april, 2021, URL: <https://www.msmanuals.com/professional/psychiatric-disorders/paraphilic-disorders/sexual-masochism-disorder>.

## Conclusie.

In deze scriptie heb ik een filosofische verdediging uitgewerkt voor de verenigbaarheid van kwalia en het MT-functionalisme van Hilary Putnam. Mijn methode was om eerst de relevante delen van de technische concepten te beschrijven die inherent zijn aan de probleemstelling. Daarna begon ik met het beschrijven van het afwezige kwalia-argument. Dat was het argument van Ned Block dat ik omschreef aan de hand van de literatuur van David Chalmers. Ik belichtte de consequenties van dat argument tegen het MT-functionalisme. Vervolgens heb ik die consequenties afgezwakt met een tegenargument van William Lycan. Ik bekritiseerde dat tegenargument en maakte vervolgens een positieve aanvulling aan de hand van het ACT van Daniel Friedman & Eirik Sørvik. Daarna doorliep ik dezelfde drie stappen voor het omgekeerde kwalia-argument, maar dan met andere auteurs. Ik beschreef het omgekeerde kwalia-argument van John Locke aan de hand van de verzamelde literatuur van Peter H. Nidditch. Daarna beschreef ik een tegenargument van Michael Tye. Ik beschreef de impact van dat tegenargument. Tenslotte bekritiseerde ik dat tegenargument en vulde het op een positieve manier aan. Met deze scriptie hoop ik het MT-functionalisme te hebben begeleid in de richting van verenigbaarheid met kwalia.

Overwegingen voor verder onderzoek kunnen zijn: (1) De functionele gevolgen binnen gedrag en geest na of tijdens de causaliteit van de diepewortelde basis van genen en evolutionaire wetten in organismen enerzijds, en het basis input/output systeem en de fysica in machines anderzijds. Een succesvolle beschrijving van beide kan dan streven naar een correcte analogie tussen geïmpliceerde functionele toestanden. Vanuit deze diepere kern kan omgekeerde kwalia als mogelijk worden begrepen, en niet in tegenspraak zijn met MT-functionalisme; (2) De bewustzijnsrelatie tussen inter-structureel-hiërarchische organen en subjectief bewustzijn. Als deze relatie grond krijgt, krijgen de argumenten die een complex FO lokaliseren en dus zodanig een subjectief bewustzijn impliceren ook grond om die claim te maken. Daarmee verliest het argument van de afwezige kwalia steeds meer het vermogen om EGSn niet-bewust te karakteriseren.

## Bibliografie.

Block, Ned Joel. "Troubles with Functionalism." *Philosophy of Science* 9, (1978): 261-325. Minnesota Studies.

Brown, George R. "Sexual Masochism Disorder." East Tennessee State University. Laatst gewijzigd April, 2021. <https://www.msmanuals.com/professional/psychiatric-disorders/paraphilic-disorders/sexual-masochism-disorder>.

Chalmers, David J. "Absent Qualia, Fading Qualia, Dancing Qualia." In *Conscious Experience*, bewerkt door Thomas Metzinger, 309-330. Kansas: Allen Press, 1995.

EMBL's European Bioinformatics Institute. "What is functional genomics?" Functional genomics I. Geraadpleegd op 1 juni, 2022. <https://www.ebi.ac.uk/training/online/courses/functional-genomics-i-introduction-and-design/what-is-functional-genomics/>

Friedman, Daniel & Søvik, Eirik. "The ant colony as a test for scientific theories of consciousness." *Synthese* 198, nr. 5 (7 februari 2019): 1-24. DOI:10.1007/s11229-019-02130-y.

Herculano-Houzel, Suzana. "The human brain in numbers: a linearly scaled-up primate brain." *Frontiers in Human Neuroscience* 3, nr. 31 (9 november 2009): 1-11. DOI:10.3389/neuro.09.031.2009.

Kant, Immanuel. *Critique of Pure Reason*. Vertaald door Marcus Weigelt. London: Penguin Classics, 2003.

Levin, Janet. "Functionalism." In *The Stanford Encyclopedia of Philosophy*, bewerkt door Edward N. Zalta. Stanford: Metaphysics Research Lab, Winter 2021. <https://plato.stanford.edu/archives/win2021/entries/functionalism/>.

Locke, John. *An Essay Concerning Human Understanding: Clarendon Edition of the Works of John Locke*, bewerkt door Peter H. Nidditch. Oxford: University Press, 1975.

Lycan, William G. *Consciousness*. Massachusetts: Institute of Technology, 1995.

NBC News. "Queen Pardons Computing Giant Alan Turing 59 Years After His Suicide." World. Laatst gewijzigd 1 april, 2014. <https://www.nbcnews.com/news/world/queen-pardons-computing-giant-alan-turing-59-years-after-his-n241>.

Putnam, Hilary Whitehall. "Minds and Machines." In *Dimensions Of Mind: A Symposium*, bewerkt door Sidney Hook, 148-179. New York: University Press, 1960.

Putnam, Hilary Whitehall. "The Nature of Mental States." In *Readings in Philosophy of Psychology, Volume I*, bewerkt door Ned Block, 223-231. Cambridge: Harvard University Press, 1980.

Putnam, Hilary Whitehall. *Mind, Language and Reality: Philosophical Papers, Volume 2*. Cambridge: University Press, 1975.

Putnam, Hilary. "Brains and behavior." In *Analytical Philosophy: Second series*, bewerkt door Ronald J. Butler, eerste hoofdstuk. Oxford: Basil Blackwell, 1965.

Smart, J. J. C. "The Mind/Brain identity Theory." In *The Stanford Encyclopedia of Philosophy*, bewerkt door Edward N. Zalta. Stanford: Metaphysics Research Lab, Spring 2017. <https://plato.stanford.edu/archives/spr2017/entries/mind-identity/>.

Tye, Michael. *Ten Problems of Consciousness: A Representational Theory of the Phenomenal Mind*. Second printing. Massachusetts: Institute of Technology, 1996.

Watson, James C. "Overview of Pain." Mayo Clinic College of Medicine and Science. Laatst gewijzigd april, 2020. <https://www.merckmanuals.com/home/brain,-spinal-cord,-and-nerve-disorders/pain/overview-of-pain?redirectid=553?rulerredirectid=30>.