

## 10.3. Hersenen en bewustzijn. Neurale en mentale gebeurtenissen

Paul A.M. van Dongen © 2021

*Of en hoe hersenprocessen mentale processen veroorzaken, wordt het Harde Probleem genoemd. Het Harde Probleem is onoplosbaar, omdat er geen natuurwetten zijn over mentale gebeurtenissen.*

### Samenvatting

Mensen voeren allerlei complexe handelingen uit, en in de regel kunnen ze zich bewust zijn van wat ze doen en waarom. Maar bij patiënten met hersenletsel of andere aandoeningen is er een bizar verband tussen wat zij doen en wat ze zich bewust zijn. De reden die mensen voor hun gedrag vermelden, is niet altijd de werkelijke oorzaak – ook niet als ze oprecht zijn. Als we spreken over hersenen en bewustzijn, moeten we een onderscheid maken tussen gebeurtenissen in de fysieke wereld (oorzaak en gevolg) en die in de mentale wereld (redenen en intenties). In een heldere bespreking moeten ‘oorzaak’ en ‘reden’ strikt worden onderscheiden. Experimenteel hersenonderzoek toont dat de hersenen het initiatief tot actie al genomen hebben, enkele seconden voordat de actor zich dat bewust werd.

Er zijn verscheidene standpunten in de discussie over het verband tussen hersenen en geest, of tussen neurale en mentale gebeurtenissen. Deze standpunten worden hier kort besproken. De sterke en zwakke punten van de afzonderlijke visies komen aan de orde. Er zijn argumenten dat er geen natuurwetenschappelijke verklaring kan zijn voor het verband tussen neurale en mentale gebeurtenissen.

#### Samenvatting

1. Inleiding
2. Hersenen en bewustzijn
  - 2.1. Bewustzijn na hersenletsels
  - 2.2. Temporeel verloop tussen hersenprocessen en bewustzijn
  - 2.3. Vrije wil en verantwoordelijkheid
3. Filosofie: de verschillende visies
  - 3.1. Dualistische visies
  - 3.2. Materialistische visies
  - 3.3. Functionalisme
  - 3.4. Idealisme: eliminatie van fysieke gebeurtenissen
  - 3.5. Overzicht van de visies
4. Bewustzijn bij dieren?
5. Besluit

### 1. Inleiding

#### Mentale en neurale gebeurtenissen

U leest dit stuk van papier of van een beeldscherm. U ziet ●. Wat gebeurt er dan? Vanaf ● komen rode lichtquanten. U hield uw oog op ● gericht, zodat de rode lichtquanten op uw gele vlek vielen, en daar door kegeltjes geabsorbeerd werden. Dan komt het verhaal van cellen in de retina, de oogzenuw en de hersenen. En dan ziet u ●. Alles is recht-toe-recht-aan exacte wetenschap met causale relaties tussen gebeurtenissen, totdat er iets nieuws gebeurt: u ziet ●. Deze waarneming is een mentale gebeurtenis, het is een bewustzijnstoestand. Hetzelfde geldt voor geluid, geur, pijn enz. en ook voor herinneringen, gevoelens en intenties.

Wat zijn de verschillen tussen mentale en fysieke (en neurale) gebeurtenissen?

1. De neurale gebeurtenissen vinden ergens plaats: daar waar de neuronen zijn. Maar waar vinden de gewaarwordingen plaats? Waar zit uw sensatie rood? Als u de Dom van Utrecht ziet, hoe groot is die waarneming dan? Het is geen zinnige vraag om naar de plaats van mentale gebeurtenissen te vragen. De categorie plaats is niet van toepassing op mentale gebeurtenissen (verg. Hospers 1967, p. 379-381).
2. Iedereen kan in principe fysieke objecten en gebeurtenissen waarnemen, ze zijn publiek waarneembaar. En ook de activiteit van zenuwen en het zenuwstelsel kan men met speciale apparaten waarnemen. Maar mijn ervaring van ● is alleen voor mij toegankelijk. Als u naar ● kijkt, heeft u uw

ervaring van ● maar niet de mijne. U kunt in principe neurale gebeurtenissen in mijn brein waarnemen (Mitchell e.a. 2008, Just e.a. 2010, Just en Varma 2007, Binder e.a. 2016, Huth e.a. 2016). Maar dan heeft u uw beleving van mijn neurale gebeurtenissen, en niet mijn ervaring van ●. Dit is het probleem het 'heteropsychologische' (Carnap 1928, § 140), of van 'other minds' (Avramides 2001). Men kan de belevingen, overtuigingen en gedachten van een ander te weten komen door er naar te vragen. Maar dan neemt men de woorden, en niet de belevingen van de ander waar. Dat levert niet altijd de juiste informatie: mensen kunnen zich vergissen of liegen. Neurologen en psychiaters weten dat bepaalde patiënten andere redenen voor hun gedrag noemen dan de werkelijke oorzaak die bij de neuroloog of psychiater bekend is.

3. Neurale gebeurtenissen zijn gebeurtenissen van neuronen, maar mijn mentale gebeurtenissen zijn ervaringen van mij. (Alleen ... vraag me niet wat of wie die "ik" dan is, want dat tart iedere beschrijving. <sup>1</sup>)

Dit omschrijft de betekenis van mentale en neurale gebeurtenissen. Deze begrippen kunnen niet in simpelere termen verwoord (gedefinieerd) worden. Het zijn basisbegrippen zoals 'denken', 'weten', 'willen', 'voelen', 'zien', 'horen' ('*semantic primes*', Goddard en Wierzbicka 2007). Deze begrippen kunnen alleen omschreven worden zoals hier boven gebeurd is: door gebeurtenissen te benoemen die - zoals we gemakshalve aannemen - vertrouwd zijn voor de lezer.

### Subjectief en intersubjectief

Als men enige tijd naar een fel-gekleurd vlakje staart, en daarna naar een egaal grijs vlak, dan ziet men op het grijs veld een vlakje in de 'complementaire kleur'. Na het kijken naar een rood vlakje, ziet men een groen na-beeld, en na het kijken naar een blauw vlakje, ziet men een geel na-beeld. Als een persoon een na-beeld waarneemt, is dat een subjectieve, private mentale gebeurtenis. Maar men kan na-beelden voorspelbaar bij mensen opwekken. Dan spreekt men van 'intersubjectieve waarneming'. Intersubjectieve waarnemingen zijn de basis voor iedere empirische wetenschap.

Er zijn allerlei subjectieve mentale gebeurtenissen, zoals het aflezen van een meetinstrument, het waarnemen van een kleur, het waarnemen van een na-beeld, en het waarnemen van een optische illusie. Het waarnemen van een na-beeld, en het waarnemen van een optische illusie zijn intersubjectieve waarnemingen, die onderwerp zijn voor empirisch onderzoek. De na-beelden en de illusies zijn niet het gevolg van actuele lichtstimuli, maar ze worden bij (nagenoeg) alle mensen gemaakt door het zenuwstelsel. Er zijn ook andere, meer individuele mentale gebeurtenissen, zoals het waarnemen van een hallucinatie, of het waarnemen van visioenen. Het waarnemen van hallucinaties of visioenen, zijn subjectieve waarnemingen, die alleen privé toegankelijk zijn, en niet geschikt zijn voor empirisch wetenschappelijk onderzoek. (Hooguit voor empirisch onderzoek naar geneesmiddelen die hallucinaties verminderen.)

### Een 'psyche'?

Neurale toestanden zijn toestanden van neuronen. Maar tot wie of wat behoren mentale toestanden dan? Het voor de hand liggende antwoord is: van de 'geest', de 'psyche', de 'ziel', de 'mind'. Maar ... wat is die psyche dan? Als ik droom of geniet, is het niet mijn lichaam dat droomt of geniet, maar 'ik ben het zelf'. De psyche is dan 'mijn wezenlijke zelf'. Maar wat is mijn zelf dan? "*Wanneer ik zeer intiem binnentreed in wat ik mezelf noem, dan struikel ik altijd over een of andere afzonderlijke perceptie van warm of koud, licht of schaduw, liefde of haat, pijn of plezier. Ik kan nooit mezelf op enig moment vangen zonder perceptie en nooit iets waarnemen behalve de perceptie.*" (Hume 1739, book1, part 4, section 6). Mijn gedachten zijn van mij. Maar, opnieuw, wat wordt bedoeld met 'ik/mij'? 'Ik' is zo'n basaal begrip, dat het niet in meer elementaire woorden verwoord (gedefinieerd) kan worden <sup>2</sup> (ook dit is een '*semantic prime*', Goddard en Wierzbicka 2007). In het vervolg spreek ik liever niet over 'geest', 'psyche' of 'ziel', omdat deze zelfstandige naamwoorden suggereren dat er een of ander ding is. Ik spreek hier verder van mentale gebeurtenissen die eerder aan de hand van voorbeelden omschreven zijn.

<sup>1</sup> Kant (1781) sprak over „*Ich denke*“ en „*die transzendente Einheit des Selbstbewußtseins*“, en Vendler (1984) over „*the transcendental Self*“.

<sup>2</sup> Dit wordt ook wel de *left brain interpreter* genoemd (Phelps en Gazzaniga 1992).

## 2. Hersenen en bewustzijn

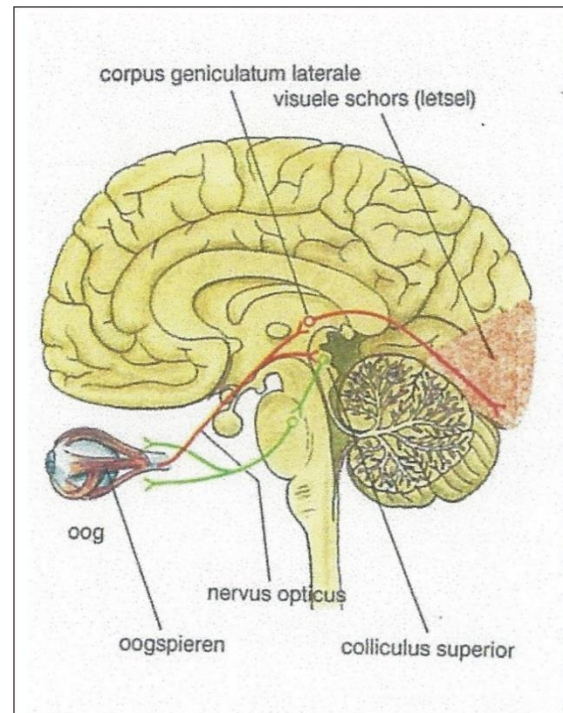
### Emergente eigenschappen?

Mentale gebeurtenissen worden wel beschouwd als emergente eigenschap van een brein dat complex genoeg is (Jerison 1973, Bunge 1977). In hoofdstuk 10.2. zijn 'emergente eigenschappen' verder toegelicht. Het onderscheid in niveaus is hier belangrijk: men kan niveaus zodanig kiezen dat processen op één niveau causaal verklaard worden door processen op een lager niveau (van Dongen en van den Bercken 1981). Processen op een bepaald niveau kunnen causaal worden verklaard door processen op één niveau lager. Een eigenschap op een hoger niveau is een emergente eigenschap, die ontstaat door de eigenschappen één niveau lager. Zijn mentale gebeurtenissen nu ook op dezelfde manier simpel een emergente eigenschap? Sommige denkers vinden van wel (Bunge 1977). Maar er is minstens één duidelijk verschil. In principe vormen gebeurtenissen op één niveau een sluitende causale verklaring voor gebeurtenissen op één niveau hoger. Maar dat geldt niet voor mentale processen. Het is onduidelijk hoe mentale gebeurtenissen tot stand komen.

### 2.1. Empirisch onderzoek naar belevingen

#### Adequaat, complex gedrag zonder belevingen

In speciale situaties kunnen mensen adequaat, complex gedrag vertonen, terwijl alles erop wijst dat zij zich niet daarvan bewust zijn. Een patiënt had een letsel opgelopen aan zijn visuele schors: hij kon niets meer zien, hij was blind. Maar subcorticale delen van het visuele systeem waren nog intact (figuur 1). Een onderzoeker plaatste deze patiënt voor een scherm met daarop een lichtpuntje dat op en neer bewoog. De onderzoeker gaf de patiënt de opdracht met zijn ogen dit puntje te volgen. De patiënt protesteerde: hij zag niets, dat kon hij niet, wat was dat voor flauwekul. Maar de onderzoeker bleef aandringen. En jawel, de patiënt volgde met zijn ogen zo'n lichtpuntje, hoewel hij oprecht zei niets te zien (Pöppel e.a. 1973). Dit verschijnsel wordt '*blindsight*' genoemd. *Blindsight* is door verscheidene onderzoekers bevestigd (Covey en Stoerig 1991). Vermoedelijk komt dit tot stand via intacte verbindingen met een hersendeel, de colliculus superior, die indirect de spieren voor oogbewegingen bestuurt, terwijl het lichtpuntje geen toegang had tot de taalcortex, en niet gepaard ging met een bewuste ervaring.



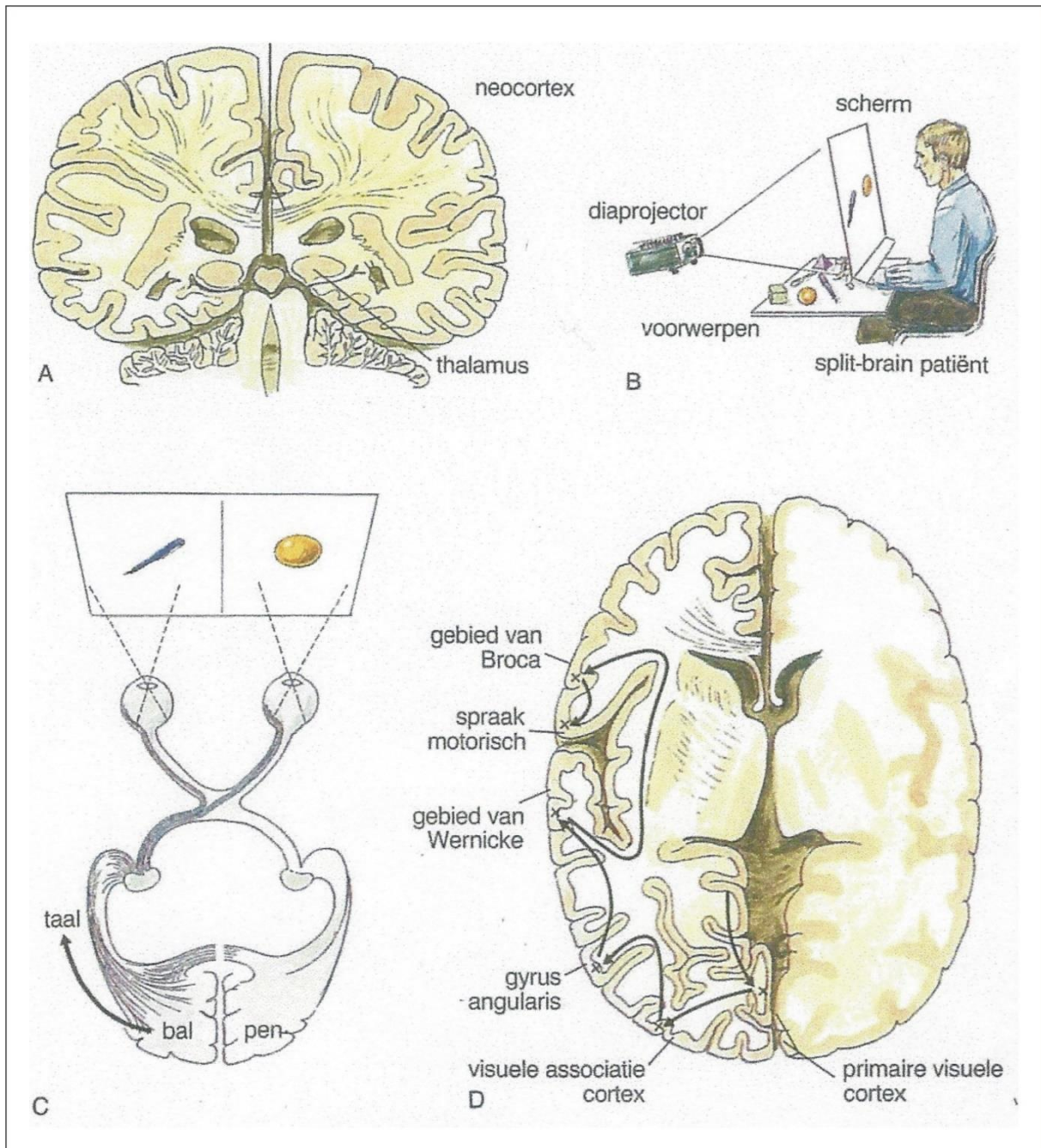
Figuur 1. Schematisch overzicht van de beschadigde en intacte delen van het visuele systeem van een patiënt met blindsight. Hier is getekend hoe een bewegend lichtpuntje nog correcte oogbewegingen kon veroorzaken, terwijl de patiënt zei het puntje niet te zien.

#### Split-brain patiënten

Bij epileptische patiënten wordt soms de grote verbindingssbaan tussen beide cortexhelften doorgesneden, het corpus callosum, van 200 miljoen zenuwvezels (figuur 2). Dat doen neurochirurgen om uitbreiding van de epileptische aanvallen te voorkomen. Dit noemt men '*split-brain*' patiënten, omdat beide cortexhelften gescheiden functioneren. Sperry (1970) zette een *split-brain* patiënt voor een scherm waarop hij in een korte flits twee voorwerpen projecteerde: bijv. links een pen en rechts een bal.

- De representatie van de bal in het rechter gezichtsveld kwam via de linker visuele cortex naar taalcentra in de linker hersenhelft (figuur 2). De patiënt kan de bal benoemen, en dit lijkt op een alledaagse bewuste ervaring.
- De representatie van de pen in het linker gezichtsveld kwam via de rechter visuele cortex, maar bleef in de rechterschors en kon taalcentra niet bereiken: de patiënt kon de pen niet benoemen. Maar een *split-brain* patiënt kan met zijn linkerhand, die door de rechterschors bestuurd wordt, de pen wel pakken zonder visuele controle. Als men de

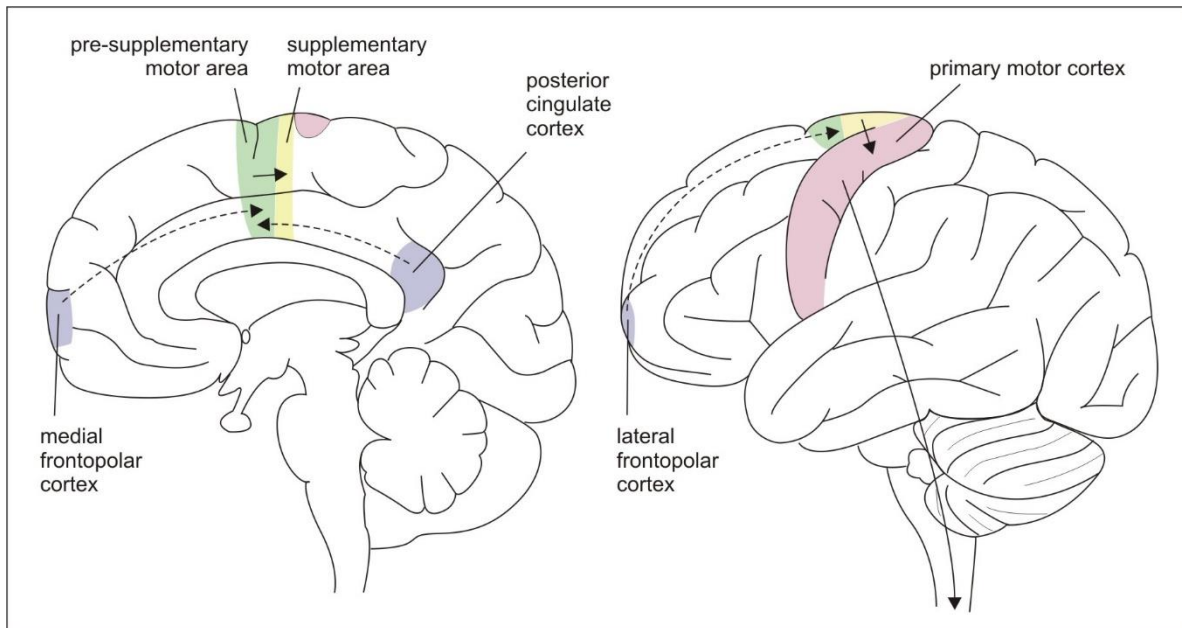




Figuur 2. Schema van het onderzoek bij split-brain patiënten. A. Overzicht van de hersenverbinding die doorgesneden is. B. Overzicht van de opstelling om de vaardigheden van split-brain patiënten te testen. C. Waar komen welke beelden in de hersenen terecht? D. Hoe neurale codes voor de bal spraakcentra kunnen bereiken, en de neurale codes voor de pen de bewegingen van de linkerhand kunnen beïnvloeden.

patiënt vraagt waarom hij een pen gepakt heeft, komt hij met allerlei smoesjes, omdat hij de oorzaak van zijn gedrag niet kan weten, en niet kan verwoorden. Dat toont de werking van de *left-brain interpreter* (Gazzaniga 1992), een module in de linker cortex die probeert een samenhangend ver-

haal te maken van de beschikbare informatie. Deze module wordt ook wel denigrerend 'kwebbeldoos' genoemd (Lamme 2010). De patiënt kan wel adequaat, complex gedrag vertonen, maar dat gaat niet gepaard met bewuste gewaarwording of een correcte taaluiting.



Figuur 3. Enkele cortexdelen die een rol spelen bij de voorbereiding van 'vrijwillige' bewegingen (zie Soon e.a. 2008, Haggard 2008, Fried e.a. 2011). De continue pijlen geven directe zenuwverbindingen weer, en de gestippelde pijlen indirecte verbindingen (zie ook figuur 4).

## 2.2. Temporeel verloop tussen hersenprocessen en bewustzijn

### 2.2.1. Hersenactiviteit en handbewegingen

#### De *Bereitschaftspotential*

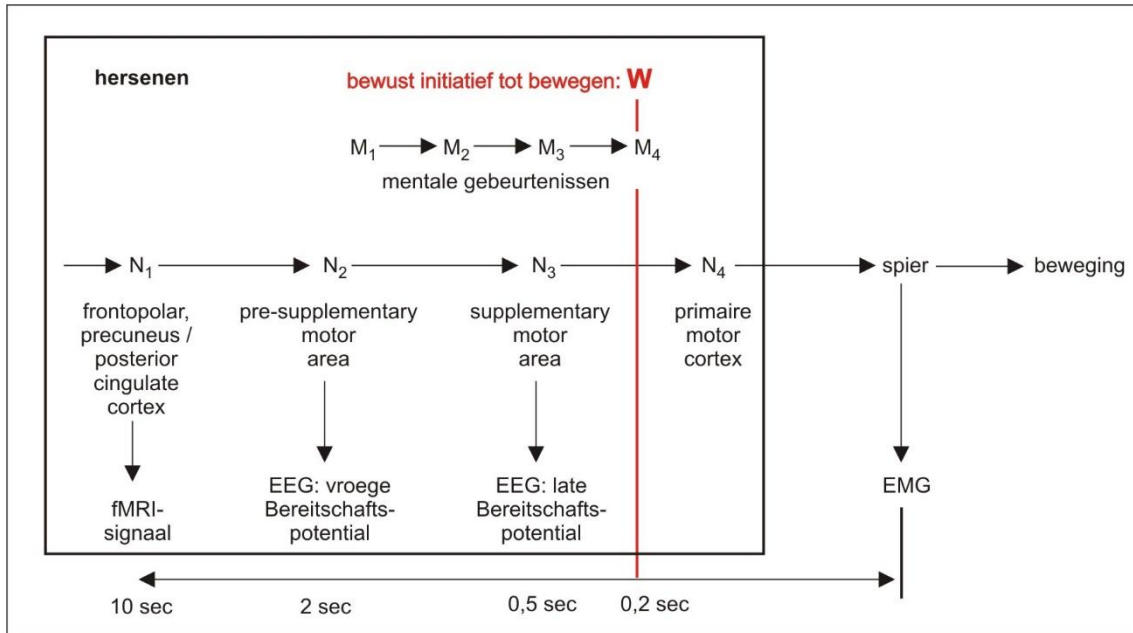
Voor een onderzoek naar het verband tussen hersenprocessen en bewegingen zaten proefpersonen in een rustige kamer, en zij hadden de opdracht om in afzonderlijke sessies op een zelfgekozen moment plots een vinger te bewegen. Van deze proefpersonen werd de elektrische activiteit van de hersenen (elektro-encefalogram, EEG) en spieren (elektro-myogram, EMG) afgeleid. Deze signalen werden in een computer gevoerd. De hersenactiviteit vóór de beweging werd gemiddeld, waarbij het EMG-sig-naal gebruikt werd om het precieze begin van de beweging vast te stellen. In het EEG was een trage negatieve golf te zien, 1 - 2 sec voordat de proefpersoon een beweging maakte (Kornhuber en Deecke 1965). Dit werd de *Bereitschaftspotential* (BP) genoemd. De BP is het grootst in de contralaterale precentrale cortex; daar is de amplitude 10 - 15  $\mu$ V. Dus 1 - 2 seconde voor een zelfgekozen beweging is er al systematische hersenactiviteit. Er was al hersenactiviteit voor de beweging, en dat was niet raar, want hersenactiviteit veroorzaakt de beweging. Verscheidene onderzoekers hebben dit gerepliceerd (reviews Shibasaki en Hallett 2006, Haggard 2008). Nu onderscheidt men een vroege en een late BP: de

vroege BP begint 2 sec voor de beweging, en de late ongeveer 0,5 sec voor de beweging (Shibasaki en Hallett 2006).

Hersenactiviteit kan ook met fMRI (functionele *magnetic resonance imaging*) bestudeerd worden: dat toont welke hersendelen in welke situaties extra actief zijn. Met fMRI hebben Soon e.a. (2008) hersenactiviteit gevonden veel vroeger dan de BP. Zelfs tot 10 seconden voor de beweging was er al activiteit in de fronto-polare cortex en de posterior cingulate cortex gekoppeld aan de beweging (zie figuur 3).

#### De *Bereitschaftspotential* en intenties

Men heeft niet alleen het verband tussen hersenprocessen en bewegingen bestudeerd, maar ook tussen hersenprocessen en belevingen. Libet e.a. (1983) plaatste zijn proefpersonen in dezelfde opstelling als in de vorige alinea beschreven. Maar zij hadden ook een wijzerplaat met een 'secondewijzer', die in 2,56 seconden rond ging. Toen konden de proefpersonen achteraf zeggen wanneer zij besloten hadden te bewegen met een nauwkeurigheid van 40 ms. De proefpersonen vertoonden een BP op 1 - 2 sec voor de beweging, maar ze meldden dat ze ongeveer 0,2 sec voor de beweging besloten hadden te reageren (figuur 4, Libet e.a. 1983, Libet 1985). De hersenen hadden dus al de beslissing genomen te bewegen, voordat de proefpersonen dit bewust waren. De bewuste beleving van de intentie kwam dus



Figuur 4. De Bereitschaftspotential. Schema voor de achtergronden van 'vrijwillige' bewegingen. Verscheidene seconden voor de spier-activiteit is er al bewegings-gerelateerde activiteit in allerlei hersengebieden. Die activiteit is er seconden voordat de proefpersoon zich bewust is van het initiatief tot bewegen (Kornhuber en Deecke 1965, Libet e.a. 1983, Soon e.a. 2008).

na het begin van de hersenactiviteit tot bewegen. Dus de intentie (als beleving) kan deze hersenactiviteit niet veroorzaakt hebben. Dit is door veel andere onderzoekers bevestigd. Bij bewegingen op een 'zelfgekozen' moment hebben de hersenen al eerder besloten tot beweging, voordat de persoon zich dat bewust is. Wellicht leiden deze bevindingen ertoe dat het concept 'vrije wil' herzien of afgeschaft moet worden (Roskies 2010).

Libet e.a. (1983) vroegen hun proefpersonen te doen alsof ze een bewegingen wilden maken op een zelfgekozen tijdstip, en die beweging af te breken op een afgesproken tijdstip. "Ondanks dit vreemde verzoek," leek het toch alsof de proefpersonen dit konden, en werden er soms andere EEG-signalen gevonden (Libet e.a. 1983, p. 369). Deze observaties zijn aanleiding voor veel discussie (Libet 1985, Walsh e.a. 2010).

#### De Bereitschaftspotential en de hersenen

Hier noem ik de hersendelen die een rol spelen bij bewegingen op een 'zelfgekozen' moment in omgekeerde temporele volgorde (figuur 3).

- De primaire motor-cortex bevat cellen met uitlopers (axonen) direct naar motoneuronen in het ruggenmerg.
- De *supplementary motor area* ligt vlak voor de motor-cortex. Dit gebied genereert de late BP in het EEG, ongeveer 0,5 sec voor de beweging (Shibasaki en Hallett 2006).
- De *pre-supplementary motor area* ligt vlak voor de *supplementary motor area*. Dit gebied genereert de vroege BP in het EEG, ongeveer 2 - 0,5 sec voor de beweging (Shibasaki en Hallett 2006).
- De *anterior cingulate cortex*. Bij patiënten met epilepsie is met diepte-electrodes de activiteit van afzonderlijke neuronen afgeleid. Sommige neuronen vuurden vaker, en andere juist minder vaak voordat de patiënten de intentie tot bewegen rapporteerden (Fried e.a. 2011).
- De *frontopolar cortex* en de *precuneus / posterior cingulate cortex*. Hier is met fMRI extra activiteit gevonden tot 10 sec voordat de proefpersonen de intentie tot bewegen rapporteerden (Soon e.a. 2008).

In het meeste onderzoek is de tijd van de eerste bewuste aandrag 0,2 sec voor het EMG-signaal. Maar als je proefpersonen tijdens het experiment vraagt of ze dachten te gaan bewegen op het moment dat een toon klonk, was het gemiddelde interval wel 1,4 sec (Matsushashi en Hallett 2008), maar met een grote spreiding (de standaard deviatie was 50% van het gemiddelde). (1) Hetzij kunnen proefpersonen het tijdstip van bewuste gewaarwording niet betrouwbaar beoordelen, of (2) het tijdstip van bewuste gewaarwording is zeer variabel.

### 2.2.2. Oogbewegingen en 'inzicht'

Het registreren van oogbewegingen toont gerichte activiteit van de hersenen, zonder dat de persoon zich daarvan bewust hoeft te zijn. Voor een experiment zaten proefpersonen voor een computerscherm waarop 5 letters in een cirkel rond een fixatiepunt verschenen. De opdracht was om van 4 letters een woord te maken, terwijl de 5<sup>e</sup> letter een afleiding was. Als ze het woord gevonden hadden, moesten ze dat zo snel mogelijk zeggen. Tijdens dit experiment werden hun oogbewegingen gemeten, zodat de onderzoekers konden weten naar welke letters de proefpersonen gekeken hadden. De proefpersonen moesten ook zeggen of ze de oplossing opeens zagen (dat noemde men 'inzicht, hoofdstuk 3.5.'), of dat ze letter-voor-letter de oplossing hadden opgebouwd. Uit de oogbewegingen kon men concluderen wanneer de hersenen de oplossing gevonden hadden. In dit experiment hadden de hersenen de oplossing gevonden gemiddeld 2,3 – 2,4 seconden, voordat de proefpersoon de oplossing rapporteerde (Ellis e.a. 2011). Ook hier was de hersenactiviteit sneller dan de gerapporteerde bewuste ervaring.

## 2.3. Vrije wil en verantwoordelijkheid

### Vrije wil

In mentale termen: mensen hebben het gevoel dat zij zelf kunnen kiezen wat te doen of te laten (Wegner en Wheatley 1999). De vraag of mensen of dieren een vrije wil hebben, kan echter niet empirisch beantwoord worden. Ik denk zelfs dat het begrip 'vrije wil' niet helder omschreven kan worden. Sommige auteurs verwerpen categorisch de opvatting dat mensen een vrije wil hebben (Dennett 1991, Wegner 2002, Swaab 2010, Lamme 2010), hoewel

ze niet helder kunnen omschrijven wat ze verwerpen<sup>3</sup>. We kunnen hooguit zeggen dat andere mensen en dieren zich gedragen alsof de impuls tot gedrag van binnenuit komt (zie ook hoofdstuk 10.2.). Maar ook robots kunnen zo geprogrammeerd zijn dat ze zich gedragen alsof de impuls tot gedrag van binnenuit komt. In het algemeen geldt: als een apparaat zich onvoorspelbaar gedraagt, dan lijkt dat apparaat over vrije wil te beschikken – ongeacht of dit onvoorspelbare gedrag door een speciaal programma of door incidentele storing veroorzaakt is (Ebert en Wegner 2011)<sup>4</sup>.

Hier lopen verschillende discussies door elkaar:

- Is de werking van zenuwcellen en daarmee de werking van de hersenen gedetermineerd, d.i. volledig causaal bepaald (hoofdstuk 10.2.)? In het zenuwstelsel is er in ieder geval onbegrepen activiteit van zenuwcellen. Bijvoorbeeld zenuwcellen in het visueel en auditief systeem vuren actiepotentialen af in volledige duisternis of volledige stilte.
- Is er een onstoffelijk iets (een geest of een zelf, wat dat ook moge betekenen) die zenuwcellen kan beïnvloeden en die zelf in ieder geval niet volledig door zenuwcellen aangestuurd wordt?

De kwantummechanica kan het probleem van de vrije wil niet oplossen. Volgens de Kopenhagen-interpretatie van de kwantummechanica is de natuur op micro-schaal wezenlijk onbepaald of chaotisch. De 'vrije wil' daarentegen wordt geacht in staat te zijn tot doelgerichte acties.

### Verantwoordelijkheid

Discussies over de vrije wil gaan ook vaak over de consequenties van het ontbreken van een vrije wil: hoe kan een ouder, een rechter of een maatschappij mensen verantwoordelijk stellen voor hun daden, als hun daden volledig causaal bepaald zijn<sup>5</sup>? Vanuit een pragmatische visie is het niet relevant of een mens al dan niet vrij is. Voor de samenleving is het belangrijk dat mensen zich netjes gedragen, en dat doen ze vooral als ze verantwoordelijk gesteld worden voor hun daden (= beloond of gestraft worden). Dus voor de maatschappelijke orde moeten mensen verantwoordelijk gesteld worden, ongeacht ethische/filosofische bedenkingen. Het neurale equivalent van de kans dat men verantwoordelijk gesteld kan worden,

<sup>3</sup> Men kan hooguit zeggen of men wel of niet gelooft in het causaal determinisme, maar causaal determinisme is alleen van toepassing in formele wetenschappen (hoofdstuk 10.1.).

<sup>4</sup> Sommige mensen worden kwaad op dergelijke 'onwillige' apparaten.

<sup>5</sup> Het is bizar om het oordeel over een uitspraak te laten hangen van de consequenties van die uitspraak.



maakt dan deel uit van het (onvrije?) hersensysteem dat gedrag aanstuurt. In mentale termen, is de anticipatie van straf dan onderdeel van de 'vrije wil'. Daardoor kan maatschappelijk ongewenst of misdadig gedrag minder waarschijnlijk worden.

### 3. Filosofie: de verschillende visies

Ik geef hier een kort overzicht van de verschillende filosofische standpunten over het verband tussen neurale en mentale gebeurtenissen.

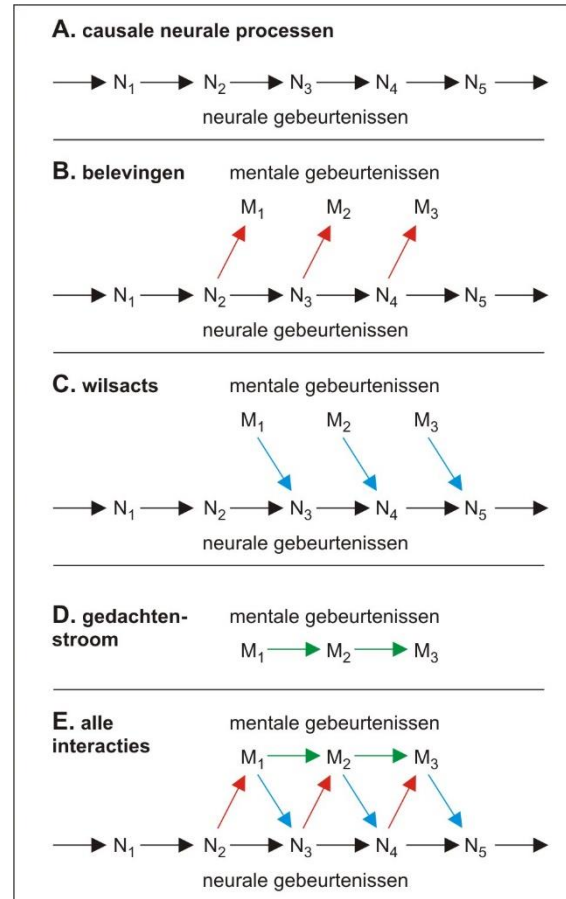
#### 3.1. Dualistische visies

##### 3.1.1. Interactionisme

Een stimulus, zoals ●, veroorzaakt fysische gebeurtenissen (absorptie en reflectie van lichtquanten), die neurale gebeurtenissen in het oog en de hersenen veroorzaken, die op hun beurt mentale gebeurtenissen veroorzaken: u ziet ●. Zo gaat dat - denkt men. En ook mentale gebeurtenissen veroorzaken fysische gebeurtenissen. Bijvoorbeeld: toen ik dit stukje tekst schreef, wilde ik het interactionisme uitleggen. Dit plan veroorzaakte activiteit in mijn hersenen, waardoor mijn spieren zulke bewegingen maakten, dat duizenden letters via mijn wordprocessor, op papier of op uw beeldscherm kwamen. Vele mensen beschouwen dit als een duidelijk voorbeeld van de wil die fysische processen veroorzaakt. Dat is interactionisme: fysische processen veroorzaken mentale processen en *vice versa* (Popper en Eccles 1977). Deze theorie sluit direct aan bij het alledaagse beleven. Daarom is interactionisme de meest algemene theorie hierover bij het grote publiek - zelfs voordat men erover nagedacht heeft.

##### Mentale gebeurtenissen en causaliteit

In eerste instantie lijkt causaal onderzoek naar mentale gebeurtenissen simpel. Als je een proefpersoon een visuele stimulus presenteert en hem vraagt wat hij waarneemt, kan hij dat zeggen. Als je dit meerdere malen doet, zegt hij (ongeveer) hetzelfde, en als je dat bij verschillende proefpersonen doet, zeggen zij ongeveer hetzelfde (het effect is dus reproduceerbaar). Als je de stimuli varieert, veranderen de responsies in de verwachte richting (het effect is valide).



*Figuur 5. De relaties tussen neurale en mentale gebeurtenissen. De zwarte pijlen geven causale relaties weer tussen neuronen. De pijlen met andere kleuren geven hypothetische 'invloeden' weer, maar dat zijn geen causale invloeden volgens de criteria van hoofdstuk 10.2. Er staan bijvoorbeeld groene pijlen tussen de afzonderlijke mentale gebeurtenissen, maar het is niet duidelijk wat men kan bedoelen met de opmerking 'een mentale gebeurtenis veroorzaakt een andere mentale gebeurtenis.'*

Dus stimulus en responsie zijn sterk gecorreleerd, de stimulus gaat vooraf aan de responsie, en manipuleren van de stimulus varieert de responsie in de voorspelde richting. Aan alle voorwaarden lijkt voldaan om te concluderen dat de stimulus de responsie veroorzaakt (vergelijk hoofdstuk 10.2.). Waarom zou je dan niet zeggen dat de stimulus de beleving veroorzaakt?

##### Interacties

De relevante causale verbanden volgens het interactionisme worden getoond in figuur 5. Het simpelste geval is dat neurale gebeurtenissen andere neurale gebeurtenissen ver-



oorzaken (fig. 5A). Dat verloop volgens de natuurwetenschappelijke interpretatie van veroorzaking (hoofdstuk 10.2.). Binnen het interactionisme gaat men ervan uit dat neurale gebeurtenissen mentale gebeurtenissen (belevingen) veroorzaken (figuur 5B), maar de natuurwetenschap heeft geen wet voor het ontstaan van belevingen. Binnen het interactionisme gaat men ervan uit dat 'de wil' gedrag en dus neurale gebeurtenissen veroorzaakt (figuur 5C), maar het is onduidelijk hoe men dit in overeenstemming brengt met natuurwetten over het behoud van energie en behoud van impuls. Men beleeft dat gedachten elkaar opvolgen, en daarom denkt men dat de ene gedachte de andere veroorzaakt (figuur 5D), maar in een natuurwetenschappelijk kader kan men niet verwoorden wat men bedoelt met 'de ene gedachte veroorzaakt de andere'.

### 3.1.2. Psychofysisch parallellisme

Hoe neurale processen mentale processen kunnen veroorzaken, is onduidelijk, en helemaal hoe mentale processen een oorzaak kunnen zijn van neurale processen - zonder dat fundamentele wetten uit de natuurkunde geschonden worden. Om dit probleem te omzeilen is het psychofysische parallellisme geformuleerd. Voor iedere mentale gebeurtenis is er een neurale gebeurtenis, maar er is geen causaal verband tussen. Er is een reeks neurale gebeurtenissen waarvan de eerdere steeds een oorzaak van de latere is. Tezamen met sommige neurale gebeurtenissen treden mentale gebeurtenissen op, zonder dat de laatste door de eerste veroorzaakt zijn: het zijn twee reeksen parallel lopende gebeurtenissen. De filosoof Gottfried Wilhelm Leibniz (1714) meende dat fysieke en mentale gebeurtenissen in principe onafhankelijk verlopen, maar dat ze dankzij God prachtig parallel verlopen.

Parallellisme is vergeleken met een film of een videoband met een geluidsspoor: ieder beeld correspondeert met een geluid, maar het is niet zo dat het beeld het geluid veroorzaakt of *vice versa*. Maar verder gaat deze vergelijking mank. Bij direct opgenomen beelden wordt de representatie van beeld en geluid veroorzaakt door het opgenomen feitelijke voorval, netjes volgens natuurkundige principes. Maar het geluid kan ook nagesynchroniseerd worden. De parallellist ziet dat neurale en mentale gebeurtenissen verbonden zijn, maar - zegt hij - niet causaal verbonden. Waarom weigert hij zo'n systematisch verband causaal te noemen? Is hij gewoon koppig? Er is echter een goede reden om dit verband niet causaal te

noemen: sinds Hempel (1965) bedoelen we met 'causaal verklaren' herleiden tot een natuurwet, en in het geval van neurale en mentale gebeurtenissen is zo'n natuurwet niet voorhanden.

### 3.1.3. Epifenomenalisme

Sommige denkers beschouwen mentale gebeurtenissen als een 'bijproduct' van neurale gebeurtenissen (Huxley 1874, Heil 2013). Dit is wel vergeleken met een klok: een klok is gemaakt om de tijd aan te geven, maar slingeruurwerken maken ook een tikkend geluid. Dat geluid is een bijproduct. Digitale klokken maakten aanvankelijk geen, maar men kan veel digitale klokken zo instellen dat ze een tikkend geluid maken: dat is een bijproduct van een bijproduct. De neurale processen gaan hun causale loop, en de mentale processen (de belevingen) zijn daar een afspiegeling van. Alsof een psyche via een videocamera op een beeldscherm processen ziet gebeuren, zonder er een invloed op te kunnen hebben.

Wat epifenomenalisme voor velen zo onaantrekkelijk maakt, is het idee dat mentale gebeurtenissen machteloos, impotent zijn: dat 'mentale processen er niets toe doen'. Alle gebeurtenissen in onze fysieke wereld, de hele geschiedenis, al onze daden en uitspraken, dat alles zou hetzelfde zijn geweest, als er geen mentale gebeurtenissen zouden zijn. Kathedralen zouden gebouwd zijn, en boeken geschreven, dank zij onze hersenen, ook wanneer er geen bewustzijn geweest zou zijn. Dat is strijdig met onze belevingen. *"Als het echter zou blijken dat ons geloof, verlangens en intenties causaal ineffectief zijn, dan lijkt het dat we moeten concluderen dat het een illusie is te denken dat we handelen door intenties of redenen."* (Davis 2013, p. 37).

### 3.1.4. Het Harde Probleem

Maar het probleem hoe mentale gebeurtenissen ontstaan, is veel groter dan men zich wenst te realiseren. In 1872 claimde Du Bois-Reymond dat het probleem van het ontstaan van bewustzijn (*"Wat is de oorsprong van de bewuste ervaring in onbewuste zenuwen?"*) nooit natuurwetenschappelijk opgelost zou kunnen worden (hoofdstuk 9.). Vooral het standpunt "we zullen het niet kunnen oplossen" stuitte op veel verzet bij vooruitgangsendenkers. Het denken van Du Bois-Reymond is in de Engelstalige literatuur verloren gegaan. Maar interactionisten zagen wel het probleem om het ontstaan van mentale processen causaal te

verklaren. Dit probleem kreeg in 1995 een naam: het Harde Probleem (*'the Hard Problem'*) van Chalmers (1995).

Volgens mij is er geen natuurwetenschappelijke verklaring mogelijk voor belevingen, omdat belevingen niet in het domein van de natuurwetenschap vallen. In het gekozen ontologische stelsel is er geen plaats voor belevingen, en zijn belevingen geen feitelijke gebeurtenissen. Chalmers (1995) noemde dit *'the Hard Problem'*, om dit te onderscheiden van *'easy problems'* die met natuurwetenschappelijke (materialistische) methodes opgelost kunnen worden. Maar Chalmers houdt een oplossing in principe voor mogelijk. *"Het Harde Probleem is een hard probleem, maar er is geen reden dat het permanent onopgelost zal blijven."* (Chalmers 1995, p. 218). ... Je kunt nooit weten. Immers, vroeger dacht men ook dat het onmogelijk was om doelgericht gedrag te verklaren op basis van causale processen.

### 3.1.5. Kritiek op het dualisme

Dualisme sluit direct aan bij de dagelijkse ervaring. We beleven dat we iets willen doen, even later doen we het, en we krijgen de sensorische terugmelding dat we het gedaan hebben. We nemen dingen waar, en de visuele, auditieve en tactiele belevingen gaan zo netjes hand in hand dat we vanzelf aannemen dat er fysieke objecten zijn die deze belevingen veroorzaken. We beleven een uitwendige wereld, en we gaan ervan uit dat een uitwendige wereld bestaat. Dat is het standpunt van iedereen, voordat hij of zij erover is gaan nadenken.

Er zijn twee opties: (1) er is een niet-stoffelijk 'zelf' of een 'ziel' die beleeft, of (2) belevingen zijn een onbegrepen eigenschap van materie als deze zo complex is als een brein.

Maar een algemeen wetenschappelijk en filosofisch principe is: eenvoud eerst. Als je de hele wereld kunt verklaren met één klasse verschijnselen, beperk je dan daartoe. Of *in concreto*: een materialist zegt dat mentale gebeurtenissen helemaal niet nodig om de fy-

sieke wereld te verklaren, dus beter achterwege kunnen blijven. En dat moet een dualist toegeven: er is niets in de fysieke wereld dat we dankzij mentale gebeurtenissen kunnen verklaren<sup>6</sup>. Bovendien wordt tegengeworpen, dat het strijdig lijkt met natuurwetten dat neurale gebeurtenissen een oorzaak kunnen zijn van mentale gebeurtenissen, en *vice versa*.

## 3.2. Materialistische visies

Materialistische denkers kunnen de volgende uitspraken doen. *"Gevoelens zijn niet meer dan fysisch-chemische processen in de hersenen."* *"Pijn is niets meer dan zenuwactiviteit in c-vezels."* *"Zien is alleen maar de absorptie van lichtquanten in het netvlies."* Hier geef ik twee materialistische visies.

### 3.2.1. Mentaal = neuraal: de 'identity theory'

De *'identity theory'* stelt dat 'mentale processen neurale processen in de hersenen zijn' (Place 1956, Feigl 1958, Smart 1959, Quine 1960, Davidson 1963, 1995)<sup>7</sup>. *"Alle mentale toestanden, gebeurtenissen en processen zijn toestanden, gebeurtenissen of processen van het centraal zenuwstelsel van gewervelde dieren."* (Bunge 1977, p. 506). Net zoals de Ochtendster en de Avondster bij nader inzien dezelfde planeet Venus zijn, zo zullen ook mentale en neurale gebeurtenissen bij nader inzien dezelfde gebeurtenissen blijken te zijn. Gooien we dan de eerdere analyse over het onderscheid tussen mentale en neurale gebeurtenissen overboord? Ja. Aanhangers van de *identity theory* laten aan de hand van voorbeelden uit de natuurwetenschap zien, dat twee aanvankelijk onderscheiden processen soms toch hetzelfde blijken te zijn, zoals bliksem en een elektrische ontlading in de lucht. Zo ook, zeggen ze, zijn mentale en neurale processen hetzelfde. De *identity theory* heeft op veel denkers een grote aantrekkingskracht. De *identity theory* is geworteld in de natuurwetenschap, en die heeft een groot prestige, en als monistische visie is de *identity theory* simpeler dan dualistische visies.

taal (Smullyan 1957). Mijn voorkeur is (bij nader inzien) zelf-referentie te vermijden.

<sup>7</sup> Deze visie wordt ook wel 'materialistisch reductionisme' genoemd. Maar ik prefereer het om de term reductionisme/reduceren te gebruiken als men een gebeurtenis op één niveau causaal verklaart door gebeurtenissen op een lager niveau. Daarvan is in het 'materialistisch reductionisme' geen sprake. Daar definieert men twee gebeurtenissen simpel als gelijk.

<sup>6</sup> Een speciaal geval is 'zelf-referentie'. *"Het neerschrijven of uitspreken van "Ik ben me bewust van ..." is een fysische gebeurtenis waarvan een mentale gebeurtenis een oorzaak is; het filosofisch standpunt dat mentale processen geen oorzaak zijn van fysische processen, lijkt onjuist."* (Van Dongen 1980, stelling bij een proefschrift). Maar zelf-referentie leidt nu eenmaal tot paradoxen [https://en.wikipedia.org/wiki/List\\_of\\_paradoxes](https://en.wikipedia.org/wiki/List_of_paradoxes). Als een taal zelf-referentie toelaat, wordt dit daardoor een zwakkere

**Psychofysisch reductionisme**

**Mentale processen causaal verklaard?** Stel iemand doet de uitspraak "Gevoelens zijn niet meer dan fysisch/chemische processen in de hersenen." Of processen in de psychologie herleid kunnen worden tot processen in de biologie, hangt ervan af welke psychologie men aanhangt.

- In de traditie van **William James** (1890) is psychologie de wetenschap van het mentale leven. *"Psychologie is de Wetenschap van het Mentale Leven, zowel van zijn verschijnselen als van de voorwaarden. De verschijnselen zijn dingen zoals gevoelens, verlangens, cognitie, redeneringen, beslissingen en dergelijke."* (James 1890, chapter 1). Volgens mij kunnen mentale processen niet herleid worden tot fysische processen (dit hoofdstuk en 10.2).
- In de traditie van **Watson** (1913, 1924), **Lashley** (1923) en **Skinner** (1938, 1953) is psychologie de wetenschap van gedrag. Zij bestudeerden uitsluitend direct waarneembare processen (= gedrag); belevingen maakten geen deel uit van causale verklaringen van gedrag. Gedrag is in principe herleidbaar (reducerbaar) tot biologie.

**Cultuur causaal verklaard?** In de antropologie en sociologie staat 'cultuur' centraal. Hier lopen verscheidene discussies door elkaar.

- Culturele gedragingen bij **dieren**. Bij dieren zijn er verscheidene gedragingen die door sociaal leren op de volgende generaties worden overgedragen. Zo ontstaan er lokaal verschillende varianten (hoofdstuk 3.2.). Ik heb voorgesteld om dit ook bij dieren 'cultuur' te noemen.
- Cultuur bij **mensen**. Uiteraard gelden hier dezelfde principes als bij psychologie. (1) Als men met 'cultuur' uitsluitend bedoelt de uiterlijk waarneembare (materiële) aspecten, dan kan die cultuur bij mensen herleid worden tot cultuur bij dieren; dit kan in principe herleid worden tot biologie. (2) Als men met 'cultuur' bij mensen ook de mentale aspecten bedoelt van de gewoonten van een volk - zoals nagenoeg alle antropologen dat doen -, dan kan men 'cultuur' bij mensen niet herleiden (reduceren) tot de gedragscultuur bij dieren. Als men echter met 'cultuur' bij dieren ook de mentale aspecten bedoelt, dan kan cultuur van mensen in principe herleid worden tot cultuur bij dieren. Maar helaas kan dit niet empirisch getoetst worden.

**Kritiek op de identity theory**

Twee belangrijke bezwaren tegen de *identity theory* zullen volgen.

Ten eerste, wordt gezegd, is het een **categoriefout** om te zeggen dat mentale en neurale gebeurtenissen identiek zijn. En een categoriefout is een doodsteek voor een wetenschappelijk idee. Als twee dingen of gebeurtenissen (A en B) identiek zijn, moet alles wat voor A geldt, ook voor B gelden. Welnu, neurale gebeurtenissen hebben een plaats in de hersenen, terwijl mentale gebeurtenissen geen plaats hebben; hoe kunnen ze dan hetzelfde zijn? Neurale gebeurtenissen zijn opeenvolgende toestanden van neuronen, terwijl mentale toestanden van mij zijn; hoe kunnen ze dan hetzelfde zijn? Een hersenonderzoeker kan in principe uit zijn meetsignalen van mijn brein weten wat ik denk. Maar ook dan weet hij hooguit mijn gedachten, terwijl ik mijn gedachten heb. Hij weet wanneer ik pijn heb, maar hij heeft mijn pijn niet. Filosofen discussiëren nog steeds of de grondgedachte van de *identity theory* een categoriefout is.

Ten tweede, wordt gezegd, hanteren aanhangers van de *identity theory* een **fout begrip** van mentale gebeurtenissen: hun begrip dekt niet het begrip in zijn oorspronkelijk bedoelde betekenis. Maar aanhangers van de *identity*

*theory* claimen dat mentale gebeurtenissen illusies zijn.

**3.2.2. Eliminatie van 'mentaal'**

Wat is het verband tussen mentale en neurale gebeurtenissen? Een origineel standpunt is het eliminiatiematerialisme (Ramsey 2013), voor het eerst duidelijk verwoord door Rorty (1965), en vooral verdedigd door Rey (1983) en het echtpaar Paul M. en Patricia S. Churchland (1984, 1986). Het eliminatiematerialisme gaat ervan uit dat onze belevingen illusies zijn<sup>8</sup>. *"Is fysikalisme al met al een verwerping van mentale objecten, of van een theorie over hen? Verwerpt het de mentale toestand van pijn of woede ten gunste van zijn fysische begeleider, of stelt het de mentale toestand gelijk met de fysieke toestand van het organisme (en dus de toestand van het fysiek organisme met de mentale toestand)?"* (Quine 1960, p. 265). Het is de vraag of er wel een relevant verschil is tussen de 'identity theory' en het eliminatiematerialisme.

Als we toetsbare kennis willen verwerven, moeten we hersenwetenschap bedrijven, in plaats van ons bezig te houden met voelen, horen, beleven en verlangen. Weg met de

<sup>8</sup> In de psychofysica spreekt men van een 'illusie' als een stimulus zodanig is dat gewone

proefpersonen iets waarnemen dat fysisch niet aanwezig is.

volkpsychologie. Wij kunnen toetsbaar praten over neurale gebeurtenissen: dan kunnen we precies verwoorden wat we bedoelen. Maar het gepraat over mentale gebeurtenissen leidt tot niets. Alle woorden voor mijn belevingen verwijzen naar niets anders dan mijn strikt persoonlijke belevingen. Communicatie hierover met u is bij voorbaat hopeloos als we wetenschappelijke helderheid willen verwerven, omdat mijn belevingen ontoegankelijk zijn voor u<sup>9</sup>. Men kan in de wetenschap maar beter helemaal niet meer spreken over mentale gebeurtenissen.

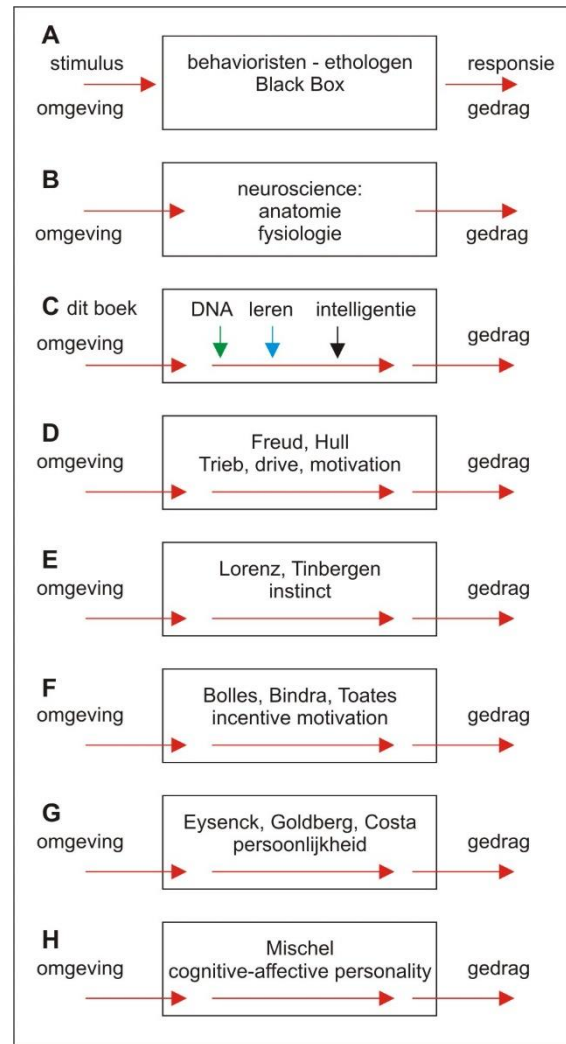
#### Kritiek op het eliminatie-materialisme

Het is nogal pijnlijk, als belevingen geëlimineerd worden, want het **enige** wat er voor mensen toe doet, zijn belevingen, want we kunnen niets anders beleven. “Sinds Descartes zijn *cogito* geformuleerd had, zijn filosofen overtuigd dat minstens sommige opvattingen evident zijn. Bijvoorbeeld: welk niet-circulair bewijs zou men kunnen aanvoeren voor het geloof dat men nu bewust is?” (Boghossian 2006, p. 116-117). Een eliminatie-materialist geeft de primaat aan de wereld van fysieke objecten, en elimineert alle belevingen. Maar volgens interactionisten kunnen mensen alleen door belevingen de wereld van de dingen kennen; een eliminatie-materialist zou de bron van alle kennis elimineren. Eliminatiematerialisten zoals Paul en Patricia Churchland zijn bijzonder, omdat zij jarenlang artikelen en boeken schrijven over het begrip dat zij zeggen geëlimineerd te hebben.

#### 4.3. Functionalisme

Ik heb mijn eigen ervaringen. Maar de ervaringen van een ander kan ik niet hebben. Ik kan zelfs niet vaststellen of een ander wel ervaringen heeft. Dit is het probleem van 'de psyche van een ander'. Heeft een ander mens een psyche? Hebben dieren een psyche? Hebben marsmannetjes een psyche - als ze bestaan natuurlijk - ? Hebben computers een psyche? Geen van deze vragen zijn logisch of experimenteel oplosbaar. Het is natuurlijk onbevredigend om over hersenen en mentale gebeurtenissen te discussiëren, zonder dat je een echt sterk argument hebt of je gesprekspartner al dan niet belevingen heeft.

Een functionalist omzeilt al deze problemen: als een functionalist waarneemt dat een ander organisme, een buitenaards wezen of een



Figuur 6. Allerlei biologische en psychologische theorieën zijn hypotheses over de invulling van de black box.

computer hetzelfde reageert op stimuli als hijzelf, dan neemt hij aan dat deze systemen gelijkaardig zijn aan hemzelf. Hij is niet geïnteresseerd wat er in de black box gebeurt, en hij bekommert zich niet over mentale processen. Het probleem van 'de psyche van een ander' is dan verschoven naar een analyse van input-outputrelaties (stimulus/responsrelaties, omgeving/gedragrelaties), zoals in de Turing-test (Turing 1950).

Juist zoals functionalisten zijn behavioristen en de meeste ethologen beroepsmatig niet geïnteresseerd in wat er binnen het organisme gebeurt. Zij bestuderen het verband tussen de buitenwereld en het gedrag van organismen.

<sup>9</sup> Ik raad wel aan om in een romantische context over belevingen te communiceren.



**Hebben computers een psyche?**

Functionalisten concluderen of een mens, dier of ding een psyche heeft uit input/output-reacties. Hebben computers dan een psyche? Dit criterium is verwant aan een criterium voor een intelligente computer. Laten we het volgende gedachte-experiment doen, de test van Turing (1950). Ik ga er gemakshalve van uit dat ik intelligent ben. Stel: u belt mij op en u krijgt niet mij maar mijn sprekende computer aan de lijn - zonder dat u dat weet overigens. U kunt er met de slimste vragen niet achterhalen of u met mij of met mijn computer spreekt. Dan moet ik fair zijn, en niet alleen zeggen dat ik intelligent ben, maar mijn computer ook. Moet ik dan ook niet fair zijn, en op zijn minst de mogelijkheid openlaten dat mijn computer een psyche heeft? Vooral Fodor (1983) heeft een lans gebroken voor de mogelijkheid dat computers een psyche hebben. Dit is overigens een filosofisch standpunt met juridische en sociale implicaties. Als computers belevenissen kunnen hebben, kunnen ze in principe ook pijn en leed ondervinden, En hebben ze dan geen recht op wetsbescherming? Als je een ander mens niet mag beschadigen of doden, mag je dan een intelligente computer wel slaan of kapot maken of zelfs uitzetten ... ? Een consequentie is: bejaardente-huizen voor afgeschreven intelligente computers, die aangesloten blijven op de netspanning, met onderhoudsmonteurs en voldoende afwisselende programma's.

#### 4.4. Idealisme: eliminatie van fysische gebeurtenissen

Bijna alle denkbare standpunten over het verband tussen mentale en neurale gebeurtenissen hebben nu aanhangers, behalve één opvallende afwezige: het filosofisch idealisme. Sommige filosofen concluderen: alle woorden voor 'dingen' zijn woorden voor herhaald voorkomende belevingen. Dit standpunt is het filosofisch idealisme. Voor mijn kennis zijn mijn belevingen primair. Secundair - want van belevingen afgeleid - zijn de dingen buiten mij en hun bestaan. Fysieke objecten zijn volgens het idealisme reeksen van belevingen. En belevingen 'bestaan' niet, als ze niet beleefd worden. Dus: fysieke objecten bestaan niet, als ze niet waargenomen worden. Tot zover in het kort het idealisme, dat de absolute tegenhanger van het eliminatie-materialisme is. Een bezwaar tegen het filosofisch idealisme is dan ook, dat deze denkers de veronderstelde bron van ons bestaan (de fysieke wereld) geëlimineerd hebben. Maar dat vinden de aanhangers geen bezwaar, maar juist een verdienste van hun denken.

#### 4.5. Een nieuwe metataal

**Proberen het onzegbare toch te zeggen?**

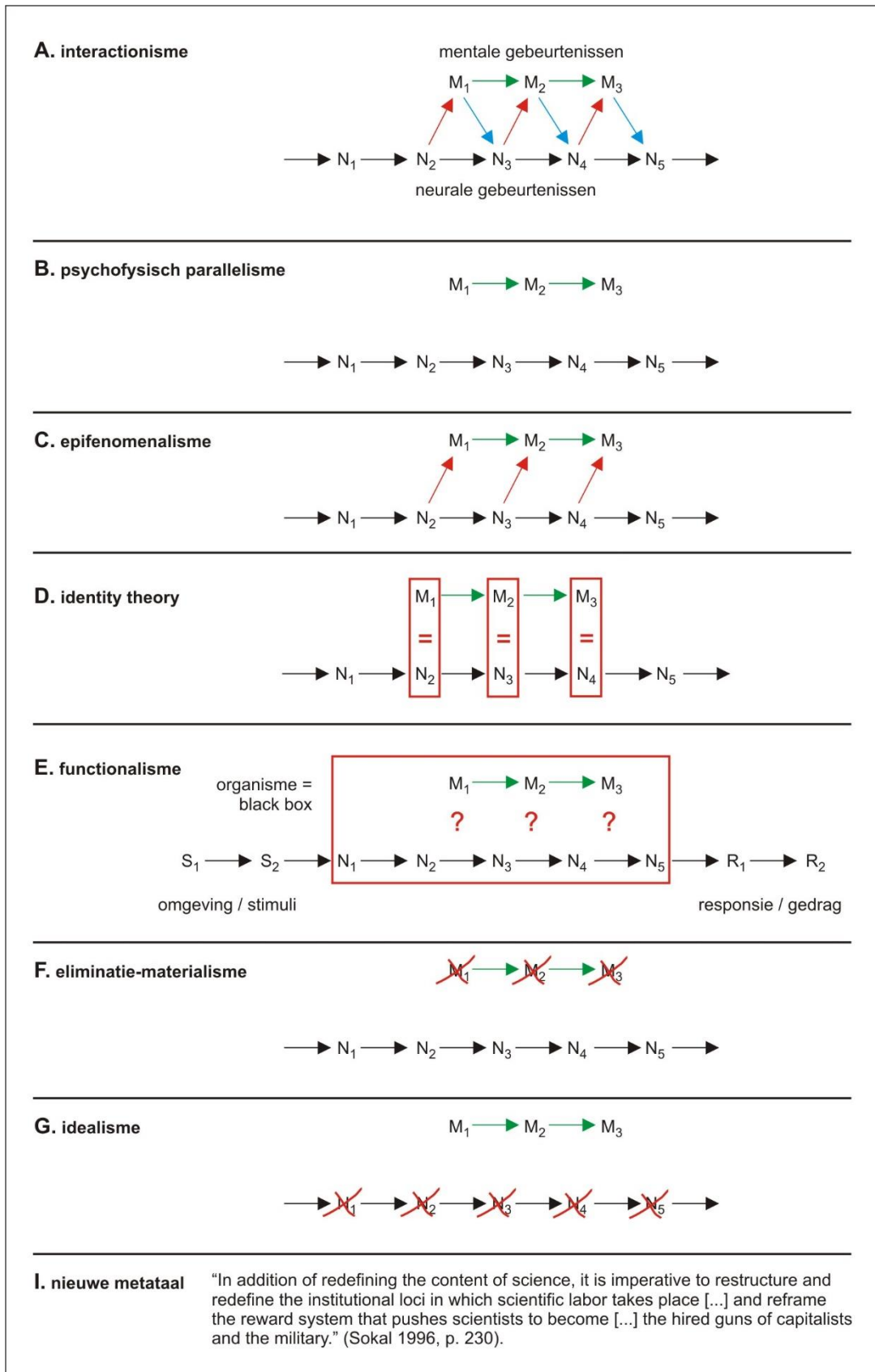
In het westerse denken worden lichaam en geest als afzonderlijk beschouwd. Deze dualistische visie is verwoord door René Descartes (1664). Maar het dualisme is nu uit de mode, en sommigen vinden het zelfs een denkfout (Dennett 1991, Damasio 1994). De westerse taal is doordrongen van dualistische termen, zodat vragen over bewustzijn en het zelf niet helder gesteld kunnen worden, en

zeker niet beantwoord. Een mogelijk uitweg is het ontwikkelen van een nieuwe metataal waarin dualistische termen niet meer voorkomen. Zo'n metataal kan ook een nieuw materiebegrip inhouden: een materie waarbij overal het mentale aanwezig is: panpsychisme (Seager en Allen-Hermanson 2015). Overigens is het een open vraag of een dergelijke metataal voor ons mensen (al dan niet uit een westerse cultuur) überhaupt mogelijk is. Zo'n metataal leidt gemakkelijk tot transcendente zinnen, die de een als geklets (Sokal 1996, Maudry 2011), en de ander als de Diepste Waarheid zal beschouwen.

#### 4.6. Overzicht van de visies

**4.6.1. De visies****Visies over hersenen en bewustzijn**

Het brain/mind-vraagstuk is een conceptueel vraagstuk (wat is "*mind*"?) en een causaliteitsvraagstuk. In hoofdstuk 10.2 probeer ik een heldere visie op causaliteit te presenteren; zo'n heldere visie is noodzakelijk voor een heldere discussie over het brain/mind-vraagstuk. Figuur 7 geeft een schematisch overzicht van de belangrijkste standpunten in de brein/psyche-discussie (b.v. Hospers 1967, Churchland 1984). Al eeuwen hebben filosofen zich systematisch afgevraagd wat de status van onze kennis over de wereld is. Als ik ● zie, ben ik zeker dat ik ● zie, maar ik kan nooit zeker zijn dat ● bestaat. Het hangt af van ons ontologisch uitgangspunt (Voorwoord), wat we kunnen bedoelen met 'dat ● bestaat'.



Figuur 7. De verbanden volgens de afzonderlijke visies. De zwarte pijlen geven causale relaties weer tussen neuronen. De groene pijlen suggereren relaties tussen mentale gebeurtenissen. De rode pijlen suggereren dat neurale gebeurtenissen mentale gebeurtenissen veroorzaken. De blauwe pijlen suggereren dat mentale gebeurtenissen neurale gebeurtenissen en gedrag veroorzaken.

Tabel 1. De sterke en zwakke punten van de verschillende visies op de relatie tussen brein en psyche.

	sterke punten	zwakke punten
<b>dualistische visies</b> <ul style="list-style-type: none"> <li>• interactionalisme</li> <li>• psychofysisch parallelisme</li> <li>• epifenomenalisme</li> </ul>	<ul style="list-style-type: none"> <li>• accepteert materie en psyche</li> <li>• sluit aan bij de dagelijkse ervaring</li> <li>• psyche heeft macht/invloed</li> <li>• accepteert materie en psyche</li> <li>• elimineert het causale probleem brein-psyche</li> <li>• accepteert materie en psyche</li> </ul>	<ul style="list-style-type: none"> <li>• kan niet aangeven hoe de interactie verloopt</li> <li>• kan niets extra's verklaren</li> <li>• kan niet aangeven hoe het verband brein-psyche in elkaar zit</li> <li>• kan niets extra's verklaren</li> <li>• psyche wordt wel geaccepteerd, maar is machteloos</li> <li>• kan niet aangeven hoe het verband brein-psyche in elkaar zit</li> <li>• kan niets extra's verklaren</li> <li>• psyche wordt wel geaccepteerd, maar is machteloos</li> </ul>
<b>functionalisme</b>	<ul style="list-style-type: none"> <li>• beperkt zich tot publiek waarneembare gebeurtenissen</li> <li>• elimineert het probleem van de psyche van een ander</li> </ul>	<ul style="list-style-type: none"> <li>• lost het brein/psyche-probleem niet op, maar omzeilt het</li> </ul>
<b>materialisme</b> <ul style="list-style-type: none"> <li>• <i>identity theory</i></li> <li>• eliminatie-materialisme</li> </ul>	<ul style="list-style-type: none"> <li>• eenvoudig (monisme)</li> <li>• sterke band met de natuurwetenschap</li> <li>• geeft het primaat aan de (veronderstelde) bron van ons bestaan (de stoffelijke wereld)</li> <li>• eenvoudig (monisme)</li> <li>• sterke band met de natuurwetenschap</li> </ul>	<ul style="list-style-type: none"> <li>• basisgedachte is wellicht categoriefout</li> <li>• hanteert psyche-begrip in een oneigenlijke betekenis</li> <li>• ontkent mentale gebeurtenissen</li> </ul>
<b>idealisme</b>	<ul style="list-style-type: none"> <li>• geeft het primaat aan de bron van alle kennis (mentale gebeurtenissen)</li> <li>• eenvoudig (monisme)</li> </ul>	<ul style="list-style-type: none"> <li>• ontkent de (veronderstelde) bron van ons bestaan (de stoffelijke wereld)</li> </ul>
<b>nieuwe metataal</b>	<ul style="list-style-type: none"> <li>• elimineert semantisch dualistische problemen</li> <li>• leidt tot een fraai transcendentaal wereldbeeld</li> </ul>	<ul style="list-style-type: none"> <li>• metataal bestaat nog niet en is wellicht onmogelijk</li> <li>• leidt tot vaag transcendentaal geklets</li> </ul>

### Geneesmiddelen, drugs en belevingen

Als iemand bepaalde stoffen inneemt, kan hij gaan hallucineren. En omgekeerd, als iemand hallucineert kunnen geneesmiddelen deze hallucinaties verminderen of wegnemen. Hallucinaties zijn mentale gebeurtenissen die door substanties opgeroepen en weggenomen kunnen worden. Dit lijken voorbeelden dat substanties onomstotelijk mentale gebeurtenissen veroorzaken of teniet doen. Dus lijkt het dat het interactionisme (of eventueel ook het epifenomenalisme) correct is, en de overige opvattingen niet. Zitten de andere visies

er naast? Niet noodzakelijk. Psychotrope stoffen hebben in het dualistische en in het materialistische denken invloed op neuronen. Het functionalisme omzeilt verhalen over het verband tussen deze neurale en mentale gebeurtenissen. En het is de vraag of een nieuwe metataal een samenhangend verhaal kan bieden.

De voorbeelden over hallucinaties zijn eerder voorbeelden dat het interactionisme diep in ons voorwetenschappelijk denken en onze voorwetenschappelijke taal geworteld zit, dan dat het duidelijke argumenten ten gunste van het interactionisme zijn

### De Modes van de Tijd

Filosofen hopen dat filosofie over Wijsheid gaat, maar in de praktijk is de inhoud van filosofische bespiegelingen afhankelijk van de Mode van de Tijd. Figuur 8 geeft een overzicht van de relevante denkrichtingen vanaf het jaar 1700. Het filosofisch idealisme was in die tijd nauwelijks vertegenwoordigd. Tot 1800 was het dualisme populair, hetzij uit overtuiging, of om straf van de kerk of overheid te vermijden. Na de Verlichting ontstond de eeuw van het materialisme. Binnen het materialisme, behaviorisme, de ethologie en het functionalisme zijn toetsbare uitspraken goed mogelijk. Denkers uit deze richtingen vonden het dualisme ronduit fout: het was 'de fout/vergissing van Descartes' (Dennett 1991, Damasio 1994). Maar na 1995 werd het dualisme geleidelijk weer respectabel met de erkenning van het Harde Probleem (Chalmers 1995, 1997). Daarnaast ontstond de cognitieve psychologie met 'Syntactic structures' van Noam Chomsky (1957). Cognitief psychologen zijn vaak niet expliciet over het verband tussen mentale en fysieke processen.

### Een keuze voor een visie

Het lijkt erop dat de afzonderlijke visies consistent of volledig zijn, maar niet beide. Vooral eliminatie-materialisten en idealisten hebben een consistent verhaal. Maar het belangrijkste argument tegen deze visies is, dat ze hetzij alle belevingen, hetzij de (veronderstelde) bron van ons bestaan geëlimineerd hebben - al zijn ze zelf juist zo trots op die eliminatie. Voor een positiebepaling heeft men dus een aantal alternatieven: een keuze uit de verschillende vormen van dualisme, materialisme, functionalisme of idealisme met al hun voor- en nadelen; of de hoop stellen op een nog te ontwikkelen metataal met diens voor- en nadelen. Mijn uitgangspunt is om zoveel mogelijk toetsbare zinnen te formuleren, zo-

dat foute zinnen snel door logica of observaties verworpen kunnen worden. Omdat uitspraken over mentale gebeurtenissen niet toetsbaar zijn (Carnap 1928, § 140), wil ik uitspraken over mentale gebeurtenissen zoveel mogelijk vermijden. Daardoor zijn mijn wetenschappelijke uitspraken in overeenstemming met het functionalisme of het eliminatie-materialisme. Dat is niet omdat ik belevingen onbelangrijk zou vinden. Integendeel: belevingen zijn het **enige** wat er voor mensen toe doen, want mensen kunnen niets anders beleven dan belevingen. Maar de inhoud van belevingen is niet toetsbaar.

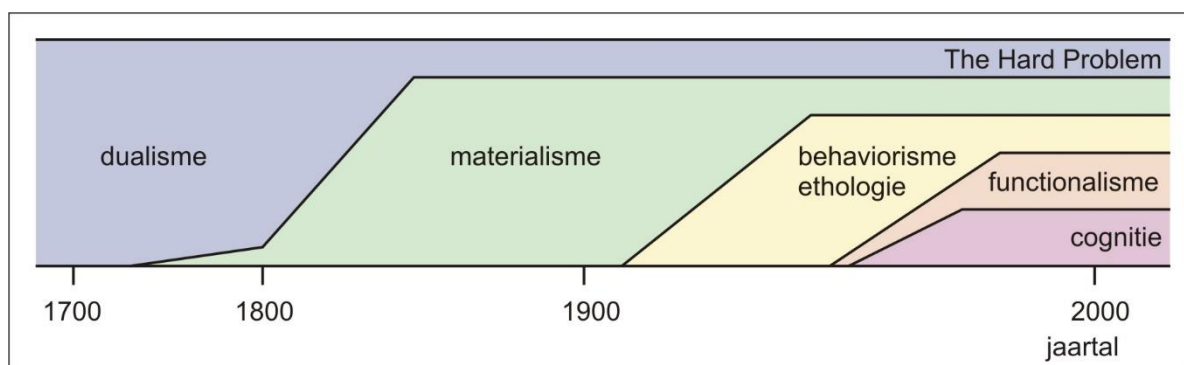
### 4.6.2. Denkers en visies

#### Denkrichtingen

Figuur 9 geeft een overzicht van enkele denkers over hersenen en bewustzijn. Voor dit schema heb ik deze denkers (nogal bruto) in drie categorieën ingedeeld.

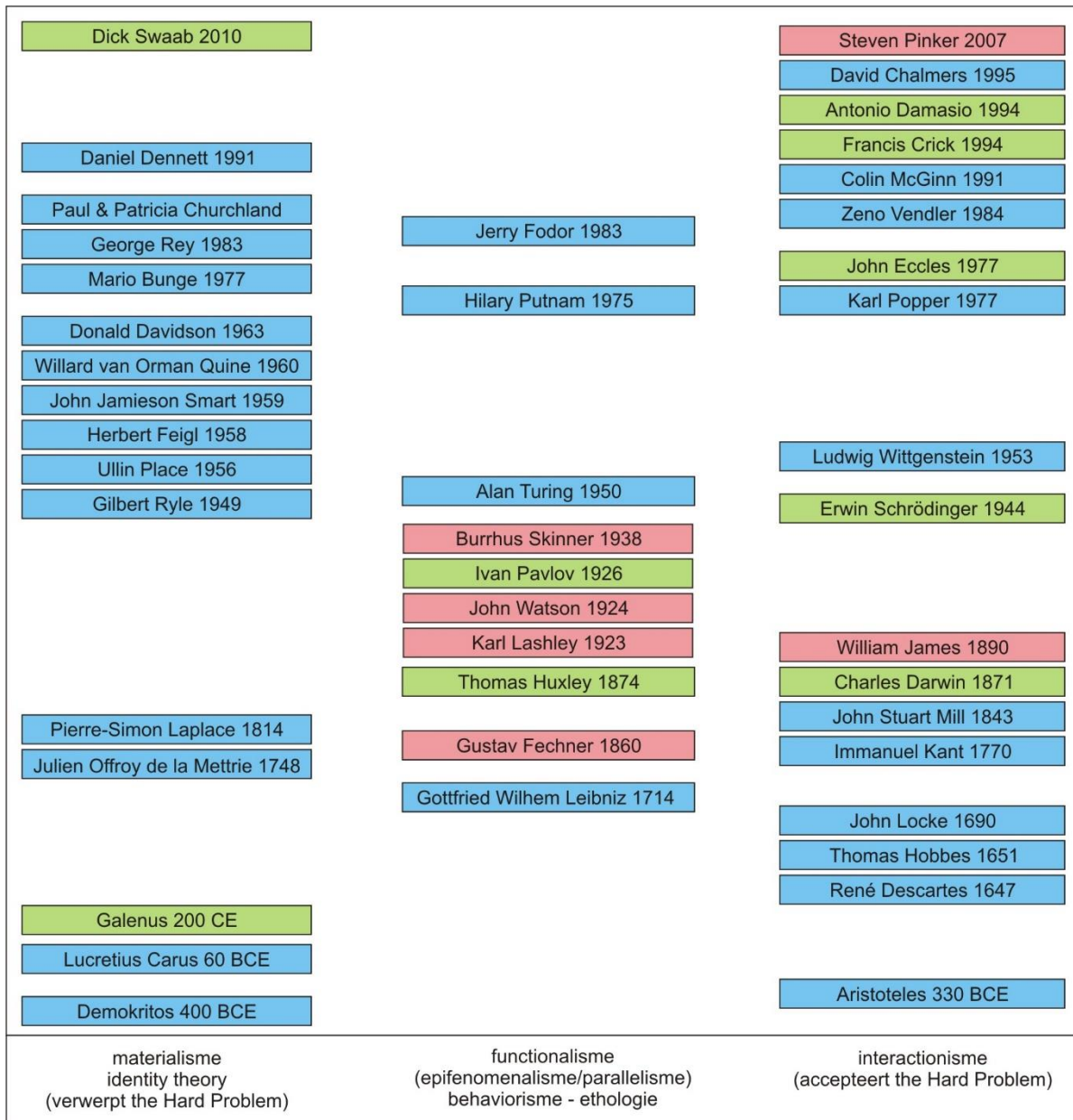
1. **Interactionisme.** Tot deze categorie reken ik denkers die mentale en fysieke processen als relevant beschouwen. Voor recente denkers houdt dit vaak in dat ze het Harde Probleem accepteren.
2. **Functionalisme.** Tot deze categorie reken ik denkers die primair in input/output-relaties geïnteresseerd zijn en niet zozeer in mentale processen of de inhoud van de black box. Functionalisten, behavioristen of ethologen zijn beroepsmatig niet in het mentale geïnteresseerd, maar alleen in het verband tussen buitenwereld en output/gedrag.
3. **Materialisme.** Tot deze categorie reken ik denkers die vooral geïnteresseerd zijn in de causale werking van het machientje en niet zozeer in belevingen.

Bij enkele denkers geef ik hier toelichting en citaten.



Figuur 8. Modes in het denken over het verband tussen neurale en mentale processen sinds het jaar 1700.





Figuur 9. Denkers over het verband tussen neurale en mentale processen, graf ingedeeld in 3 categorieën (zie tekst). Blauw: filosofen; groen: natuurwetenschappers; rood: psychologen.

### Mentale en fysische taal

Mentale gebeurtenissen worden direct beleefd (dit is een tautologie). Een mentale gebeurtenis gaat vooraf aan iedere taaluiting: er is een directe koppeling tussen mentaal en taal. Ik heb hier bijvoorbeeld voorgesteld een strikt onderscheid te maken tussen de woorden 'reden' (mentaal) en 'oorzaak' (fysisch). Maar in het dagelijks taalgebruik in het Grieks, Latijn, Frans, Engels, Duits en Nederlands worden deze woorden door elkaar gebruikt (Bunge 1979). Ook het woord 'doel' heeft mentale en fysische betekenissen (hoofdstuk 10.2.). Het is gemakkelijk om mentale processen te verwoorden, maar het is erg moeilijk om consequent

alleen in fysische (materialistische) termen te spreken. In de praktijk gebruiken verscheidene filosofen en psychologen die eigenlijk alleen materialistische uitspraken willen doen, toch nog verscheidene mentale termen (zie onder). Verscheidene psychologen verwerpen het interactionisme, maar zij blijven spreken in termen van belevingen en fysische processen; dit is een bizarre combinatie van ideeën. De meeste denkers doen uitspraken die alleen in een interactionistisch kader zinnig zijn.

### Thomas Hobbes

Hobbes (1651) wilde een materialistische leer verkondigen, maar "*Hobbes blijft inderdaad*

termen gebruiken die naar allerlei psychische processen verwijzen; daarmee hebben we bij Hobbes een soort materialistische psychologie zoals de meeste huidige gangbare psychologie tekstboeken.” (Wilm 1925, p. 88).

### Descartes

Volgens Descartes (1664) bestaan mensen uit twee 'substanties': materie (lichaam) en ziel, terwijl dieren uitsluitend uit materie bestaan zonder ziel. Descartes beschrijft fictieve wezens, die God geschapen zou hebben met de bedoeling dat ze zoveel mogelijk op ons, mensen, zouden lijken. Hij beschrijft de werking van deze wezens puur mechanisch, maar niettemin noemt hij ze 'mensen'<sup>10</sup>.

*“Deze mensen bestaan net als wij uit lichaam en ziel. Eerst beschrijf ik het lichaam op zich; daarna de ziel op zich; en tenslotte toon ik hoe deze twee naturen samengevoegd zouden moeten worden en verenigd.”* (Descartes 1664, p. 1). *“Helaas heeft Descartes niet al deze beloften ingelost: hij heeft alleen het lichaam besproken en bijna niets gezegd over de ziel en diens interacties met het lichaam.”* (Lokhorst 2015).

*Dualisme.* De meeste uitspraken van Descartes zijn dualistisch. Zintuiglijke indrukken worden door zeer fijne snelle deeltjes (*animal spirits*)<sup>11</sup> via holle zenuwen naar de hersenholttes gebracht. Als deze via de epifyse op de ziel inwerken, ontstaan er belevingen (Descartes 1664, p.73-74). Descartes (1637, p. 51) zegt expliciet dat er geen 'andere ogen in de hersenen' zijn die het beeld bekijken, waarmee hij bij voorbaat de suggestie van een 'Cartesian theatre' (Dennett 1991) verwierpt. De ziel of andere oorzaken leiden tot kleine bewegingen van de epifyse, die bewegingen van het lichaam veroorzaken. *“En omgekeerd is de machine van het lichaam zo samengesteld dat als deze klier op een of andere manier bewogen wordt door de ziel of door een andere oorzaak, deze dan de aanwezige geesten leidt naar de poriën van de hersenen die hen door de zenuwen naar de spieren leidt; zo laat deze klier door de geesten de ledematen bewegen.”* (Descartes e.a. 1985, p. 341). *“En de werking van de ziel bestaat volledig in het feit dat eenvoudig door iets te willen dat zij zorgt dat deze kleine klier waarmee ze zo nauw verbonden is, zo beweegt dat het gevolg tot stand komt dat correspondeert met deze wilsdaad.”* (Descartes

1649, p. 359).

*Materialisme.* Descartes besprak uitgebreid de werking van het lichaam, maar hij slaagde er niet in de werking van de ziel te bespreken (Lokhorst 2015, zie ook 'Een wijze raad van Descartes' in sectie 5). In gesprekken van Descartes en koningin Christina van Zweden wilde zij over de ziel en de liefde spreken, en hij over de mechanische werking van het lichaam; hun relatie bekoelde.

La Mettrie (1748) meende dat Descartes eigenlijk een materialist was, maar dat hij daar niet voor uitkwam uit vrees voor het Vaticaan. Overigens hebben anderen aan Descartes iedere mogelijke brain/mind-opvatting toegeschreven (Lokhorst 2015).

### Immanuel Kant

In zijn inaugurele rede als hoogleraar in 1770 benadrukte Kant het onderscheid tussen 'sensualitas' (waarnemingsvermogen) en 'intelligentia' (verstand). *“Waarnemingsvermogen is de ontvankelijkheid van een subject waardoor het mogelijk is dat zijn toestand de aanwezigheid van een voorwerp representeert. Intelligentie (rationaliteit) is het vermogen van een subject waardoor hij in staat is dat te representeren wat door zijn hoedanigheid de zintuigen niet kan passeren.”* (Kant 1770, section 2).

'Sensualitas' gaat over het empirisch onderzoek naar de mens, wat Kant 'antropologie' noemde, en waarover hij van 1772 tot 1796 college gaf (Kant 1798). *“Kant bedoelde met antropologie wat we tegenwoordig gedragspsychologie of experimentele psychologie noemen.”* (Brook 2013). 'Intelligentia' gaat over rationeel kennen en oordelen, wat hij in zijn 3 'Kritiken' behandelde. Ik ben van mening dat het denken van Kant vooral dualistisch was. Kant nam mentale processen serieus (Kant 1781), en hij nam fysieke processen serieus (Kant 1770, 1798). Hij benadrukt het onderscheid tussen rationaliteit en empirie. Volgens hem heeft de mens ratio om te bedenken wat juist gedrag is – en daar vervolgens naar te handelen. Het past alleen in het interactionisme dat mentale processen (ratio) een oorzaak zijn van fysieke processen (moreel juist gedrag).

Volgens Meerbote (1989) en Lau (2014) zijn de opvattingen van Kant vooral functionalistisch, maar daar ben ik het niet mee eens. Functionalisten zijn niet geïnteresseerd in wat

<sup>10</sup> Dit was een literaire truc om een veroordeling door het Vaticaan te vermijden, maar later besloot Descartes de tekst niet te publiceren (Descartes e.a. 1985, p. 79). De tekst is pas na het overlijden van Descartes gepubliceerd.

<sup>11</sup> Jan Swammerdam ontcrachtte snel het idee van *animal spirits* (Cobb 2002). Maar essentieel in het denken van Descartes is dat hij een materialistische verklaring zocht; die werd later gevonden in axonen en ionkanalen.

er in de black box gebeurt, en dus niet in mentale processen.

### John Stuart Mill

De uitspraken van John Stuart Mill vallen binnen het interactionisme. *“Laat het bijvoorbeeld aangetoond worden dat een zeer complexe reeks oorzaken en gevolgen elkaar opvolgen in het oog en in de hersenen om een beleving van kleur te produceren [...] dan is er nog op het einde van al deze bewegingen iets wat geen beweging is: er is de beleving van kleur. [...] De manier waarop ieder van deze bewegingen een andere beweging produceert, is misschien ontvankelijk voor verklaring door een of andere bewegingswet: maar de manier waarop de laatste beweging de beleving van kleur veroorzaakt kan niet door enige bewegingswet verklaard worden.”* (Mill 1843 Book V, Ch. V, section 3).

### Charles Darwin.

Darwin heeft verscheidene interactionistische uitspraken gedaan. *“Het feit dat lagere dieren door dezelfde emoties geprikkeld worden als wijzelf, is zo goed aangetoond dat ik de lezer niet ga vermoeien met details. Verschrikkingen werken bij hen op dezelfde manier als bij ons, waardoor de spieren trillen, het hart klopt van angst, de sfincters ontspannen, en het haar overeind gaat staan. Achterdocht, het kind van vrees, valt op bij de meeste wilde dieren.”* (Darwin 1871, p. 39). Het boek *'The expression of emotions'* (Darwin 1872) gaat niet over emoties, maar over gedrag (de **expressie** van emoties). In het slotwoord van dit boek schrijft Darwin (1872): *“... in dit hele boek heb ik vaak het gevoel gehad dat het moeilijk is de woorden de wil, bewustzijn en intentie correct te gebruiken.”*

Daarmee verwoordt Darwin het probleem dat het goed mogelijk is om consistent uitsluitend in mentale termen of in fysieke termen te spreken. Je kunt consistent spreken in termen van oorzaken en gevolgen. En je kunt consistent spreken in termen van redenen en intenties. Maar als je beide types concepten verbindt, resulteert het in niet-toetsbare zinnen.

### William James

Voor William James en veel andere psychologen is psychologie het wetenschappelijk onderzoek van bewuste ervaringen (of belevingen). Bewustzijn is altijd aanwezig geweest in levenloze en later in levende dingen. *“Als evolutie gladjes verloopt, dan moet bewustzijn in*

*een of andere vorm aanwezig zijn geweest bij de allereerste oorsprong van de dingen.”* (James 1890, p. 148-149). James (1890) beschouwt een 'ziel' als een voor de hand liggend idee: *“Volgens mij leidt het tot het minste logisch verzet te denken dat de ziel op een mysterieuze manier door hersentoestanden beïnvloed wordt, en daarop reageert met bewuste affecties van zichzelf.”* (James 1890, p. 348 - 349). Maar het denken van James lijkt nog het meeste op het panpsychisme, van een bezielde levende en levenloze natuur (hoofdstuk 8.2.).

### John Watson

John Watson is de grondlegger van het behaviorisme. Watson accepteerde wel dat er mentale processen zijn, maar die kunnen niet wetenschappelijk onderzocht worden. *“In de ogen van een behaviorist, is psychologie een puur objectieve experimentele tak van de natuurwetenschap. Het theoretische doel ervan is gedrag te voorspellen en te beïnvloeden.”* (Watson 1913, p. 158). *“Mijn doelstelling is niet om bewustzijnstoestanden te beschrijven en te verklaren.”* (Watson 1913, p. 168)<sup>12</sup>. Het is verstandiger mentale processen buiten beschouwing te laten. *“Twee tegengestelde gezichtspunten zijn nog dominant in het Amerikaans psychologisch denken: introspectieve of subjectieve psychologie, en behaviorisme of objectieve psychologie.”* (Watson 1924). *“Het behaviorisme claimt dat bewustzijn noch een welomschreven noch een bruikbaar concept is.”* (Watson 1924, p. 2). Volgens Watson is introspectie een foute aanname voor een toetsbare, wetenschappelijke psychologie. *“Door de basisaanname dat er zo iets als bewustzijn is en dat we het met introspectie kunnen bestuderen, vinden we evenveel analyses als er individuele psychologen zijn. [...] De pakweg 30 vruchteloze jaren sinds de vestiging van het laboratorium van Wundt, hebben definitief bewezen dat de zogenaamde introspectieve psychologie uit Duitsland op de verkeerde hypothesen is gebaseerd.”* (Watson 1924, p. 5). Watson beschouwt het gebruik van mentale termen in de wetenschap alleen maar als verwarrend.

### Niko Tinbergen

*“Om een voorbeeld te noemen: de conclusie dat een dier jaagt omdat hij hongerig is, zal veel mensen meteen tevreden stellen. Toch is het woord 'hongerig' onvoldoende duidelijk,*

<sup>12</sup> De doelstelling van Watson is daarmee hetzelfde als van mij: gedrag causaal verklaren in termen van toetsbare processen.

omdat het gebruikt kan worden als (1) een **geschikte beschrijving** van de toestand van een dier op basis van subjectieve en objectieve criteria. Als het woord zo gebruikt wordt, is die conclusie hooguit voorlopig en niet bevredigend voor de wetenschapper die wil weten wat er in het dier gebeurt als het in deze toestand is. Hij zal proberen uit te vinden welke impulsen de spieren prikkelen bij het voedsel zoeken, waar die impulsen vandaan komen, enz. Maar als we de conclusie dat een dier jaagt omdat hij honger heeft, letterlijk nemen, dus (2) als **causale verklaring**, en als de claim dat het subjectieve verschijnsel van honger een van de oorzaken is van voedsel-zoekgedrag, dan worden fysiologische en psychologische beschrijvingen door elkaar gegooid. Zoals eerder opgemerkt – de etholoog wil het eventuele bestaan van subjectieve verschijnselen bij dieren niet ontkennen, maar hij vindt het nutteloos deze als verklaring te presenteren, want ze kunnen niet waargenomen worden met wetenschappelijke methoden. [...] Het onderzoek naar doelgerichtheid, het onderzoek van subjectieve verschijnselen, en het onderzoek naar oorzaken, zijn drie verschillende manieren om over gedrag te denken; ieder daarvan is consistent bij de toepassing van zijn eigen methoden. Maar wanneer zij andermans terrein betreden, is verwarring het gevolg.” (Tinbergen 1951, p. 4-5).

“Zij weten dat mensen vaak intense emoties ervaren tijdens bepaalde fases van gedrag, en zij merken op dat het gedrag van veel dieren vaak lijkt op ons ‘emotioneel’ gedrag; daarom concluderen zij dat dieren dezelfde emoties beleven als wijzelf. Maar velen gaan nog verder, en houden vol dat emoties en andere subjectieve factoren causale factoren zijn, in de wetenschappelijke betekenis van dat woord.” (Tinbergen 1951, p. 4). Er is ook ‘spontaan gedrag’ dat opeens in een constante omgeving optreedt. Het is een uitdaging voor de wetenschap om interne processen te identificeren die dit ‘spontaan gedrag’ veroorzaken. Sommigen dachten dat die interne processen uitsluitend zuiver subjectieve, mentale processen waren. Maar “dat kwam door de verwarring van ‘spontaniteit’ met het wezenlijk subjectieve begrip ‘vrije wil’.” (Tinbergen 1951, p. 71).

### Ludwig Wittgenstein

Ludwig Wittgenstein heeft verscheidene uitspraken gedaan die het best in een interactionistisch kader passen. Hij erkent mentale processen: “Waarom zou ik ooit ontkennen dat er een mentaal proces is? Alleen: “Er heeft zojuist

het mentaal proces van herinneren in mij plaatsgevonden...” betekent niets meer dan “Ik heb me net herinnerd...” Het mentaal proces ontkennen betekent het ontkennen van herinneren: ontkennen dat iemand zich ooit iets herinnert.” (Wittgenstein 1953, §306). Op de vraag of hij een behaviorist was, antwoordde hij: “Nu lijkt het wel alsof we mentale processen ontkend hebben. Maar die willen we natuurlijk niet ontkennen.” (Wittgenstein 1953, §308). Ook het onderscheid dat hij maakte tussen oorzaak en reden (hoofdstuk 10.2.), past alleen in het interactionistisch denken.

### Roger Sperry

De neuropsycholoog en neurobioloog Roger Sperry (1970) heeft onderzoek gedaan aan split-brain patiënten (zie boven). Eccles, die zelf expliciet een interactionist is, beweert dat ook Sperry uitspraken doet die alleen in een interactionistisch kader passen. Sperry (1980) noemde zichzelf ‘mentalist’ en hij verwierp het dualisme, mede omdat hij dat associeerde met bovennatuurlijke verschijnselen en het voortbestaan van de ziel na de dood. Maar ik ben het eens met de opmerking van Eccles dat uitspraken van Sperry vooral interactionistisch zijn.

### Karl Popper en John Eccles

John Eccles was een Australisch neurofysioloog die in 1963 de Nobelprijs kreeg voor zijn onderzoek aan ruggenmergreflexen. Karl Popper was een Oostenrijks-Engels filosoof, die zijn belangrijkste idee, het falsificatieprincipe ontleend heeft aan de Duitse Joodse psycholoog Otto Selz (1881 – 1943, zie Ter hark 2004).

Eccles en Popper hebben samen een boek geschreven, ‘*The self and its brain*’ (Popper en Eccles 1977). Hierin beschrijven ze hun interactionistische visie van de ‘wereld van de dingen’ en de ‘wereld van belevingen’<sup>13</sup> die elkaar causaal beïnvloeden. Die interactie zou plaatsvinden in het ‘*liaison brain*’<sup>14</sup>. Eccles (1994) speculeert dat de *supplementary motor area* (tekstkader ‘De *Bereitschaftspotential* en de hersenen’) onderdeel is van het *liaison brain*

### Daniel Dennett

De filosoof Daniel Dennett (1991) is uiterst negatief over het dualisme, maar hij weigert zijn positie te bepalen in heldere woorden. “Ik weiger één formele, correct gekwalificeerde uitspraak te doen waarmee mijn theorie in een notendop zou worden samengevat.” (Dennett e.a. 1993, p. 523). “Uiteraard ben ik min of

<sup>13</sup> Daarnaast onderscheiden zij een derde wereld: de wereld van wetenschappelijke kennis.

<sup>14</sup> ‘*Liaison*’ betekent vooral ‘stiekeme seksuele relatie’.



meer een 'teleofunctionalist' ..." (Dennett e.a. 1993, p. 524). Hij vermijdt het gebruik van toetsbare zinnen.

#### Antonio Damasio

De neuroloog Antonio Damasio leek het dualisme te verwerpen; de titel van zijn boek was dan ook 'Descartes' error'. Toch doet hij interactionistische uitspraken: *"De elementaire geheimen van de geest verblijven binnen de interactie van de vuurpatronen gegenereerd door veel circuits van neuronen, lokaal en globaal, van moment tot moment, in het brein van een levend organisme."* (Damasio 1994, p. 259).

#### David Chalmers

De Australische filosoof David Chalmers (geboren 1966) onderscheidt twee type problemen.

- Gemakkelijke problemen zoals object discriminatie en het formuleren van zinnen. Deze kunnen in principe causaal verklaard worden door de werking van neuronen. Hiervan geeft hij verscheidene voorbeelden.
- *"Het echte Harde Probleem van bewustzijn is het probleem van beleving, [...] Het is breed geaccepteerd dat beleving ontstaat vanuit een fysieke basis, maar we hebben geen goede verklaring waarom en hoe bewuste beleving ontstaat."* (Chalmers 1995, p. 201). Helaas geeft Chalmers geen heldere causale formulering ('ontstaat vanuit') en gebruikt hij het complexe woord 'waarom'.

Chalmers beschouwt het Harde Probleem niet als fundamenteel onoplosbaar. *"Het Harde Probleem is een hard probleem, maar er is geen reden dat het permanent onopgelost zal blijven."* (Chalmers 1995, p. 218).

Chalmers noemt zijn eigen visie "naturalistic dualism". Het is dualisme, omdat in zijn denken mentale gebeurtenissen ontologisch anders zijn dan fysieke gebeurtenissen, en niet tot fysieke gebeurtenissen gereduceerd kunnen worden. Chalmers is kritisch over andere theorieën over bewustzijn. *"De meeste bestaande theorieën over bewustzijn hetzij ontkennen het bestaan ervan, of verklaren iets anders. Of zijn verheffen het probleem tot een eeuwig mysterie."* (Chalmers (1995, p. 217).

## 4. Bewustzijn bij dieren?

Na bovenstaande bespreking van mentale processen bij mensen, zal het duidelijk is dat het zeer moeilijk is om in toetsbare zinnen over het bewustzijn van dieren te spreken.

#### Descartes

Volgens Descartes (1662) gedragen dieren zich als automaten op basis van hun 'passies', maar alleen mensen hebben een ziel en vrije wil, zodat ze tegen hun passies kunnen ingaan. Dieren reageren weliswaar op positieve of negatieve prikkels, maar dat zouden volgens Descartes slechts automatische reacties zijn.

#### 'Inzicht' bij dieren?

Wolfgang Köhler<sup>15</sup> (1921) confronteerde chimpansees met allerlei problemen die met stokken of dozen oplosbaar waren. Als zo'n chimpansee opeens efficiënt een oplossing toepaste, noemde Köhler dat 'inzicht'. *"De chimpansees vertonen gedrag met inzicht, zoals we dat ook bij mensen zien."* (Köhler 1921, p. 191). Köhler zag een fundamenteel onderscheid tussen leren-met-inzicht en *trial-and-error* leren. *"... de term 'inzicht' van Köhler kan slechts betekenen dat het dier niet eenvoudig op reflex-achtige manier reageert op bepaalde zintuigelijke stimuli, maar dat het dier in zekere zin relevante of essentiële relaties grijpt tussen eigenschappen of dingen in de totale situatie."* (McDougall en McDougall 1931, p. 237-238). 'Inzicht' vergt een overzicht over de aanwezige situatie, maar ook anticiperen op een toekomstige oplossing ('vooruitzien'). Ik heb operationele criteria vermeld voor gedrag-met-inzicht (hoofdstuk 3.5.), maar het grote probleem met 'inzicht' is dat het onvermijdelijk associeert met bewustzijn.

#### Behaviorisme

Uitspraken over het al dan niet voorkomen van gevoelens bij dieren, zijn vrijblijvend en onoplosbaar, daarom is het begrip 'bewustzijn' geschrapt door behavioristen zoals Lashley (1923), Watson (1924) en Skinner (1938, 1953). Deze onderzoekers wilden van de psychologie een exacte wetenschap maken, en daartoe spraken zij uitsluitend in termen van gedrag, en vermeden zij te spreken in termen in niet-observeerbare interne processen.

#### Ethologie

Sommige ethologen accepteren dat dieren wellicht bewuste belevingen kunnen hebben, maar deze kunnen niet wetenschappelijk bestudeerd worden. *"Hoewel een etholoog, zoals*

<sup>15</sup> Hij is ook een van de voormannen van de Gestaltpsychologie.

we eerder gezegd hebben, het eventuele bestaan van subjectieve verschijnselen bij dieren niet wil ontkennen, claimt hij dat het nutteloos is dit als oorzaken te presenteren, omdat ze niet met wetenschappelijke methoden geobserveerd kunnen worden." (Tinbergen 1951, p. 4-5). Ik ben het met Tinbergen eens.

### Donald Griffin

De bioloog Donald Griffin (1976, 1978) claimde dat er continuïteit is in de evolutie van mentale processen, en hij concludeerde dat ook dieren bewustzijn hebben. Dat noemde hij 'cognitieve ethologie'<sup>16</sup>. Griffin (1984, p.37) probeerde een gedragscriterium te formuleren voor bewustzijn bij dieren: "Dus het lijkt waarschijnlijk dat een breed toepasbaar, zo niet alles omvattend, criterium voor bewust besef bij dieren is veelzijdige aanpasbaarheid van het gedrag aan veranderende omstandigheden en uitdagingen." Ristau (1991) en Bekoff (1995) hangen ook dit standpunt aan. Maar naar mijn mening is flexibel, doelgericht gedrag een criterium voor intelligentie (ook van kunstmatige intelligentie van computers) en niet voor bewustzijn (hoofdstuk 3.5.).

### Frans de Waal

De bioloog Frans de Waal heeft verdedigd om gedachten en emoties van dieren onderdeel te laten zijn van een wetenschappelijke verklaring van gedrag. Volgens De Waal (1982, p. 38) hebben chimpansees "het vermogen om doelgericht te denken". En "De vaardigheid om eerdere ervaringen te combineren om een doel te bereiken wat het beste omschreven kan worden als redeneren en denken – betere woorden bestaan niet." Ik zeg liever dat chimpansees leren en intelligent gedrag vertonen, inclusief *goal-oriented* gedrag, wat empirisch aangetoond kan worden (hoofdstuk 3.5.). Maar liever gebruik ik de mentalistische termen 'redeneren' en 'denken' niet, omdat dit niet empirisch aangetoond kan worden.

Volgens Frans de Waal is het belangrijk "om vragen die toch niet beantwoord kunnen worden, te vermijden, en om emoties te beschouwen als mentale en lichamelijke toestanden die gedrag bevorderen dat past bij de uitdagingen van de omgeving. [...] Het taboe dat zolang geregeerd heeft in het gedragsonderzoek van dieren, lijkt verouderd." (de Waal 2011, p. 191). Volgens mij produceer je juist onbeantwoordbare vragen, als je emoties tegelijk als mentale en lichamelijke toestanden beschouwt. Er is niet zozeer een taboe op dergelijke woorden,

maar het is een tactische, verstandige spelregel om vervelende oeverloze discussies te vermijden over onderwerpen die toch niet opgelost kunnen worden.

### Hillary Kornblith

De filosoof Hillary Kornblith (2002, 2005) wilde een verband leggen tussen kennis bij dieren en bij mensen, waarbij hij kennis als mentaal proces beschouwde. Ik denk dat dieren en mensen op dezelfde manier kennis verwerven, en dat alle kennis 'leerinhoud' is (hoofdstuk 3.2.). In die betekenis is 'kennis' bij mensen en andere dieren het gevolg van causale leerprocessen. Ik licht dit toe aan de hand van het gedrag van trekvogels. Ik ga ervan uit dat trekvogels de eerste maal geen erfelijke 'kennis' hebben van hun wintergebied; maar de volgende keren hebben ze aangeleerde kennis.

De belangrijkste spelregel in mijn werk is dat alleen **toetsbare uitspraken** worden geaccepteerd. Een gevolg daarvan is dat in dit intellectuele spel fysieke stimuli, gedrag, de waarneembare gevolgen van gedrag, en de voorgeschiedenis van het individu als verklaring worden geaccepteerd, en mentale gebeurtenissen niet.

## 5. Besluit

### Maak het juiste onderscheid

Speciaal als men toetsbare uitspraken wil doen over het verband tussen hersenen en bewustzijn, moet men zijn woorden zorgvuldig kiezen.

1. Als men toetsbare uitspraken wil doen, moet men formele en empirische wetenschappen onderscheiden (hoofdstuk 10.1.).
2. Als men toetsbare uitspraken wil doen, moet men neurale en mentale processen onderscheiden. Daarom onderscheid ik oorzaak en reden, en ook gevolg en *goal* en *purpose* (hoofdstuk 10.2.). Het is denkbaar dat neurale en mentale processen hetzelfde zijn (dus dat aanhangers van de *identity theory* uiteindelijk toch gelijk krijgen). Maar om tactische redenen benoem ik neurale en mentale processen verschillend. Als ze toch hetzelfde zijn, kunnen de uitspraken eenvoudig samengevoegd worden. Maar als neurale en mentale processen als hetzelfde benoemd zijn, en ze zijn uiteindelijk toch verschillend, dan zijn de eerdere uitspraken onbruikbaar.

<sup>16</sup> Het woord 'cognitief' heeft vaak een mentale lading.

**“Bewustzijn herwonnen, maar toetsbaarheid verloren”****“Consciousness regained, testability lost”**

Men is uiteraard vrij in een keuze voor wetenschappelijke verklaringen met of zonder mentale gebeurtenissen. Dit is gerelateerd aan de keuze of men processen in de wereld (in principe) al dan niet wil verklaren door causale processen; en dat is weer gerelateerd aan de keuze voor of tegen reductionisme. Maar iedere keuze heeft gevolgen; “bewustzijn herwonnen” (Humphrey 1984), toetsbaarheid verloren. *“Hoewel een etholoog, zoals we eerder gezegd hebben, het eventuele bestaan van subjectieve verschijnselen bij dieren niet wil ontkennen, claimt hij dat het nutteloos is dit als oorzaken te presenteren, omdat ze niet met wetenschappelijke methoden geobserveerd kunnen worden.”* (Tinbergen 1951, p. 5). Dit is niet alleen ‘nutteloos’, maar ook schadelijk, want het leidt tot langdurige, onoplosbare (en vervelende) discussies. Ethologen en behavioristen vermijden mentale gebeurtenissen in hun wetenschappelijke verklaringen. Cognitief psychologen daarentegen, hechten eraan mentale processen te gebruiken in hun verklaringsmodellen. Daarvoor betalen ze de prijs dat hun zinnen niet langer toetsbaar zijn – al vinden sommigen het winst als niemand kan aantonen dat ze ongelijk hadden.

**Taal en bewustzijn**

In de voorbeelden van *blindsight* en *split-brain* patiënten vertoonden de patiënten complex gedrag, terwijl ze niet konden zeggen waarom ze iets gedaan hadden. Ondanks het complexe gedrag, waren er niet de corresponderende uitspraken en – denken we – niet de corresponderende mentale processen. Toch is het niet correct om als criterium voor bewustzijn te hanteren dat men de inhoud in taal kan verwoorden. In het hoofdstuk over taal (7.2.) heb ik getoond (1) dat er bij dieren en kinderen concepten waren, voordat er woorden waren, en (2) dat er veel woordloze concepten zijn. Men kan van geuren en smaken een heldere beleving en geheugen hebben, zonder dat men dit helder in woorden kan omschrijven. Sommige grote wetenschappelijke ontdekkingen zijn gedaan door wetenschappers die in beelden en niet in woorden dachten (Pinker 1994, p. 61-62).

We gaan ervan uit dat het neuraal commando voor een taaluiting er milliseconden eerder was dan de taaluiting, en dat dit neuraal commando de uitgesproken taaluiting echt veroorzaakte. In mentale termen: we beleven dat we iets willen zeggen, en we zeggen het daarna.

**De oorsprong van bewustzijn**

Keith (1948), Jerison (1973) en Dennett (1983, 1991) denken dat mentale processen vanzelf in de evolutie ontstaan, als de hersenen voldoende complex zijn. Fodor (1981) meent dat bewustzijn ontstaan is toen er terugkoppelverbindingen in de hersenen ontstaan zijn; maar let wel (1) al bij ‘lagere’ organismen zijn er veel terugkoppelverbindingen in het zenuwstelsel, en (2) oude terugkoppelsystemen regelden o.a. bloedsuikerspiegel,

bloeddruk en lichaamstemperatuur, en die werken onbewust. In dit denken zijn mentale processen een indirect product van de evolutie van de hersenen. Als iets het product van biologische evolutie is, dan is het in de regel ‘evolutionair voordelig’, wat inhoudt dat het een gunstige invloed heeft op overleven of voortplanten van het individu of zijn verwanten. Zijn dan mentale gebeurtenissen evolutionair voordelig? Of anders gezegd; is een organisme met mentale processen in de strijd om het bestaan beter af dan zijn broertje met dezelfde hersencapaciteit, maar zonder mentale processen? Is er iets dat we dank zij mentale processen kunnen doen, terwijl we dat anders niet (of niet zo goed) zouden kunnen<sup>17</sup>? Blijven we met de wellicht teleurstellende gedachte zitten dat er geen dwingende argumenten zijn dat mentale gebeurtenissen gedrag kunnen beïnvloeden. Het is zeer de vraag of mentale processen konden ontstaan door Darwinistische evolutie. Eigenlijk zijn dergelijke vragen over de evolutie van het bewustzijn nutteloos, omdat ze toch niet toetsbaar te beantwoorden zijn. Hierover denken is tijdsverspilling. Maar het is wel wetenschappelijk relevant het verband te onderzoeken tussen hersenen en intelligenties van mensen en andere dieren (hoofdstuk 3.5.). Ik heb besloten mijn tijd niet te verdoen met de onoplosbare vraag wanneer het bewustzijn is ontstaan.

<sup>17</sup> In deze vraag moeten we ‘kunnen doen’ beperken tot uitwendig waarneembaar gedrag. Immers alle deelnemers aan de discussie zijn het eens dat uitwendig waarneembaar gedrag

invloed heeft op overleven of voortplanten; er is geen eensgezindheid of dit ook geldt voor mentale gebeurtenissen.

**Oneigenlijk gebruik van het woord 'verantwoordelijk'.**

In sommige contexten komen we het woord 'verantwoordelijk' tegen:

- Soms worden levenloze dingen 'verantwoordelijk' genoemd; de betekenis is 'veroorzaakt'. Bijvoorbeeld "... heeft een gen ontdekt dat verantwoordelijk was voor een vorm van erfelijke amyotrofe laterale sclerose..." (Aebischer en Kato 2007).
- Soms worden abstracties 'verantwoordelijk' genoemd. Bijvoorbeeld: "*Verschillen in alle frequenties tussen individuen en tussen volken [...] zijn verantwoordelijk voor fenotypische variabiliteit ...*" (Al-Yahyaee e.a. 2007). Hier lijkt de betekenis zoiets als 'veroorzaakt', maar ik vind het oneigenlijk taalgebruik dat een abstractie een oorzaak is.

Dit zijn voorbeelden van ongelukkig metaforisch gebruik van het woord 'verantwoordelijk'. Soms gebruikt men "A is verantwoordelijk voor B", in de betekenis "A is een oorzaak van B", waarbij A en B soms toestanden/gebeurtenissen zijn van (fysieke) objecten, en soms van abstracte concepten.

**Het Harde Probleem**

Mill (1843), Du Bois-Reymond (1872), Nagel (1979) en McGinn (1991) concludeerden dat mensen de interactie tussen fysieke en mentale processen nooit causaal zullen begrijpen<sup>18</sup>. Dat noemde Chalmers (1995) het "Harde Probleem". Dennett (1991) heeft het boek '*Consciousness explained*' gepubliceerd, maar "**Dennett heeft het bewustzijn niet verklaard**".

**Te klein verstand.** Volgens Pinker is het menselijk verstand te klein om de brain/mind-interactie te begrijpen, al is het mogelijk dat "*een nog ongebooren genie – een Darwin of Einstein van het bewustzijn – komt met een verbijsterend nieuw idee dat opeens alles helder maakt.*" (Pinker 2007).

**Ontologie.** Volgens mij is de sterkste reden om te denken dat het Harde Probleem onoplosbaar is, dat er geen natuurwetten zijn over mentale gebeurtenissen. De ontologische positie van mentale gebeurtenissen is onduidelijk (zie Voorwoord). Maar wetenschap gaat onvoorspelbaar voort. Immanuel Kant had in zijn tijd zeer goede redenen om te geloven dat er geen causale verklaring kon komen voor doelmatigheid en doelgerichtheid. Maar Kant had niet kunnen vermoeden dat er een evolutieer en cybernetica zou komen.

**Zijn bewuste processen relevant en grijpbaar?**

De discussie over neurale en mentale processen wordt al eeuwen gevoerd. Voor de meeste deelnemers eraan is niet het zomaar een vrijblijvende discussie over overigens neutrale filosofische concepten. Nee, eigenlijk gaat het over dierbare ideeën zoals wat wij mensen eigenlijk zijn, over de fysieke werkelijkheid, over de zin van het bestaan, over God enz. (hoofdstuk 6.3.). De discussie bleef niet altijd wetenschappelijk en beschaafd.

Er is voortdurend ongelooflijk veel neurale activiteit in de hersenen van de mensen en het overgrote deel daarvan kan men niet bewust ervaren. Vergeleken met die hersenactiviteit is het aantal bewuste gewaarwordingen per tijdseenheid verwaarloosbaar klein. Maar omdat bewuste belevingen alles zijn wat mensen kunnen ervaren, besteden allerlei onderzoekers en denkers onevenredig veel aandacht aan belevingen. Pogingen om belevingen empirisch te onderzoeken lopen steeds spaak. Je kunt mensen wel vragen wat ze beleven, maar dan onderzoek je hun uitspraken (= gedrag) en niet hun belevingen. In de oorspronkelijke proefopzet van Libet e.a. (1983) konden proefpersonen melden wanneer zij de eerste intentie voelden om te bewegen. Maar wanneer ze die beweging moesten afbreken, was het niet duidelijk wat er eigenlijk gebeurde. Ook als je proefpersonen bij een toontje vraagt of ze dachten te gaan bewegen, wordt het tijdstip van de eerste gemelde intentie ongrijpbaar: integere proefpersonen zeggen maar wat, met erg grote spreidingen. De resultaten van Libet en Ellis maken het onwaarschijnlijk dat mentale gebeurtenissen zoals intenties een grote invloed op gedrag hebben. Ik vrees dat 'objectieve' informatie over subjectieve belevingen onmogelijk is.

**Een wijze raad van Descartes**

Als wetenschapper kun je hooguit helderheid zoeken over fysieke en mentale gebeurtenissen; dat is wat ik hier geprobeerd heb. Als mens van vlees en bloed kun je hooguit proberen een aangenaam en goed leven te leiden. De 52-jarige René Descartes en de 20-jarige theologiestudent Frans Burman bespraken in 1648 in Egmond (Nederland) hun ideeën. Descartes merkte op: "*Let wel: je moet niet te lang stilstaan bij de Méditations of metafysische vragen, en deze niet uitdiepen in commentaren. [...] anders verwijderd je verstand zich te*

<sup>18</sup> Flanagan (1991) noemde dergelijke denkers '*New Mysterians*'.

*ver van fysieke en waarneembare dingen, en kan je deze niet meer bestuderen. En dat is toch de meest gewenste bezigheid voor mensen, omdat dit hen voordelen voor het leven oplevert.”* (uit Descartes en Bridoux 1953, p. 1381).

In mijn boek heb ik deze wijze raad van Descartes gevolgd. Brain/mind-vraagstukken kan je hooguit verhelderen, en je kan subjectieve keuzes maken. Maar toetsbare oplossingen zijn volgens mij niet haalbaar.



