

'Van geest naar materie'

maart 2008

'Het is alleen onze trots die grenzen legt waar er geen zijn.'

- Julien Offray de Lamettrie

"De mens een machine."

In de moderne wetenschap zijn er een aantal grote eindvragen... Nu we de 'simpele' problemen [schijnbaar] hebben opgelost, blijven er een paar hoge bergen staan waar we de top [schijnbaar] [nog] niet van hebben bereikt. Een van deze bergen heet **bewustzijn**, en hier zal gekeken worden in hoeverre deze berg te beklimmen, of misschien zelfs wel te temmen valt. Er zal begonnen worden bij de eerste grote denker op het gebied van de filosofie van de geest, die talloze problemen formuleerde. Uiteraard wordt René Descartes bedoeld. Daarna zal zijn visie worden afgezet tegen de 'stand' van zowel de moderne wetenschap alsook de filosofie van de geest, om vorderingen en grote problemen te kunnen herkennen.

In 1637 verscheen **René Descartes'** *Discours de la méthode*¹. Een klein boekje waarin hij een nieuwe methode [ook te benoemen als epistemologie] schetst. Hij vertelt aan welke regels hij zich zelf dwong te houden, op weg naar ware kennis. In het vierde deel staan de befaamde woorden *Ik denk, dus ik ben*, naar aanleiding waarvan hij vlak daarna ook meteen maar even het Bewijs voor het bestaan van god levert. Maar dat terzijde gelaten. Deze befaamde woorden bevatten de compacte definitie van zijn *dualisme*. Hij maakt een scherp onderscheid tussen lichaam [materie] en geest. De eerste is een *res extensa* [deelbaar en vooral *uitgebreid* ding], de tweede een *res cogito* [ondeelbaar en vooral *denkend* ding]. Hij pakt ook [elders] het probleem van het interactionisme aan, door te stellen dat de plek waar de ziel/geest contact met het lichaam maakt de *pijnappelklier* is. Hoe een immaterieel iets een materieel iets kan beïnvloeden legt hij echter nergens uit, waarschijnlijk omdat dat *per definitie* niet kan.

In het vijfde deel tracht hij aan te tonen hoezeer mensen van machines [wat hij dieren vond te zijn] verschillen. Hij 'bedacht' daarvoor een soort van voorloper van de *philosophers' zombie* [en tevens een aanpak – van buitenaf, met behulp van gedrag bestuderen – die op *behaviorisme* lijkt]. Een 'iets' dat qua uiterlijk niet van de mens te onderscheiden valt, maar qua innerlijk wél. Het eerste

1 R. Descartes, (2002) *Over de methode*, Amsterdam: Boom. 6^e herziene druk. (Vert. Th. Verbeek.)

verschil is dat deze wezens nooit '(...) woorden zouden kunnen gebruiken, of andere willekeurige tekens (...)'. Daarna stelt hij dat '(..) *zelfs* als deze machines bepaalde dingen even goed of wellicht beter zouden doen dan een mens, er *onherroepelijk andere dingen zijn waarin ze tekort schieten* (...)'

[Mijn cursivering –MG]. Deze tweede redenering lijkt sterk analoog te zijn aan het [AI effect](#). Dit houdt in dat men, zodra een computer in staat is gebleken een eerst voor exclusief menselijk gehouden taak te verrichten, intelligentie gaat herdefiniëren zodat de niet-zo-exclusief-menselijk gebleken taak daar buiten valt. Descartes doet dit voor het bewustzijn door te stellen dat als een machine uitingen-van-als-bewustzijn-herkenbaar-gedrag vertoont, dit geen teken van bewustzijn is, maar van de noodzaak de definitie van de attributen van bewustzijn enger te maken.

De eerstvolgende te bespreken persoon is de Fransman **Julien Offray de Lamettrie**, die in zijn in 1748 [onder een schuilnaam] uitgegeven *L'homme machine*² Descartes bestrijdt. En op zo'n manier bestrijdt, dat te stellen valt dat hij het dualisme toen al zodanig ontkracht heeft, en het materialisme zodanig bepleit, dat het hele dualisme-debat van toen af aan als een *non-issue* gezien mag worden.

Hij kritiseert allereerst de theologen – hoe zouden schriftgeleerden iets relevant kunnen zeggen over zaken die het lichaam en de geest en niet kennis van de bijbel aangaan? Lamettrie zoekt het liever bij de kenners – de '*artsen die tegelijk filosoof zijn*'!³ Men moet van dichtbij kijken en geen aristotelische leunstoeldenker zijn die zijn handen niet vies wil maken. Vervolgens bespreekt hij [met veel empirische voorbeelden] achtereenvolgens de effecten van ziekten, slaap, bloedsomloop, verdovende middelen [opium en koffie ed.], de maaltijd, ouderdom en het klimaat op het lichaam en dwingt zo erkenning van de afhankelijkheid van de geest van het lichaam af, en vervolgens de identiekheid van die twee, met de ruim tweeënehalve eeuw op de wetenschap vooruitlopende woorden: '(...) daar de verschillende vermogens van de ziel zozeer afhankelijk zijn van de specifieke organisatie van de hersenen en het lichaam in het algemeen, dat zij schijnbaar alleen maar die organisatie *zijn*, hebben we kennelijk te maken met een zeer verlichte machine.'

(sic!)

Met dit betoog kwam een einde aan het dualisme en bestonden er [in relevante zin] voortaan alleen nog maar monismen, noem het fysicalisme, noem het materialisme, het komt neer op het afwijzen van hogere, ongrijpbare entiteiten [ziel, geest, metafysische noties of wat dan ook] ten behoeve van het standpunt dat alles wat 'is' materieel is, en dat alles wat 'is' verklaard kan worden

² Gebruikt: De Lamettrie, *De mens een machine*, Boom (Meppel/Amsterdam 1978). Vert. H. W. Bakx.

³ Leuk om te zien is dat de mensen die nu aan het filosofie-van-de-geest-debat mee willen doen zich wel *moeten* verdiepen in neurobiologie, (evolutie-)psychologie, en andere relevante vakgebieden, willen ze mee kunnen praten.

door middel van beschrijvingen van materiele zaken/processen, zonder toevlucht te nemen tot het poneren van verklaring-ontduikende niet-materiële substanties.

De intellectuele beweging van dualisme naar materialisme werd natuurlijk ook ondersteund door de gedachtegangen van onder andere Spinoza [*deus sive natura* - *pantheïsme*, oftewel de *identity-theory* voor god i.p.v. voor het bewustzijn] en Occam [*Occam's razor*- het maken van zo min mogelijk aannamen m.b.t. een theorie]. Hoe dan ook, voor elk weldenkend mens was het materialisme halverwege de 18^e eeuw een feit.

Een goede methode om de werking van een ingewikkeld apparaat te bestuderen is door te kijken naar een defecte versie ervan. Bij mensen is het natuurlijk niet ethisch verantwoord om expres schade toe te brengen, en daarom hebben wetenschappers altijd moeten 'wachten' op ongelukken die hen een richting op zouden kunnen sturen. Eén zo'n bekend ongeluk [en dus goede onderzoeksmogelijkheid] is het geval van de 19^e-eeuwse spoorwegaarbeider Phineas Gage die een lange stalen staaf door het voorste gedeelte van zijn hersenen geblazen kreeg. Dat gaf inzicht in de functie van dat gedeelte van de hersenen. Aangaande Gage werden ingrijpende persoonlijkheidsveranderingen gerapporteerd.⁴ Dit bracht veel schot in de neurologische zaak en is een van de eerste gevallen van aftasting van de mogelijkheid van de reductie/gelijkstelling van mentale- tot/met fysische toestanden, wat nu wordt omschreven als NCC's – *neural correlates of consciousness*.

De verdere geschiedenis van de filosofie van de geest laat daarna de bloei van een aantal indrukwekkende filosofische stromingen zien, die elk op elkaar voortbouwen.

Eerst het in de psychologie toonaangevende *behaviorisme* onder aanvoer van **Skinner** en **Watson**, dat de geest radicaal afwees, en zich toeleegde op de bestudering van observeerbaar gedrag, wat neerkwam op het toedienen van stimuli om te zien welke acties dit tot gevolg had, elk tussenliggend proces negerend. De beperkingen hiervan werden snel duidelijk en een uitbreiding van deze theorie was het *functionalisme*, dat wél enige affiniteit met tussen stimuli en activiteit liggende processen had. Deze stroming typeerde mentale staten als in [causaal] verband staande met stimuli, gedrag én [als uitbreiding op het behaviorisme] andere mentale staten.

Ook dit bleek echter te ongenueerd. Filosofen van de geest moesten de toevlucht nemen tot een multidisciplinaire aanpak. In verschillende wetenschapsgebieden wordt nu uitgebreid [en vruchtbaar] onderzoek naar het bewustzijn gedaan.

Programmeurs proberen kunstmatig intelligente modellen op te stellen die specifieke

⁴ Recentere studies proberen de vermeende ernst van dit geval te nuanceren aangezien dit klassieke verhaal bijna mythische proporties aangenomen heeft. Zie: *Trouw Dossier nr 20* - Laan., M v/d, *Hersengeheimen* (Amsterdam 2002) [Rainbowpocket]; Damasio, A.R. (1994) *Descartes' error*. London: Picador.

bewustzijns capaciteiten na moeten bootsen – soms door een systeem te maken dat gebaseerd is op hersenonderzoek, soms door een aanpak te hanteren die pragmatisch is en het systeem simpelweg hetzelfde resultaat probeert te laten verkrijgen. Soms gaat dit erg goed, zoals bijvoorbeeld bij de nabootsing van geheugen en rekenkracht of schaken, soms ook erg moeilijk zoals bij spraak- of gezichtsherkenning [alhoewel **PDP** grote capaciteiten insinueert, en dit is [jawel] een redelijk accurate alhoewel gesimplificeerde nabootsing van de werking van de neuronen in de menselijke hersenen]. Elke geslaagde poging is een overwinning in die zin dat het bewijst dat er voor die specifieke capaciteit geen hogere entiteiten of toestanden nodig zijn, behalve materiele.

Neurologen maken op verschillende manieren hersenscans terwijl mensen bepaalde dingen denken of handelingen verrichten, en proberen zo af te leiden welke hersenfuncties waar zitten.

Psychologen voeren tests uit die de capaciteiten en beperkingen van de menselijke geest laten zien. Belangrijke bevindingen zijn die van **blindzicht** en **change blindness**, die even kort behandeld zullen worden.

Van *blindzicht* is sprake als iemand claimt in een bepaald gebied van het gezichtsveld niets meer te kunnen zien, terwijl diegene, wanneer gevraagd te gokken wat er daar te zien is, ver boven gemiddeld goed gokt. Dit toont aan dat er meerdere niveaus van bewustzijn zijn. Blindzichters zijn zich er dus niet van bewust bepaalde informatie in 'handen' te hebben, terwijl zij [maar niet hun 'ik'] daar wél gebruik van maken, zij het onder-[zelf-]bewust.

Om *change blindness* uit te kunnen leggen moet eerst worden verteld over *saccades*. Saccades zijn de minieme en vliegensvlugge [20 á 200 milliseconden durende] sprongetjes die je oog elke seconde waagt. Iedereen heeft de sterke intuïtie de wereld te onthouden zodra hij zijn ogen sluit; iedereen denkt een scherpe en gedetailleerde representatie van de wereld in zijn hoofd te bewaren, maar dit blijkt zeer onaannemelijk. Er zijn experimenten verricht waarbij een plaatje of tekst waarnaar gekeken werd *tijdens een saccade* veranderd werd [een kerktoren verdwijnt of verschuift 1/3 van de breedte van de foto]. Niemand is in staat dit op te merken.

De bevinding van *change blindness* impliceert mijns inziens dat de *qualia* hier een slag verloren hebben. Hier raakt het onderzoek het spanningsveld tussen *verklaren* en *wegverklaren*. Als wordt aangenomen dat wij geen gedetailleerde representatie van de wereld om ons heen bewaren, scheelt dit aanzienlijk in de denk-, reken- en geheugen-capaciteiten die wij de hersenen anders wél zouden moeten hebben toegedicht. Bovendien zou de daarmee verbonden **quale** [intrinsieke en onoverdraagbare subjectief ervaren kwaliteiten verbonden aan stimuli] dientengevolge niet bestaan en dus ook niet verklaard hoeven worden. Maar is die quale dan verklaard of weg-verklaard? Het standpunt van ondergetekende in deze is dat dat eigenlijk geen relevante vraag is, of dat die vraag in

ieder geval niets aan *dit debat* toevoegt. De beantwoording van de vraag welke van de twee opties nou logisch geldig is, neemt niet weg dat een ervaring, een subjectieve toestand, door de wetenschap te vatten is gebleken. Het lijkt logisch dat dit de komende tijd met veel andere [subjectieve] verschijnselen ook zal gebeuren. Chirurgen die bepaalde elektrische stroompjes de hersenen in sturen kunnen de onderhavige persoon levendige ervaringen bezorgen, chemicaliën kunnen gemoedstoestand, redelijkheid en perceptie veranderen, bepaalde hersengebieden blijken sterke causale verbanden te vertonen met spraak, gezichtsherkenning en taalgebruik. Sterker nog - neurowetenschappers kunnen aflezen of iemand aan een *gezicht* of aan een *plaats* denkt.⁵

Ondertussen zijn wij sterk genoeg bewapend om terug te keren naar de drie door Descartes geschetste problemen.

Ten eerste maakte Descartes een onderscheid tussen een deelbaar en *uitgebreid res extensa*, en een ondeelbaar en *denkend res cogito*. Dit dualisme blijkt niet houdbaar. Niet alleen is het principieel onmogelijk dat een immaterieel iets een materieel iets beïnvloedt [noch andersom], het is daarnaast ook nog eens gebleken dat wat Descartes het *res cogito* noemde deelbaar is, hersenaandoeningen kunnen selectieve denkcapaciteiten [lange- of korte-termijngeheugen, gezichtsherkenning, benoemingsvermogen] aantasten of onmogelijk maken.

Ten tweede het gebruik van taal of een ander tekensysteem. Nagegeven moet worden dat Descartes deze bewustzijnsvoorwaarde niet al te scherp heeft gedefinieerd. Echter, computers blijken in staat tot spraakherkenning [= ontcijferen van input – reuma-patienten gebruiken dit om e.g. hun PC te besturen, maar ook mijn telefoon belt bepaalde personen nadat ik de goede naam duidelijk uitgesproken heb], én in staat tot het talig rapporteren. [SHRDLU](#) is in staat zijn blokkenwereld m.b.v. visuele input te herkennen en te beschrijven. En natuurlijk zijn programma's nog veel beter in staat *machinetaal* te begrijpen. Zij het langzaam, zij het knuddig, zij het grof – machines blijken, *contra* Descartes, woorden zowel als in- én output te kunnen gebruiken.

Descartes derde en laatste bezwaar, dat als analoog aan het *AI-effect* benoemd is, zal ook nog even besproken worden. Het 'onherroepelijke' feit dat Descartes aandraagt, dat als het terrein van de kunstmatige intelligentie⁶ met bewustzijn verbonden capaciteiten zou kunnen 'bouwen', er andere capaciteiten over zouden blijven, lijkt ten eerste niet principieel onherroepelijk, omdat de geschiedenis een trend laat zien naar het kunnen nabootsen van steeds méér vroeger voor inherent menselijk en onnabootsbaar gehouden capaciteiten [*Edgar Allan Poe* noemde het een belachelijk en

5 Pinker, S, (2002) *The blank slate*, Viking: Harmondsworth, p. 41-2.

6 Eigenlijk beter te noemen 'synthetische intelligentie', omdat 'kunstmatig' associaties heeft met 'nep' en 'schijn', terwijl computerprogramma's wel degelijk kunnen schaken [mbv. heuristisch die ondergebracht is in algoritmen], en mensen op de maan hebben gezet (Zie: Haugeland, J. *Artificial intelligence: The very idea*. Cambridge: MIT Press.

absurd idee dat een machine ooit zou kunnen schaken]. Waarom zou dat proces ineens stoppen? Als dat zo is merken we het pas als we het merken, het is voorbarig te zeggen dat je weet hoe hoog de berg is als er nog mist om de top heen hangt.

Deze bespreking van de filosofie van de geest is ietwat verschoven naar kunstmatige intelligentie, maar dat heeft natuurlijk alles te maken met Descartes' uitdagingen, die allemaal aangenomen en ontkracht lijken te zijn. Hier wordt niet betoogd dat *machine-bewustzijn* mogelijk is, dat had een heel andere opbouw/aanpak gevergd. Nee, wat aangetoond is, is ten eerste dat het dualisme vermoord is door de executie van Lamettrie; ten tweede dat Descartes' verdere standpunten aangaande de mogelijke capaciteiten van machines voor een groot deel achterhaald zijn gebleken; en ten derde *niét* dat de mist rond de top van de berg die **bewustzijn** heet is opgeklaard, dat de top in zicht is - maar wél, dat we, als we niet weten hoe ver het is, één ding zeker weten, namelijk dat we, als we naar beneden kijken, duidelijk kunnen zien dat we al een hele klim hebben gemaakt!