**Bestaat er een digitale ziel?**

**Door James Hoskins**

n onze westerse cultuur neemt de overtuiging steeds meer toe dat computers uiteindelijk bewust zullen worden. In de afgelopen drie jaar alleen al hebben mensen vanuit verschillende achtergronden van het maatschappelijke spectrum zowel de angst als de hoop uitgesproken dat dit zal gebeuren. Wetenschappers als Stephen Hawking, Elon Musk en Bill Gates zijn allemaal verontrust over het feit dat kunstmatige intelligentie (AI) een bedreiging vormt voor de veiligheid van de mensheid (1).

Een presbyteriaanse predikant in Florida zei dat hij van plan is om het evangelie met de robots te delen als ze eenmaal wakker zijn (2). In de literatuur en in de filmwereld verkende men de interessante vraag of we verliefd kunnen worden op onze nieuwe machine-superieuren zoals in de films: Ex Machina, Transcendentie en HAAR.

Rondom het mogelijke bestaan van A.I. rijzen ook geestelijke, wereldbeschouwelijke vragen. Is het bijvoorbeeld mogelijk dat een computer bewust wordt? Als dit het geval is, zou die dan een ziel hebben? Zo niet, wat betekent dat dan voor onze ziel? Zou het op een of andere manier weerleggen dat we een ziel hebben? Zou dat dan de waarheid van de Bijbel bestrijden? Enzovoorts. Veel christenen zullen zich waarschijnlijk afvragen wat ze allemaal zoal moeten geloven.

Het goede nieuws is dat er de vrijheid is om het met elkaar oneens te zijn. Vanuit de christelijke wereldbeschouwing denkt men verschillend over aspecten van A.I.. Of iemand denkt dat computers zelfbewust kunnen zijn, hangt uiteindelijk af van welke van de twee populaire visies op de ziel en het menselijk bewustzijn hij of zij aanhangt: het niet-reductionistisch fysicalisme of het substantie-dualisme. Voordat ik beide visies toelicht, moet er eerst een belangrijk onderscheid worden gemaakt.

**Het verschil tussen ‘smart’ en ‘bewust’**

Veel mensen stellen het intelligentieniveau van een computer ten onrechte gelijk aan bewustzijn, maar dit zijn twee verschillende dingen. Een computer kan potentieel oneindig 'slim' of 'intelligent' zijn als men daarmee bedoelt dat hij een ​​onbegrensd aantal berekeningen met onbegrijpelijke snelheden kan maken. Maar toch is dit technisch gezien niet hetzelfde als bewustzijn. Bewustzijn omvat veel meer dan alleen een kwantitatieve meting van intelligentie. Het omvat kwalitatieve ervaringen zoals subjectief bewustzijn, begrip, intentionaliteit en een ‘holistische’ identiteit. Zelfs als een computer intelligent genoeg zou zijn om het te doen voorkomen dat het kwalitatieve ervaringen had, bijvoorbeeld dat het leek alsof het pijn had of verliefd was, dan weten we nog steeds niet zeker of het echt bewust is. Een beroemd gedachte-experiment laat zien waarom.

**De Chinese kamer**

De filosoof John Searle vroeg aan mensen zich het volgende voor te stellen. Stel, iemand die geen Chinees kent en wiens moedertaal Engels is, bevindt zich in een kamer met alleen dozen vol met Chinese karakters en een reeks Engelse instructies (3).Mensen buiten de kamer sturen kaarten naar binnen met Chinese symbolen erop, die eigenlijk vragen zijn zonder dat de persoon binnen de kamer weet dat het vragen zijn. Door de Engelse instructies te volgen, kan de persoon Chinese symbolen terugsturen die de juiste antwoorden op de vragen zijn. Mensen buiten de kamer zouden kunnen denken dat de persoon binnen de kamer goed Chinees spreekt. Maar de realiteit is dat de persoon in de kamer het helemaal niet begrijpt. Hij heeft geen idee wat de karakters betekenen. Hij volgt gewoon de juiste instructies om de juiste output te geven.

Het punt van Searle is dat dit precies is wat een computer doet. Het volgt een reeks instructies (een programma) om de juiste output te geven, ongeacht hoe snel of efficiënt deze in staat is om dit te doen, ongeacht hoe natuurlijk of knap zijn reacties op ons lijken. We kunnen logischerwijs niet concluderen dat de computer begrip of bewustzijn heeft. Het intelligentieniveau van een computer kan misschien de *illusie*van bewustzijn geven, maar het kan nooit het echte bewustzijn garanderen. Sorry, Siri.

**Is het onderscheid belangrijk?**

Sommigen zullen misschien denken dat het voorbeeld van de Chinese kamer te veel denkt te bewijzen. Als we het gedachte-experiment serieus nemen, zou iemand kunnen zeggen, dan kunnen we ook niet zeker weten dat andere mensen zich ook bewust zijn, omdat de enige manier waarop we in staat zijn om het bewustzijn van andere mensen te detecteren, is door te kijken naar hun output, dus naar hoe ze zich gedragen in reacties op hun omgeving. Maar we nemen aan dat andere mensen de hele tijdbewust zijn; dus waarom dan geen computer?  Als een computer effectief de natuurlijke reacties van een mens in elke situatie kon nabootsen, dan zouden we moeten aannemen dat het ook bewustzijn heeft, zelfs als het een programma zou volgen; er is gewoon geen andere manier om het na te gaan.

Dit is in essentie het argument dat Alan Turing gaf bij het verdedigen van wat later de 'Turing-test' zou worden genoemd (4).

Wat Turing en anderen echter niet erkennen, is dat output (gedrag) niet onze enige indicator is om te weten of andere mensen bewustzijn hebben. We hebben als mensen ook introspectie. Als menselijke wezens heeft ieder van ons een soort van derde persoon, een innerlijk bewustzijn, een objectieve ’getuige’, een ’virtuele ander’ die beoordelingen maakt vanuit onze eigen individuele stroom van ervaringen. Ik weet bijvoorbeeld dat ik op dit moment honger heb. Ik ben me ook bewust (getuige) van mijn groeiende prikkelbaarheid en ik leid daaruit af (oordeel) dat dit waarschijnlijk komt vanwege mijn lege maag. Dus door introspectie krijgen we onbetwistbare kennis van ons eigen bewustzijn (5).Maar we krijgen veel meer dan dat. We weten ook dat we organisch afstammen van andere mensen die beweren hetzelfde bewustzijn te hebben als wij. Er is een continuïteit van oorsprong, ervaring, organische materie en vorm die we delen met andere mensen die we niet delen met computers of robots. Deze continuïteit stelt ons in staat om de kennis van ons eigen individuele bewustzijn uit te breiden en zelfverzekerd te concluderen dat andere mensen ook bewust zijn. Dat betekent echter niet dat we tot diezelfde conclusies bij computers kunnen komen.

**Wat is wijsheid?**

Het onderscheid tussen intelligentie en bewustzijn is belangrijk omdat het christenen, die mogelijk theologische bezwaren hebben tegen het idee van een **bewuste**computer, toestaat om toch acht te slaan op de wijsheid van technologen zoals Bill Gates en Elon Musk over de potentiële **intelligentie**-dreiging van A.I. De waarheid is namelijk dat er werkelijk een groot gevaar in zou kunnen schuilen als we te veel controle over menselijke activiteiten geven aan A.I. Neem bijvoorbeeld de ’beurs-crash’ van 2010 (6).Zelfs als echt computer**bewustzijn**onmogelijk is, moet computerintelligentie, waarover we het allemaal eens zijn, wat betreft de toepassing ervan zeer zorgvuldig overwogen worden.

**Twee visies over menselijk bewustzijn**.

Met het onderscheid tussen intelligentie en bewustzijn in gedachten, kunnen we nu beoordelen hoe iemands kijk op de ziel en het menselijk bewustzijn van invloed is op wat hij of zij denkt over de mogelijkheid van computerbewustzijn.  Allereerst is er het fysicalisme, dit is een filosofische veronderstelling die beweert dat alles fysisch is en dat immateriële eigenschappen (zaken van biologische, psychologische, morele of sociale aard) supervenieren (12)op het fysische. Het fysicalisme staat voor de overtuiging dat al wat bestaat, in diepste wezen materie is bepaald door de universele natuurwetten. In het fysicalisme merkt men alle objecten en eigenschappen aan als door de natuurwetenschap beschrijfbaar. Deze stellingname is uitgewerkt in een 20ste-eeuwse wetenschapsfilosofie en filosofie van de geest. (Wikipedia)

Hoewel de theorie van het fysicalisme meer bekend is als heterodox (dit betekent: afwijkend van de gangbare overtuiging van het leerstellig geloof) dan als orthodox, hangen oprechte christenen beide standpunten aan, die ik hieronder zal bespreken. Houd in gedachten dat deze twee weergaven de meest voorkomende zijn, maar dat ze zeker niet de enige opties zijn. Er zijn andere, meer genuanceerde meningen waar we in dit artikel niet op in kunnen gaan (7).

**Onsplitsbaar fysicalisme**

Volgens onsplitsbaar (niet reductief) fysicalisme (ook wel 'eigendoms devitalisme' (13) genoemd) is het menselijk bewustzijn echt, dus geen illusie, maar is het volledig afhankelijk van, en uiteindelijk veroorzaakt door fysieke gebeurtenissen in de hersenen

Volgens deze visie is bewustzijn geen deel van onze ziel, en bestaat er niet zoiets als een ziel: we *zijn*eenvoudigweg ons lichaam en zijn eigenschappen, niets meer dan dat. Niet-reductieve fysici zien het bewustzijn als een eigenschap van materie die tevoorschijn komt wanneer een fysiek systeem (zoals een menselijk brein of lichaam) samengesteld is en op een bepaalde manier werkt.

Christenen die deze opvatting hebben, staan ​​wellicht het meest open voor de mogelijkheid van computerbewustzijn. Immers, als menselijk bewustzijn gewoon een eigenschap is van een fysiek brein, wie zegt dan dat we geen kunstmatig brein kunnen bouwen dat dezelfde emergente (14) eigenschap (eigenschap die spontaan optreedt)) zou kunnen produceren? Christof Koch, een van 's werelds meest toonaangevende neurowetenschappers en onderzoekers van het bewustzijn, denkt dat dat heel goed mogelijk is. Hij legt uit: "Bewustzijn is een eigenschap van complexe systemen met een bepaald 'oorzaak-gevolg' werkingsmechanisme. Ze hebben een specifieke manier om met de wereld om te gaan, zoals de hersenen, of in principe, zoals een computer dat zou kunnen. Als je een computer zou bouwen die dezelfde circuits heeft als de hersenen, zou deze computer ook een bewustzijn hebben dat ermee verbonden is. Het zou voelen als iets om deze computer te zijn”(8).

Niet-reductionistisch fysicalisme is een aantrekkelijke optie voor veel wetenschappelijk ingestelde christenen maar stuit niettemin op een aantal theologische en filosofische bezwaren. Want als er niet zoiets bestaat als een ziel en ik ben niet meer dan mijn lichaam en zijn eigenschappen, dan houd ik letterlijk op te bestaan ​​als ik sterf. De dood, bezien vanuit een fysiek gezichtspunt, is niet minder dan vernietiging, omdat mijn lichaam vergaat en het bewustzijn dat ermee verbonden is, uitdooft. Hoe zou ik **mezelf**kunnen zijn in de opstanding met niets om mijn identiteit na de dood te behouden of te veranderen? Zelfs als God alle exacte atomen zou assembleren die mijn lichaam hadden samengesteld toen ik stierf, zou het niet ervoor zorgen dat het herrezen "ik" meer was dan een verbeterde kopie of replica.

Om dit punt te illustreren, stelt de fysicus Peter van Inwagen het volgende denkbeeldige gesprek voor: "Is dat het huis dat je dochter vanmorgen met enkele blokken gebouwd heeft?" "Nee, ik heb dit opnieuw gebouwd nadat ik het per ongeluk had omgestoten. Ik heb echter alle blokken precies zo geplaatst zoals zij gedaan had. Vertel het haar maar niet“ (9).In zowel het voorbeeld van de blokken als de opgestane persoon, betoogt Van Inwagen, is het nieuw gemaakte slechts een *replica*van het origineel, dus niet *hetzelfde*als het origineel.

**Substantie-dualisme**

Volgens het substantie-dualisme zijn mensen een eenheid van twee verschillende ‘substanties’: lichaam en ziel. Bewustzijn is een eigenschap van onze ziel, die los van het lichaam na de dood kan bestaan. Aldus vermijdt substantie-dualisme met succes het identiteitsprobleem van het fysicalisme dat eerder werd genoemd. Ik kan nog steeds mezelf zijn in de opstanding, omdat mijn ziel mijn identiteit na de dood behoudt en bepaalt.

Christenen die deze opvatting aanhangen, zullen waarschijnlijk het minst de mogelijkheid van computerbewustzijn erkennen. Immers, als bewustzijn een eigenschap van de ziel is, die alleen door God wordt gegeven, dan zal een computer, ongeacht hoe intelligent deze zal kunnen worden gemaakt, nooit een ziel hebben en daarom nooit bewust zijn. Interessant is dat Alan Turing in 1950 probeerde te reageren op deze opvatting:

“De meeste mensen erkennen dat God bepaalde dingen niet kan doen, zoals het getal één gelijk aan twee maken, maar moeten we niet geloven dat Hij de vrijheid heeft om een ​​ziel aan een olifant te geven als Hij dat gepast acht? En precies hetzelfde argument kan gemaakt worden in het geval van machines. In het maken van machines moeten we niet zo oneerbiedig zijn ons aan te matigen dat we Zijn kracht hebben om zielen te maken, Niet meer dan bij de voortplanting van kinderen. In beide gevallen zijn we eerder instrumenten van Zijn wil die de huisvesting bieden voor de zielen die Hij creëert” (10).

Het betoog van Turing is in wezen dit: God kan een ziel geven aan wie Hij wil, en Hij doet dit voor nieuwe wezens wiens lichamen we door voortplanting creëren, namelijk onze menselijke nakomelingen. Waarom kon God dan ook geen ziel schenken aan een intelligente computer die we creëren? Maar het betoog van Turing gaat ervan uit dat alle dualisten geloven dat God zielen ex nihilo (iets scheppen uit het niets) schept en ze een ziel bij de conceptie verleent. Echter niet alle dualisten geloven dat.

Want veel christenen geloven dat zielen, hoewel ze echt en te onderscheiden zijn van het lichaam, toch organisch worden overgedragen via voortplanting, samen met het lichaam - een visie die 'traducianisme' (15) wordt genoemd. Volgens deze visie reproduceren mensen zich als gehele wezens - lichaam en ziel. Daarom is de enige manier voor het ontstaan van een nieuwe ​​ziel om deze organisch (via voortplanting) te laten afstammen van wezens (mensen) die ook een ziel hebben. Voor een traducianist is de continuïteit van de menselijke soort waarover ik eerder sprak een continuïteit van oorsprong, ervaring, organische materie en vorm, en dus de sleutel tot de ‘productie’ van nieuwe zielen.

**Vrijheid om het oneens te zijn**

De kwestie van computerbewustzijn doet er misschien niet toe op praktisch niveau, althans nog niet. Computers zijn al intelligent genoeg om ons voorzichtig te doen zijn in hoe we ze gebruiken; je hoeft niet in computerbewustzijn te geloven om dat toe te geven. Of computers kwalitatieve ervaringen kunnen hebben, hangt grotendeels af van onze filosofische vooronderstellingen over zielen. Ook al kunnen we het oneens zijn, toch kunnen we samenwerken aan een betere toekomst. Zal een computer ooit intelligent genoeg worden om volledig bewust te worden? Ik denk van niet. Moeten robots gered worden? Nee, maar het kan geen kwaad om de kwaliteit van de Pascal (16) te verbeteren

**James Hoskins** is een leraar, schrijver en muzikant uit Kansas City, Missouri. Hij geeft filosofie- en wetenschapslessen aan een voorbereidende hogeschool. Hij schrijft over de kruising van rede, geloof en cultuur op zijn blog PhiloLogos.net, evenals op christandpopculture.com.

Dit artikel verscheen eerder in CHRISTIAN RESEARCH JOURNAL, volume 39, nummer 02 (2016).

Vertaald door **Gerard Feller**

Literatuur

1  Michael Sainato, “Stephen Hawking, Elon Musk, and Bill Gates Warn about Artificial Intelligence,” *The Observer*, August 19, 2015, [**http://observer.com/2015/08/stephen-hawkingelon-musk-and-bill-**](http://observer.com/2015/08/stephen-hawkingelon-musk-and-bill-) gates-warn-about-artificial-intelligence/.

2  Anthony Cuthbertson, “Florida Pastor Plans to Convert Robots to Christianity,” *International Business Times*, February 6, 2015, [**http://www.ibtimes.co.uk/florida-reverendchristopher-benek-wants-convert-**](http://www.ibtimes.co.uk/florida-reverendchristopher-benek-wants-convert-) artificial-intelligence-christianity-1486912.

3  For a detailed overview, see David Cole, “The Chinese Room Argument,” *The Stanford Encyclopdia of Philosophy*(winter 2015 edition), ed. Edward N. Zalta, [**http://plato.stanford.edu/entries/chinese-room/.**](http://plato.stanford.edu/entries/chinese-room/)

4  Alan M. Turing, “Computing Machinery and Intelligence,” *Mind*59, 236 (October 1950): 433–60.

5  It is undeniable because the very act of trying to deny our own consciousness would require the use of our consciousness.

6  Ben Rooney, “Trading Program Sparked May ‘Flash Crash,’” *CNN Money*, October 1, 2010, [**http://money.cnn.com/2010/10/01/markets/SEC\_CFTC\_flash\_crash/**](http://money.cnn.com/2010/10/01/markets/SEC_CFTC_flash_crash/)

7  Such as hylemorphic dualism.

8  Antonio Regalado, “What It Will Take for Computers to Be Conscious,” *MIT Technology Review*, October 2, 2014, [**http://www.technologyreview.com/news/531146/what-it-will-takefor-computers-to-be-conscious/.**](http://www.technologyreview.com/news/531146/what-it-will-takefor-computers-to-be-conscious/)

9  Peter van Inwagen, “The Possibility of Resurrection,” *International Journal for Philosophy of Religion*9, no. 2 (1978): 114–21.

* Alan M. Turing, “Computing Machinery and Intelligence,” *Mind*59, 236 (October 1950): 443.
* [**https://nl.wikipedia.org/wiki/Fysicalisme**](https://nl.wikipedia.org/wiki/Fysicalisme)
* Supervenieren:Als in een object eigenschap A optreedt en noodzakelijkerwijs hierdoor ook eigenschap B dan supervenieert eigenschap B op A. Als je het object één op één reproduceert dan heeft het object ook eigenschap B. Met name relevant bij een aantal posities binnen de filosofie van de geest
* Het vitalisme is de leer van het bestaan van een levenskracht in de organische natuur . Devitalisme is de ontkenning daarvan.
* **Emergentie (wikipedia)**is een begrip dat met name centraal staat in de systeemtheorie en de wetenschapsfilosofie. Het betreft vaak de ontwikkeling van complexe georganiseerde systemen, die bepaalde eigenschappen vertonen die niet herleidbaar zijn door louter een reductie van hun samenstellende delen. Door interactie ontstaan eigenschappen, patronen, regelmatigheden en/of geheel nieuwe entiteiten. Een mogelijke definitie is: Een emergent verschijnsel ontstaat door interactie van factoren (die van alles zijn kunnen zijn: objecten, golven, krachten, ideeën) waardoor nieuwe entiteiten (objecten, patronen, reacties, inzichten) ontstaan die kenmerken/eigenschappen hebben die niet herleidbaar zijn uit eigenschappen van samenstellende/onderliggende factoren/bouwstenen.
* **Traducianisme**(wikipedia) is een concept uit de Christelijke theologie , wat gaat over de oorsprong van de menselijke ziel dan wel geest. Traducianisme stelt dat de ziel wordt gegenereerd door de natuur; dit betekent dat bij het ontstaan van een mens , de ziel van de mens zijn oorsprong vindt in de ziel van de vader en moeder. Vanuit het creationisme  bekeken betekent dit dat alleen de ziel van de eerste mens, Adam , door God werd geschapen. Eva, zijn vrouw, werd door God geschapen uit een rib van Adam (Genesis 2:21) en daarmee uit hem: 'mijn eigen gebeente, mijn eigen vlees (..) uit een man gebouwd.' (Genesis 2:23)
* De**PASCAL** was een door Philips  gebouwde computer. De afkorting PASCAL staat voor ***P****hilips****A****kelig****S****nelle****CAL****culator*. De PASCAL was de opvolger van Peter (*Philips Experimentele Tweetallige Electronische Rekenmachine*). Hij werd gebouwd door het Natuurkundig Labaratorium van Philips, dat hem ook in gebruik nam voor eigen rekentaken. Er werd onder andere gerekend aan heteluchtmotoren, televisies en halfgeleiders. PASCAL behoorde in zijn tijd tot de snelste computers ter wereld.In 1960 werd PASCAL opgeleverd. Hij is tot en met 1972 bij het NatLab in gebruik geweest