

De intieme en ondoordringbare belevingswereld

4.1 Andere geesten en conceptueel scepticisme

Stel je voor: op verzoek van een op het eerste gezicht zeer betrouwbare wetenschapper neem je deel aan een experiment waarin je moet gaan 'chatten' met ene ELIZA. Van meet af aan lijkt het te klikken tussen jou en ELIZA. Ze luistert naar je en stelt goede vragen. Ze lijkt je te begrijpen. En in korte tijd heb je het met haar over persoonlijke, intieme zaken.

Dan komt de wetenschapper langs die het experiment heeft opgezet. Zonder iets te vragen, begint hij over je schouder mee te kijken naar je chatsessie met ELIZA. Je bent enigszins verontwaardigd over het feit dat hij zo maar 'inbreekt' in jullie gesprek en vraagt hem wat nu precies de bedoeling is. Tot je ontzetting antwoordt hij dat je je niet moet opwinden: ELIZA is geen vrouw van vlees en bloed, maar een computerprogramma dat hij heeft ontwikkeld.

Zo moet het ongeveer gelopen zijn met de proefpersonen – of misschien kan je wel zeggen: de slachtoffers – van de experimenten van Joseph Weizenbaum. Tussen 1964 en 1966 ontwikkelde Weizenbaum aan het vermaarde Massachusetts Institute of Technology (MIT) het computerprogramma ELIZA, vernoemd naar het personage Eliza Doolittle uit *My Fair Lady*.^{*} Proefpersonen konden via een toetsenbord en een scherm met ELIZA 'converseren'. Er werd hun niet verteld dat hun gesprekspartner geen mens van vlees en bloed was, maar een computerprogramma.

ELIZA was geprogrammeerd naar het model van psychotherapeutische gesprekken. In een dergelijke context is het vaak zo dat de patiënt de respons van de psychiater erg welwillend interpreteert. De patiënt gaat ervan uit dat de psychiater dingen weet en begrijpt en dat zijn of haar verbale reacties zeer betekenisvol zijn. ELIZA exploiteert en parodieert dat gegeven, doordat Weizenbaum de computer enkele relatief eenvoudige instructies heeft gegeven waarbij 'ingehaakt' lijkt te worden op wat de 'gesprekspartner' te melden heeft.

Volgens Weizenbaum was die illusie bij veel gesprekspartners van ELIZA erg sterk. Zij raakten dusdanig emotioneel betrokken dat ze bijvoor-

^{*} In *My Fair Lady* is Eliza een arme bloemenverkoopster die plat Londens spreekt en door een professor intensief 'getraind' wordt om zeer voornaam te spreken en zich voor te doen als iemand van een hogere klasse. De musical is gebaseerd op *Pygmalion* van G.B. Shaw.

* Zie Weizenbaum (1976).

beeld weigerden Weizenbaum de transcripties van het gesprek te laten lezen, omdat dat hun privacy zou schenden. Anderen vroegen Weizenbaum de kamer te verlaten tijdens hun sessie met ELIZA. Het behoeft geen betoog dat de onthulling van ELIZA's ware aard koud op het dak viel bij haar gesprekspartners.*

```

=====
EEEEEEEE L      IIIIII ZZZZZZZ      AAA
E      L      I      Z      A      A
E      L      I      Z      A      A
EEEEEE L      I      Z      A      A
E      L      I      Z      A      A
EEEEEEEE LLLLLLL IIIIII ZZZZZZ      A      A
=====
ELIZA > (PLEASE TYPE IN ALL CAPS) WHAT'S YOUR NAME DEAR ?
AMIT
ELIZA > HI, I'M ELIZA, WHAT DO YOU WANT TO TALK ABOUT ?
AMIT > CAN I TALK ABOUT MY PROBLEM ?
ELIZA > SURE... ?
AMIT > DUE TO ON GOING EXAMS IN COLLEGE, I HAVE A LOT OF STRESS.
ELIZA > PLEASE RELAX A LITTLE. SLEEP WELL.
AMIT > THANKS FOR YOUR ADVICE.
ELIZA > NO MENTION.
AMIT > BYE
ELIZA > BYE AND KEEP IN TOUCH...
=====

```

Het probleem waar de slachtoffers van ELIZA tegenaan liepen, noemen we het probleem van 'andere geesten'. Het voorbeeld illustreert niet alleen dat we ons kunnen vergissen wanneer we bewustzijn en bewustzijnsinhouden toeschrijven aan iets of iemand anders, maar ook dat we daar een zeer sterke neiging toe hebben. We interpretern gretig (en soms dus al te gretig) het gedrag van 'anderen' in termen van bewustzijnsinhouden waarmee we, strikt genomen, enkel bij onszelf vertrouwd zijn. Dat doen we natuurlijk niet zomaar. We doen het omdat we op basis van dergelijke veronderstellingen richting kunnen geven aan hoe we met 'anderen' communiceren en omgaan. Maar als dat interpretatieproces fout gaat, gaat het meestal ook goed fout; daar kennen we allemaal voorbeelden van.

In dit hoofdstuk gaan we niet bestuderen hoe we in de alledaagse praktijk betere veronderstellingen kunnen maken over het bewustzijn en de bewustzijnsinhouden van anderen. Daar gaan andere boeken en andere disciplines over. Wel proberen we het filosofisch probleem van 'andere geesten' te ontrafelen. We gaan na op welke gronden we dergelijke veronderstellingen maken en proberen de kloof te overbruggen tussen onze eigen bewustzijnsinhouden waarmee we 'als geen ander' vertrouwd zijn, en de bewustzijnsinhouden van anderen die we menen te kunnen aflezen van hun uitspraken en hun hun gedrag.

Wat houdt dat probleem van 'andere geesten' nu precies in? We gaan er niet alleen genoegzaam van uit dat er een buitenwereld bestaat, maar ook

dat er in die buitenwereld mensen rondlopen die net als wij kunnen genieten van een kop koffie of het Vioolconcert van Sibelius, die pijn kunnen hebben, liefdesverdriet, emoties, stemmingen, verlangens, opvattingen – kortom, mensen die een mentaal leven hebben dat lijkt op dat van ons, weliswaar niet tot in de details (naar verluidt zijn er mensen die niet kunnen genieten van het Vioolconcert van Sibelius), maar dan toch in grote lijnen.

Natuurlijk zijn we vaak, wat de psyche van andere mensen betreft, niet alleen benieuwd naar de grote lijnen, maar ook naar de details. Hoe beleeft een baby de wereld? Zie jij wel dezelfde kleuren als ik? Welk soort pijn heeft iemand anders? Hoeveel pijn heeft die persoon? Ervaart die persoon überhaupt wel pijn? En wat zijn de ervaringen van iemand die verlamd is of in coma ligt?

HABLE CON ELLA (Pedro Almodovar, 2002)

In deze film wordt de journalist Marco verliefd op de beroemde toreador Lydia, maar zij raakt in de arena ernstig verwond door een stier en valt in een diepe coma. In het ziekenhuis waar Lydia wordt verzorgd, maakt Marco kennis met de minzame verpleger Benigno die de zorg op zich heeft genomen voor een andere patiënte: Alicia, een danseres die hij bewondert, maar die ook in een diepe coma ligt.

Met zijn liefdevolle zorg voor Alicia toont Benigno aan Marco hoe men het best om kan gaan met een comateuze geliefde: verzorg haar lichamelijk respectvol, zet haar letterlijk en figuurlijk in het zonnetje, en vooral: 'Spreek met haar' (vandaar de titel: *Hable con ella*). Inderdaad, dagelijks vertelt Benigno Alicia honderduit over wat hij allemaal heeft beleefd en vooral ook over de dansvoorstellingen die hij heeft bijgewoond. Hoewel Marco zich afvraagt of zoveel zorg iets uitmaakt en hoewel ook persoonlijke belemmeringen hem ervan weerhouden om dezelfde zorg te geven aan Lydia, raakt hij helemaal ingenomen met de manier waarop Benigno omgaat met Alicia. En dan gebeurt er een wonder of misschien wel een drama: de comateuze Alicia blijkt zwanger.

Hable con ella speelt ontzettend mooi in op de onzekerheid die we hebben over de belevingswereld van comateuze patiënten. Benigno is ervan overtuigd dat zijn liefdevolle zorg en de verhalen die hij vertelt, een positief verschil maken voor Alicia; Marco is en blijft echter sceptisch. Maar je kan de film ook anders bekijken: zelfs als het geen enkel verschil maakt in de belevingswereld van comateuze patiënten, dan nog moeten we liefdevol met hen omgaan en zelfs tegen hen blijven spreken, want aan dat soort kwalitatieve, intermenselijke verhoudingen moeten we in alle omstandigheden blijven vasthouden.



Benigno (Javier Cámara) en Alicia (Leonor Elizabeth Ceballos Watling) in *Hable con ella*, een film van Pedro Almodovar uit 2002 (still uit de film)

Deze vragen zijn niet louter theoretisch. Verre van: de antwoorden kunnen een immens verschil maken voor hoe we ons verhouden tot andere personen en hoe we met hen omgaan. Als we bijvoorbeeld zouden weten dat iemand die in coma ligt tot op zekere hoogte kan begrijpen wat er tegen hem of haar wordt verteld, of dat de comateuze patiënt toch in zekere mate kan genieten van de zorg die we hem of haar geven, dan zouden we wellicht extra zorgvuldig omgaan met de patiënt.

Of een ander voorbeeld: als iemand zegt hoofdpijn te hebben, dan laten we hem of haar doorgaans rustig opknappen. Maar als we vermoeden dat de persoon die hoofdpijn misschien wel fingeert, dan zouden we proberen een gesprekje aan te knopen om de redenen te achterhalen waarom de betrokkene die hoofdpijn fingeert. Dan zouden we dieper willen doordringen in de binnenwereld of de belevingswereld van de ander.

Telkens wanneer we vragen stellen over de binnenwereld of de belevingswereld van iemand anders en zoeken naar antwoorden, lopen we tegen het probleem aan dat we nooit rechtstreeks toegang hebben tot wat er zich in het bewustzijn van een andere persoon bevindt of afspeelt, dat wil zeggen tot zijn of haar mentale toestanden of processen. Wat we rechtstreeks kunnen waarnemen is verbaal of non-verbaal gedrag, maar niet wat erachter steekt in zijn of haar geest.

Omgekeerd is het zo dat we exclusief toegang lijken te hebben tot onze eigen binnenwereld of belevingswereld. Als ik tandpijn heb, kan iemand anders misschien wel mijn jammerklachten aanhoren, mij naar de tandarts zien rennen en redelijkerwijs afleiden dat ik tandpijn heb, maar alleen *ik* kan mijn tandpijn ook echt voelen, letterlijk 'aan den lijve ondervinden'.

Dat dit een fundamenteel probleem is, laat zich misschien het best begrijpen als we inschatten hoe moeilijk het is om aan iemand die niet vertrouwd is met een bepaald soort ervaring, die ervaring uit te leggen. Kijk, pijn hebben we allemaal wel eens gehad, dus als jij aan iemand vertelt pijn te hebben, dan zal die persoon zich daar gemakkelijk iets bij kunnen voorstellen. Maar stel nu dat je aan een persoon die vanaf zijn geboorte blind is, moet uitleggen wat dat nu eigenlijk is, 'zien', en wat voor soort ervaringen 'visuele ervaringen' zijn. Of stel dat je iemand die om een of andere reden wel zoet, zuur en zout kan proeven, maar niet bitter, zou proberen uit te leggen hoe dat nu eigenlijk smaakt, 'bitter', en wat de eigenheid van bitterheid is onder de smaken. Is dat niet uiterst moeilijk, zo niet onmogelijk? Zolang de betrokkene geen minimum aan eigen ervaringen heeft waarop we kunnen voortbouwen en, cruciaal, zolang we geen redenen hebben om te veronderstellen dat zijn of haar ervaringen voldoende vergelijkbaar zijn met de onze, zullen onze pogingen tevergeefs zijn.

Dat de toegang die we hebben tot onze eigen belevingswereld van een totaal andere aard is dan de toegang die we hebben tot de belevingswereld

van andere personen, wordt in de filosofie het probleem van 'andere geesten' genoemd. Dat probleem bestaat uit drie deelproblemen:

HET METAFYSISCH PROBLEEM

Wat van alles wat zich buiten ons bevindt, heeft bewustzijn en wat niet? Welke dieren hebben bewustzijn? Kunnen machines bewustzijn hebben of krijgen? Zijn er soorten bewustzijn? Zijn er gradaties van bewustzijn? En zo ja, waar trekken we de grens? Wat is het criterium?

HET EPISTEMOLOGISCH PROBLEEM

Kunnen we zeker weten dat anderen bewustzijn hebben? Waarop is onze kennis van de bewustzijnsinhouden van anderen gebaseerd? Kunnen we weten of iemand anders dezelfde ervaringen heeft als wij? En zo ja, hoe?

HET CONCEPTUEEL PROBLEEM

Hoe kunnen we de begrippen die we hanteren om ons eigen mentaal leven te bevatten, ooit gebruiken om het mentaal leven van iemand anders te begrijpen? Hoe kan mijn concept van 'pijn' ooit verder reiken dan mijn 'pijn'? Of andersom: kan ik wel de 'eigenaar' zijn van mijn 'pijn', in die zin dat iemand anders de 'eigenaar' kan zijn van zijn of haar 'pijn'? Veronderstelt een concept als 'pijn' niet dat we 'eigenaars' buiten beschouwing laten?

Bij het onderscheid tussen het metafysisch, het epistemologisch en het conceptueel deelprobleem zijn filosofen het (zoals vaker) onderling oneens over de vraag welk van de drie nu het meest fundamentele probleem is. Er is ook duidelijk overlapping tussen enerzijds het metafysisch en het epistemologisch probleem en anderzijds het epistemologisch en het conceptueel probleem.

Merk op dat het probleem van 'andere geesten' niet zomaar een bijkomend toepassingsdomein is van het scepticisme, maar dat het in de gedachtegang van dit boek een verdere stap vormt; het betreft een meer radicale sceptische uitdaging. In de vorige hoofdstukken ging het nog over het probleem dat we de buitenwereld nooit rechtstreeks lijken waar te nemen, wat ontkend wordt door het direct realisme. Maar bij de binnenwereld of belevingswereld is het probleem veeleer dat er een groot contrast is tussen onze eigen bewustzijnsinhouden die we alleen zelf kunnen ervaren, en de bewustzijnsinhouden van andere mensen van wie we hoogstens het gedrag 'rechtstreeks' kunnen waarnemen en de bewustzijnsinhouden alleen maar 'onrechtstreeks' kunnen afleiden.

Kortom, op het eerste gezicht lijkt direct realisme bij deze vorm van scepticisme geen uitweg te kunnen bieden. Zelfs als we in navolging van onder meer Thomas Reid geloven dat we objecten in de buitenwereld direct

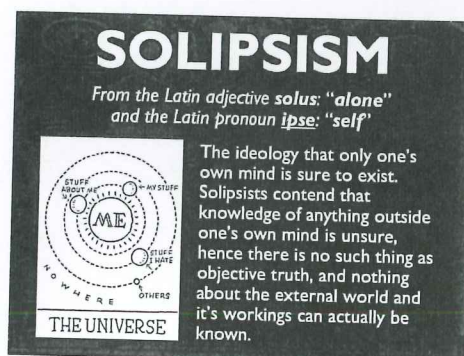
waarnemen, lopen we nog tegen het probleem aan dat we niet rechtstreeks in het hoofd van iemand anders kunnen kijken. Dit is het vraagstuk van 'andere geesten': we kunnen onmogelijk direct toegang krijgen tot anderen's binnenwereld of beleavingswereld.

4.2 Solipsisme en zelf

Een interessante en radicale, maar anderzijds louter theoretische positie tegenover het vraagstuk van het 'andere geesten' is het solipsisme. Als je een solipsist bent, dan geloof je dat er in de hele wereld maar één bewustzijnsvorm is en dat jij die ene gelukkige bent – of die ene ongelukkige, het hangt er maar van af hoe goed je als solipsist in je vel zit.

Merk op dat je door solipsist te worden, in één klap van alle filosofische zorgen verlost bent – en dat is dan weer een voordeel. Wat het metafysisch probleem betreft: er is totaal geen sprake van bewustzijnsvormen buiten jou. Epistemologisch lijkt er ook geen vuiltje aan de lucht te zijn: je bent er niet alleen zeker van dat jij bewustzijn hebt, maar bovendien kan je je onmogelijk vergissen op het gebied van je eigen bewustzijnsinhouden. En het conceptueel probleem lijkt zich ten slotte evenmin te stellen, want er zijn geen andere bewustzijnsvormen die jij met jouw begrippen zou moeten proberen te begrijpen. Voor een solipsist is er simpelweg geen metafysische, epistemologische of conceptuele kloof die op een of andere doordachte wijze overbrugd zou moeten worden.

Bron: www.sott.net



Hoewel geen enkele grote filosoof in alle ernst het solipsisme verdedigd heeft, dreigen sommige benaderingen er wel op uit te lopen. Sterker nog, sinds de vroegmoderne wijsbegeerte worden metafysische, epistemologische en conceptuele problemen vaak in 'egocentrische', bijna solipsistische termen gesteld. Descartes' *cogito*-argument dat we in hoofdstuk 1 hebben besproken, komt bijvoorbeeld aardig in die buurt.

Op het eerste gezicht is Descartes' denkstap overtuigend: *Cogito ergo sum*, 'Ik denk, dus ik besta' – er lijkt geen speld tussen te krijgen. Maar bij

nader inzien is het nog maar de vraag of deze redenering klopt. In de aanloop ernaartoe stelt Descartes dat hij alle mogelijke opvattingen in twijfel heeft getrokken. Maar als hij vervolgens de denkstap zet dat er ten minste één ding is dat niet verder betwijfeld kan worden, namelijk het bestaan van een rede die alle mogelijke opvattingen ter discussie stelt, dan lijkt hij stiekem wel degelijk uit te gaan van enkele substantiële veronderstellingen, wat betekent dat hij niet *alle* opvattingen in twijfel heeft getrokken.

Ten eerste veronderstelt Descartes dat die *res cogitans* zomaar zelfstandig kan worden gedacht, in gedachten geïsoleerd als een iets, een ding, los van alle andere dingen. Bovendien gaat hij ervan uit dat hij meteen toegang heeft tot de ware aard van dat iets, namelijk dat het in wezen een *res cogitans* is. Je zou zelfs kunnen zeggen dat Descartes in zekere zin precies datgene veronderstelt waar het hem om te doen is: dat de geest los gedacht kan worden van het lichaam. Volgens deze lezing zou je Descartes' redenering circulair kunnen achten.*

Hoe het ook zij, Descartes heeft die denkstap gezet en is daarmee een richting ingeslagen die de koers van de vroegmoderne wijsbegeerte in belangrijke mate heeft bepaald. Want het 'ik', het 'zelf' in deze redenering is intrinsiek solipsistisch: het is een bewustzijnsvorm die geen uitgebreidheid heeft in de ruimte, die niet noodzakelijk gelokaliseerd is in een lichaam en die, cruciaal, overtuigd kan zijn van zijn eigen bestaan als *res cogitans* en van zijn eigen bestaan alleen.

Let wel, Descartes gelooft dat hij vervolgens op dezelfde redelijke grond het bestaan van God en het bestaan van de buitenwereld kan 'bewijzen'. En dat is het solipsisme ver voorbij. Maar je kan je afvragen of dat wel 'dezelfde redelijke grond' is, want in tegenstelling tot zijn 'bewijzen' voor God en buitenwereld baseert Descartes de afleiding van zijn eigen bestaan als *res cogitans* uitsluitend op het gegeven dat hij exclusief toegang heeft tot zijn eigen bewustzijnsinhouden. Die exclusieve toegang noemen we introspectie.

Ook Hume beroept zich op introspectie, maar hij komt uit op een heel ander zelfbegrip dan Descartes. Waar het 'zelf' van Descartes één en ondeelbaar is, bewust en niet-uitgebreid, daar constateert Hume door middel van introspectie dat ervaringen, overtuigingen, herinneringen enzovoort continu komen en gaan, continu in beweging zijn. Het 'zelf' is bij Hume hoogstens 'een bundel' van ervaringen, overtuigingen, herinneringen enzovoort, en niet iets wat daar los van staat of los van kan worden gedacht.**

4.3 Filosofische zombies en spectruminversie

Je kunt nooit zeker weten of 'anderen' alleen schijnbaar een mentaal leven hebben dat in grote lijnen te vergelijken is met dat van jou, terwijl ze in werkelijkheid helemaal geen binnenwereld of belevingswereld hebben. Dan zijn het 'zombies' of 'leeghoofden'. De mogelijkheid van zombies is

* De meeste historici zijn het erover eens dat, opgevat als een argument, Descartes' *cogito*-redenering circulair is. Een mogelijke oplossing hiervoor suggereert Hintikka (1962): het *cogito* is eerder een soort *speech act* dan een argument in strikte zin.

** Zie de primaire tekst van Descartes, *Over de natuur van de menselijke geest*; dat deze beter te kennen is dan het lichaam, op pp. 209-215 van dit boek. Zie ook de primaire tekst van Hume, *Over persoonlijke identiteit*, op pp. 216-225.

niet alleen relevant voor het metafysisch probleem (het kan zijn dat ze bestaan), maar ook voor het epistemologisch probleem (we kunnen niet achterhalen of anderen zombies zijn, we kunnen die mogelijkheid nooit uitsluiten).

Ook robots kunnen zombies zijn. Stel bijvoorbeeld dat een robot, die er voor het overige precies zo uitziet en zich gedraagt als een normaal persoon, op een specifieke manier is geprogrammeerd: wanneer de warmtesensoren in zijn vinger oververhit geraken, wordt de vinger teruggetrokken en roept de robot 'Au!', schijnbaar met gevoel. In dat geval zullen we doorgaans geloven dat die robot pijn en bij uitbreiding een mentaal leven heeft. Alleen wanneer het gedrag van de robot in één of ander opzicht afwijkt van normaal menselijk gedrag, kunnen we misschien, als we de nodige moeite doen, achterhalen dat de robot een heel ander soort of zelfs helemaal geen mentaal leven heeft.

De mogelijkheid van zombies en de moeilijkheden die gepaard gaan met de ontmaskering ervan, worden in talrijke films geëxploreerd.

THE STEPFOORD WIVES (Frank Oz, 2004)

Deze film gaat over de succesvolle programmamaakster Joanne Eberhart die plots wordt ontslagen na een incident met een deelnemer aan een *reality show*. Om haar te helpen haar gedachten te verzetten en er psychisch weer bovenop te komen, besluit haar man Walter Kresby samen met haar te verhuizen naar het chique Stepford. Van meet af aan verbaast Joanna zich erover hoe buitengewoon gewillig en gedwee de vrouwen van Stepford zich schikken naar de wensen van hun echtgenoten en andere familiale en sociale noden. Ook Joanna's vriendin Bobby Markowitz vermoedt dat er iets niet pluis is met de vrouwen in Stepford en samen proberen ze het geheim te achterhalen.

Intussen raakt Walter Kresby meer en meer betrokken bij een besloten mannenclub onder leiding van de mysterieuze en charismatische Mike Wellington. Op zeker moment blijkt dat de vrouwen in Stepford, op verzoek van de mannen, niet alleen lichamelijk worden 'geoptimaliseerd', maar ook systematisch geestelijk worden 'ontzield' en vervangen door robots. Voor haar vriendin Bobby blijkt het al te laat te zijn, maar Joanna zelf probeert nog aan dit lot te ontsnappen.

De feministische inzet van het verhaal is natuurlijk dat mannen liever leeghoofden hebben dan vrouwen met een eigen wil, eigen verlangens, etc. Interessant is hoe Joanne en eigenlijk ook Bobby op het spoor komen van wat er aan de hand is. Hoewel de robotvrouwen uiteindelijk worden verraden door hun gedrag, komt dat niet doordat dat gedrag ongerijmd zou zijn, maar juist doordat het tot in de perfectie afgesteld is op de wensen van de mannen.



De robotvrouwen uit
The Stepford Wives, een
film van Frank Oz uit 2004
(still uit de film)

Of robots op het gebied van het mentaal leven echt het tegenovergestelde zijn van mensen, is nog maar de vraag. Om een robot menselijke taken te laten uitvoeren, zijn sensorische input en motorische output van groot belang. Volgens nogal wat theorieën in de cognitie- en bewustzijnsfilosofie verschilt datgene wat mensen in dat opzicht doen, niet of nauwelijks van hoe robots te werk gaan. In beide gevallen is het een kwestie van het manipuleren van representaties.

Volgens het zogenaamde functionalisme kunnen mentale toestanden volledig worden gekarakteriseerd in termen van de 'functionele rol' die ze spelen. Die functionele rol omvat de sensorische input waaraan mentale toestanden ontspringen, de motorische output waartoe ze leiden en de manier waarop ze zich verhouden tot andere mentale toestanden. Hoe die mentale toestanden materieel worden gerealiseerd, maakt echter niets uit. Of het nu gaat om een menselijk brein, het brein van een marsmannetje of een computersysteem: zolang hun interne toestanden dezelfde functionele rol spelen, blijven ook de mentale toestanden dezelfde.

Je kan die 'functionele rol' van mentale toestanden vergelijken met de regels die gelden voor een schaakstuk.* Of een stuk bijvoorbeeld een loper is, wordt niet bepaald door de fysische vorm die het heeft, maar door de zetten die je met die loper mag doen. Of je met een ivoren stuk, een stukje hout of een kroonkurk speelt, maakt geen enkel verschil. Iets is een loper als het bij het begin van het spel tussen paard en koning of koningin staat en als je er alleen diagonaal mee over het bord mag bewegen. Met dat soort regels kan een loper volledig bepaald worden, los van hoe het stuk materieel wordt gerealiseerd.

* Zie ook Dooremalen,
De Regt & Schouten (2010).

Functionalisten geloven dat iets soortgelijks geldt voor de karakterisering van mentale toestanden: ook die kunnen volledig worden bepaald zonder rekening te houden met hoe ze materieel worden gerealiseerd. Het gevolg is dat artificiële systemen evengoed 'mentale toestanden' kunnen hebben – en dus ook een soort binnenwereld.

WESTWORLD (Michael Crichton, 1973)



De robotcowboy uit *Westworld*, een film van Michael Crichton uit 1973 (filmaffiche)

Bezoekers die het zich kunnen veroorloven, mogen in het themapark Delos geavanceerde robots die er precies zo uitzien en handelen als mensen, bevechten en doden. Na de eliminatie worden de robots hersteld en kan het spel weer van voren af aan beginnen.

Twee toeristen die naar Delos komen, Peter Martin en John Blane, prefereren Westworld boven de andere mogelijke werelden Romanworld en Medievalworld. Zij worden belaagd door een in het zwart geklede robotcowboy, maar slagen er tot hun grote vreugde steeds in de 'revolverheld' tijdig overhoop te schieten. Maar dan gaat het mis: door een computerstoring beginnen de robots dol te draaien en de toeristen massaal om het leven te brengen. Blane wordt eerst gebeten door een ratelslangrobot en vervolgens in een duel doodgeschoten door de robotcowboy. Martin vlucht weg door Romanworld en Medievalworld; hij overwint de robotcowboy waardoor hij wordt achtervolgd en uiteindelijk blijkt hij als enige toerist het hele drama overleefd te hebben.

De film *Westworld* is een klassieker in het sciencefiction-horroggenre. Later heeft regisseur Michael Crichton opnieuw de setting van een sciencefiction-themapark gebruikt, namelijk voor *Jurassic Park*. Voor ons is het vooral interessant hoe het verschil tussen toeristen en robots in beeld wordt gebracht: robots bewegen een beetje anders en soms zie je dat een robot zichzelf even uitschakelt. Uniek aan *Westworld* is dat je als kijker toegang krijgt tot het 'innerlijk leven' van de robotcowboy, in die zin dat er computeranimatie wordt gebruikt om ons – voorbij de problematiek van het 'vreemd-psychische' – door de ogen van de robot te laten kijken.

Om het probleem van zombies echt op scherp te stellen, maken sommige filosofen onderscheid tussen twee soorten bewustzijn of twee aspecten van bewustzijn: aan de ene kant alles wat te maken heeft met representaties en het manipuleren ervan, en aan de andere kant de pure, kwalitatieve beleving daarvan, dat wil zeggen de ervaring zelf. We zouden gemakshalve kunnen zeggen dat interne representaties en het manipuleren daarvan de 'binnenwereld' vormen, en dat de pure, kwalitatieve beleving de 'belevingswereld' uitmaakt.

Zelfs als we meegaan met functionalisten die stellen dat robots en andere artificiële systemen in principe ook een 'binnenwereld' kunnen hebben, dan nog lijkt het intuïtief moeilijk aanvaardbaar dat ze daarmee ook een heuse 'belevingswereld' zouden hebben. Als een robot het Vioolconcert van Sibelius 'registreert' en 'verwerkt', dan kan er cognitief gesproken van alles in die robot gebeuren dat te vergelijken valt met menselijke cognitieve processen, maar 'hoe het is' om naar het Vioolconcert van Sibelius te luisteren, dat maakt voor ons veel uit, maar voor de robot niet(s).

De kwalitatieve belevingsaspecten die uitmaken 'hoe het is' om een bepaalde ervaring te hebben, vormen samen het fenomenaal bewustzijn. Als we dat toepassen op de mogelijkheid van zombies, dan lijken we een nieuw soort zombies te kunnen bedenken: de zogenaamde 'filosofische zombies' die qua binnenwereld volledig lijken op jou en mij, terwijl ze qua belevingswereld dramatisch van ons verschillen. Die robots zijn namelijk geheel verstoken van het fenomenaal bewustzijn dat zo karakteristiek is voor de mens en naar alle waarschijnlijkheid ook voor veel dieren.

Wat houdt het gedachte-experiment van filosofische zombies precies in? Als jij een kop koffie drinkt, dan heeft dat een specifieke smaak. Het is denkbaar dat je een zombie-tegenhanger hebt die in alle opzichten precies op jou lijkt, die op dezelfde manier in elkaar zit, dezelfde dingen ondergaat en dezelfde dingen doet en laat, maar die niet de specifieke smaakervaring heeft die jij hebt wanneer je een kop koffie drinkt.

Let wel, je zombie-tegenhanger heeft bij het drinken van de kop koffie niet zomaar een andere smaakervaring dan jij; hij heeft helemaal geen smaakervaring. Je zou kunnen zeggen dat voor hem de koffie naar niets smaakt (maar dan echt naar niets, dus ook niet naar water of zo). Sterker nog, je zombie-tegenhanger heeft helemaal geen zintuiglijke ervaringen. Van buitenaf valt je zombie-tegenhanger in niets van jou te onderscheiden, maar binnenin is er tussen jullie een wereld van verschil. Jij beleeft je zintuiglijke ervaringen op een specifieke manier, maar bij je zombie-tegenhanger valt er helemaal niets te beleven.

Sommige filosofen geloven dat dergelijke zombie-tegenhangers denkbaar en bijgevolg mogelijk zijn. Volgens hen toont de mogelijkheid van die filosofische zombies aan dat een uiterst belangrijk deel van ons bewustzijn

(het fenomenaal bewustzijn, dat wil zeggen de eigenlijke beleving van onze ervaringen) niet afhankelijk is van ons brein en de rest van ons lichaam. Hun argument is dat we in gedachten de beleving van zintuiglijke ervaringen kunnen loskoppelen van alles wat lichamelijk daarmee gepaard gaat. Andere filosofen menen daarentegen dat dergelijke 'filosofische zombies' denkbaar noch mogelijk zijn: als jouw tegenhanger werkelijk op precies dezelfde manier in elkaar zit als jij, dan zal hij ook precies dezelfde ervaringen hebben en die op precies dezelfde manier beleven.

De hamvraag is of de zogenaamde *qualia*, de zuiver kwalitatieve aspecten van de ervaring, afzonderlijk kunnen worden gedacht. Kunnen we de belevingswereld, minstens in gedachten, loskoppelen van de binnenwereld?

Merk op dat we met dergelijke gedachte-experimenten doorgaan op de weg die Descartes is ingeslagen. Descartes maakt een fundamenteel onderscheid tussen *res cogitans* en *res extensa* op grond van de mogelijkheid om, in gedachten, de geest of het bewustzijn los te koppelen van het lichaam. Hij meent ook voorbeelden te kunnen geven van cognitieve vaardigheden, zoals taalvermogen en wiskundig inzicht, die volgens hem nooit volledig gevat kunnen worden in termen van lichamelijke, materiële processen.

Maar voor het overige kan Descartes geen positieve karakterisering geven van de specifieke aard van het bewustzijn, noch van de manier waarop er interactie plaatsvindt tussen dat bewustzijn en de lichamelijke, materiële processen, zeg maar: wat er in het brein gebeurt. Als we aan de hand van filosofische zombies in gedachten een onderscheid kunnen maken tussen de binnenwereld en de belevingswereld (of technischer: tussen de intentionele of representatieve en de kwalitatieve aspecten van onze ervaringen), dan lijken we dus wat preciezer dan Descartes aan te kunnen duiden waarin verklaringen van ervaringen in termen van lichamelijke, materiële processen (of in termen van 'functionele rollen') zullen blijven tekortschieten.

De problematiek van andere geesten omvat meer dan het probleem van de mogelijkheid van zombies. Zelfs als we op een of andere manier zouden kunnen weten dat andere wezens (robots, dieren, andere mensen) bewustzijn hebben, dan nog dient de vraag zich aan of, in welke zin en in welke mate hun mentaal leven lijkt op dat van ons. Want het is mogelijk, zoals talrijke sciencefictionfilms illustreren, dat iemand zich in grote lijnen gedraagt zoals wij, terwijl de betrokkene in mentaal opzicht helemaal anders in elkaar zit. Hij heeft dan weliswaar bewustzijn, maar zijn bewustzijns-vorm is van een totaal andere orde dan die van ons.

INVASION OF THE BODY SNATCHERS (Philip Kaufman, 1978)

In San Francisco regenen kosmische zaden neer die zich in korte tijd ontwikkelen tot gewassen die menselijke individuen waarnemen, hun lichaam volledig kopiëren en eropuit zijn om letterlijk en figuurlijk de plaats in te nemen van de originelen. Perfect zijn de kopieën niet, want in hun gedrag zijn ze wat apathischer dan de originelen en er tekenen zich ook kleine verschillen af in de interactiepatronen en communicatievormen. Desalniettemin kost het gezondheidsinspecteurs Matthew Bennell en Elizabeth Driscoll ontzettend veel moeite om te achterhalen wat er precies aan de hand is, om uit te maken wie er nog mens is en wie plant, en om samen met hun vrienden Jack en Nancy Bellicec de stad te ontvluchten voordat zijzelf ten prooi vallen aan de *body snatchers* die zich razendsnel verspreiden.

De film *Invasion of the Body Snatchers* illustreert zowel de metafysische als de epistemologische problematiek van 'andere geesten'. Het thema blijkt bijzonder populair te zijn: Kaufmans film is op zichzelf al een succesvolle *remake* van het gelijknamige origineel van Don Siegel (1956), maar er is ook een recentere versie van Abel Ferrara (1993) met de kortere titel *Body Snatchers*. De genoemde films spelen in op de tijdgeest: terwijl Kaufmans film de paranoia van de jaren zestig en zeventig over een dreigende kernoorlog uitstraalt, verbeeldt Abel Ferrara het Hollywood van de jaren tachtig en negentig dat niet alleen acteurs, maar nagenoeg alle betrokkenen helemaal uitzuigt en als lege poppetjes achterlaat.*

* Voor deze interpretatie van Ferrara's film, zie Mahy (1999).



Paranoia in *The Invasion of the Body Snatchers*, een film van Philip Kaufman uit 1978 (filmaffiche)

Het probleem is dat we op basis van gedrag niet kunnen uitsluiten dat het mentale leven van iemand anders er helemaal anders uitziet dan dat van ons. Als we van iemand anders zeggen dat hij of zij een 'gesloten boek' is, dan bedoelen we doorgaans dat die persoon niet gemakkelijk vertelt over zijn of haar belevenissen, gevoelens en dergelijke meer. Dat is dan een bijzonder persoonlijkheidskenmerk. Maar in de filosofische betekenis is het mogelijk dat iedereen voor iedereen een 'gesloten boek' is. We zijn allemaal intiem vertrouwd met onze eigen bewustzijnsinhouden en hebben allemaal begrippen ontwikkeld aan de hand waarvan we met elkaar denken te kunnen communiceren over onze bewustzijnsinhouden – begrippen als 'dat smaakt bitter' of 'mijn nabeeld is blauw'. Maar hoe kunnen we weten of we met 'bitter' en 'blauw' hetzelfde bedoelen?

Nog spannender wordt het probleem als we op een graduele manier gaan nadenken over mogelijke verschillen tussen onszelf en anderen. Dan kunnen we ons voorstellen dat aspecten die in ons eigen mentale leven nauw met elkaar samenhangen – zoals bijvoorbeeld herinneringen en emoties – in het mentale leven van anderen volstrekt los van elkaar staan. Op die manier wordt het steeds moeilijker om de grens te trekken.

Om het probleem van het 'gesloten boek' tot ons te laten doordringen, hoeven we geen beroep te doen op de – soms vergezochte – scenario's van sciencefictionfilms. Enige reflectie op ons eigen mentale leven volstaat om ons te kunnen inbeelden dat bepaalde aspecten en inhouden van ons bewustzijn bij andere personen wel eens helemaal anders zouden kunnen zijn, zonder dat we dat ooit zullen kunnen achterhalen.

Merk op dat we het hier opnieuw hebben over de kwalitatieve aspecten van de ervaringen, en niet zozeer over de representatieve of intentionele aspecten. Als datgene wat ik 'bitter' noem, qua beleving een heel andere smaak is dan wat jij 'bitter' noemt, terwijl er voor het overige geen enkel verschil is tussen ons (we blijven bijvoorbeeld dezelfde groenten en bieren 'bitter' noemen), dan zullen we een 'gesloten boek' voor elkaar zijn zonder dat we het (ooit kunnen) merken. In dat geval verschillen we dus qua fenomenale aspecten van het bewustzijn zonder dat we verschillen qua representatieve of intentionele aspecten.

Laten we verder denken in deze lijn aan de hand van het gedachte-experiment van spectruminversie. Stel dat jij en ik naar een rijpe banaan kijken. Desgevraagd zullen we allebei rapporteren dat de kleur van de banaan geel is. Maar weten we wel zeker dat we dezelfde kleur ervaren? Zou het niet kunnen dat jouw geel mijn blauw is, en jouw blauw mijn geel? En dat we dat verschil niet merken, omdat we allebei van kindsbeen af geleerd hebben om de kleur die we zien wanneer we naar een rijpe banaan kijken, geel te noemen?

Mochten er kleinere verschillen zijn tussen de kleuren die jij ervaart en

de kleuren die ik ervaar, dan zouden we die wellicht kunnen opsporen. Zo kan bijvoorbeeld kleurenblindheid gemakkelijk vastgesteld worden. Maar stel nu dat de verschillen groot en systematisch zijn, dat overal waar jij de ene kleur ziet, ik de 'tegegengestelde' kleur zie, dat wil zeggen de kleur die op de kleurencirkel tegenoverliggend is. Zouden we een dergelijk systematisch verschil, zo'n spectruminversie kunnen achterhalen?

Zoals vaak lopen de meningen van de filosofen uiteen. Sommige geloven dat spectruminversie niet alleen mogelijk is, maar dat er zelfs goede redenen zijn om aan te nemen dat het werkelijk voorkomt.* Anderen beargumenteren dat spectruminversie als dusdanig misschien wel mogelijk is, maar dat we elke denkbare spectruminversie zullen kunnen detecteren. Een argument daarvoor is dat op een kleurencirkel de kleuren weliswaar netjes symmetrisch tegenover elkaar liggen, maar dat onze kleurervaringen in werkelijkheid grillig en asymmetrisch zijn. Zo onderscheiden we bijvoorbeeld veel meer tinten rood dan tinten groen. Mocht jouw rood dus mijn groen zijn en omgekeerd, dan zouden er binnen de kortste keren verschillen tussen ons opduiken in de hoeveelheid tinten die we van 'die kleur' kunnen onderscheiden.**

* Zie Nida-Rümelin (1996).

Concreet speelt uitwendig waarneembaar gedrag weliswaar een uiterst belangrijke rol bij de vraag aan wie we wanneer welke mentale toestanden toeschrijven. Maar het probleem dat door de mogelijkheid van spectruminversie aan de orde wordt gesteld, is dat de relatie tussen gedrag en mentale toestanden niet één-op-één is en dat we ons danig kunnen vergissen bij wat we aan anderen toeschrijven. Een Spartaans getraind persoon kan pijn hebben, maar geen krimp geven. En een beetje acteur kan liggen kronkelen en kermen alsof hij vergaat van de pijn, terwijl hij zich in feite kiplekker voelt.

** Zie Hardin (1997).

4.4 Pseudoproblemen?

Vormen sceptische scenario's zoals zombies en spectruminversie een ernstige bedreiging voor onze opvatting dat andere mensen (en dieren) bewustzijn hebben, voor onze rudimentaire kennis van elkaars bewustzijnsinhouden en voor de mogelijkheid om elkaars bewustzijnsinhouden te bevatten? Of zijn het veeleer pseudoproblemen, scenario's die misschien op het eerste gezicht mogelijk lijken en zo het probleem van 'andere geesten' voeden, maar die bij nader inzien toch onmogelijk zijn?

De vraag is of het wel mogelijk is dat iemands mentale leven een volkomen 'gesloten boek' is, dat iedereen voor iedereen een 'gesloten boek' is, dat de exclusieve toegang die we hebben tot onze eigen bewustzijnsinhouden exclusief blijft. We hebben toch ook termen, psychologische begrippen, waarmee we ons mentaal leven kunnen beschrijven en kunnen communiceren over wat er zich 'van binnen' allemaal afspeelt? Toont die



Ludwig Wittgenstein



Edmund Husserl

* Zie Wittgenstein (1953) waarin de gedachte wordt ontwikkeld dat de betekenis van begrippen bepaald wordt door hun gebruik, en meer in het bijzonder door de regels van de 'taalspelen' waarbinnen ze worden gebruikt.

** Husserl (1931).

mogelijkheid van communicatie daarover niet juist aan dat het 'gesloten boek' onmogelijk is?

Als we willen aantonen dat spectruminversie een pseudoprobleem is, dan kunnen we niet volstaan met zeggen dat we toch dezelfde termen gebruiken, want daarmee hebben we in onze formulering al rekening gehouden: het is mogelijk dat we dezelfde termen gebruiken voor iets wat van binnenuit bekeken, qua beleving, helemaal anders is.

De vraag is nu hoe de begrippen die uitgedrukt worden door die termen, tot stand kunnen komen. Zouden we bij wijze van spreken helemaal van binnenuit concepten als 'pijn', 'blauw' en 'bitter' kunnen bedenken? Is het mogelijk dat we helemaal op onszelf ons eigen privétaaltje ontwikkelen, los van andere mensen en taalgebruikers?

Hierboven zijn we er steeds stiekem van uitgegaan dat we onze psychologische begrippen in volstrekte isolatie kunnen ontwikkelen. We zijn bijvoorbeeld intiem vertrouwd met onze eigen pijnervaringen, en dan halen we daar door middel van abstractie het begrip 'pijn' uit. We abstraheren onze privébelevens waar alleen wij toegang toe hebben, tot een veel algemenere belevens waarvan we aannemen dat ook anderen die kunnen ondergaan.

Maar volgens Ludwig Wittgenstein klopt die hele redenering niet.* We brengen onszelf nodeloos in de sceptische problemen door ervan uit te gaan dat psychologische begrippen op die manier opgebouwd kunnen worden. Zo'n eigen privétaaltje is volgens Wittgenstein eenvoudigweg ondenkbaar; het zou zijn werk als taal niet kunnen doen en zelfs onmogelijk als een taal kunnen worden (h)erkend.

Psychologische begrippen komen veeleer tot stand in een intersubjectieve, sociale, linguïstische context. Van meet af aan is een erg belangrijk deel van hun betekenis de toepassing ervan op andere mensen. De reden waarom we gedachte-experimenten over zombies of spectruminversie zo fascinerend vinden, is dat ze de tegenhangers zijn van de levende, menselijke wezens die onze schoolvoorbeelden vormen van 'bewuste wezens' en die we van meet af aan begrijpen met behulp van psychologische begrippen. Volgens Wittgenstein kunnen we filosofische zombies en spectruminversie gerust als pseudoproblemen terzijde schuiven.

Ook in de continentale filosofie vinden we denkers die beargumenteren dat het probleem van 'andere geesten' een pseudoprobleem is. In zijn *Cartesiaanse Meditationen* stelt Edmund Husserl dat, in tegenstelling tot wat Descartes meende, de buitenwereld nooit weggedacht kan worden, omdat ervaringen altijd 'intentionaliteit' hebben.** Volgens Husserl is onze ervaring van de buitenwereld als een objectieve werkelijkheid gegrond in onze waarneming van andere personen als wezens die net zo'n ik hebben als ikzelf (alter ego's).

In zijn *Fenomenologie van de waarneming* gaat Maurice Merleau-Ponty een stap verder.* Onze ervaring van de wereld is afhankelijk van de ervaring van een sociale wereld. Omdat Merleau-Ponty de mens als een psychofysisch geheel beschouwt, is het waarnemen van een handelend, menselijk lichaam het direct waarnemen van een persoon. Tevens stelt hij dat ik onmogelijk zekerder kan zijn van mijn eigen bestaan dan van het bestaan van anderen.

Volgens Jean-Paul Sartre ontmoeten we de ander in de ervaring van het bekeken worden. Een van zijn beroemde voorbeelden in *Het zijn en het niet* is het gevoel van schaamte.** Je schaamt je nooit alleen. Sterker nog, juist in je schaamte heb je direct toegang tot de andere persoon.

Volgens Emmanuel Levinas ten slotte is de richting van de aanblik veel eerder omgekeerd: er gaat een direct ethisch appèl uit van de Ander in onze waarneming van zijn of haar (gezicht als een) gelaat. Wanneer Levinas 'de Ander' met een hoofdletter schijft, wil hij benadrukken dat het anders-zijn van de ander niet opgeheven kan worden, hoewel dat onophoudelijk doch tevergeefs wordt geprobeerd in de traditie van onder meer Descartes, die Levinas treffend de 'egologie' noemt.

In het algemeen kunnen we stellen dat in de continentale filosofie de sociale gebondenheid en verankering van de mens worden benadrukt. We zijn essentieel sociale wezens. We zijn niet alleen van anderen afhankelijk voor ons concrete bestaan, maar ook voor ons zelfbegrip. Ons bewustzijn van 'andere geesten' ligt dus aan de basis van ons zelfbewustzijn – en niet omgekeerd, zoals in de formulering van het probleem van 'andere geesten' wordt verondersteld.

4.5 Directe of indirecte toegang?

In hoofdstuk 2 hebben we Thomas Reid en de mogelijkheden en beperkingen van het direct realisme besproken als antwoord op de uitdaging van het metafysisch scepticisme. Vervolgens zijn we in de paragrafen 4.1 en 4.2 van dit hoofdstuk ervan uitgegaan dat als we de buitenwereld direct zouden waarnemen, we bij andere mensen weliswaar hun gedrag direct zouden waarnemen, maar nooit wat daar mentaal aan ten grondslag ligt. We kunnen immers niet in het hoofd van iemand anders kijken, of toch niet op dezelfde wijze als we in ons eigen hoofd kunnen kijken (want dan doen we door middel van introspectie).

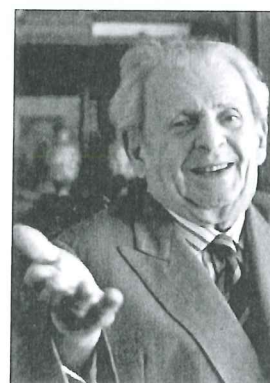
De mogelijkheid van zombies en spectruminversie (paragraaf 4.3) onderstreept bovendien dat een welbepaald soort gedrag niet stevast ontspringt aan een welbepaald soort mentale toestand. Als een man op de grond ligt te kronkelen en te kermen, dan ligt het weliswaar voor de hand om te veronderstellen dat hij pijn heeft, maar zeker weten doe je dat niet. De man kan een gesofisticeerde robot zijn die geen pijn ervaart, maar ge-



Maurice Merleau-Ponty



Jean-Paul Sartre



Emmanuel Levinas

* Merleau-Ponty (1945).

** Sartre (1943).

programmeerd is om pijnbewegingen te maken en pijn geluiden voort te brengen. Of minder vergezocht: het kan een acteur zijn die helemaal in zijn nopjes is, maar zo overtuigend kan acteren dat hij bij ons de illusie weet te wekken dat hij pijn heeft.

Merk op dat er duidelijke parallellen zijn tussen de mogelijkheid van zombies en de mogelijkheid van hallucinaties: in beide gevallen geloven we dat er iets is wat er bij nader inzien niet blijkt te zijn. En merk ook op dat er een duidelijke parallel is tussen de mogelijkheid van spectrum-inversie en de mogelijkheid van illusies: in beide gevallen geloven we terecht dat er iets is, maar vergissen we ons in een of meerdere eigenschappen die we eraan toekennen.

Als direct realisme een adequaat antwoord is op de mogelijkheid van illusies en hallucinaties, zoals bijvoorbeeld Thomas Reid heeft beargumenteerd, dan kan direct realisme op het gebied van 'andere geesten' ook een adequaat antwoord vormen op de mogelijkheid van zombies en spectrum-inversie. Akkoord, we kunnen ons af en toe vergissen op het gebied van 'andere geesten', maar dat betekent nog niet dat in al die gevallen waarin we ons niet vergissen – waarin we bijvoorbeeld terecht geloven dat een man die ligt te kronkelen en te kermen, werkelijk pijn heeft – onze opvatting afgeleid en hypothetisch zou zijn en naast allerlei sceptische alternatieven en scenario's zou moeten worden gelegd.

Volgens Thomas Reid zien we meteen dat andere mensen bewustzijn hebben; het bestaan van de andere geesten is een van de vele vanzelfsprekendheden van het gezond verstand. Welbepaalde bevindingen uit de cognitiewetenschap kunnen deze bijzondere vorm van direct realisme enigszins ondersteunen. Bekend is bijvoorbeeld dat de koppeling tussen welbepaalde basisemoties (bijvoorbeeld angst) en welbepaalde gelaatsuitdrukkingen niet alleen redelijk betrouwbaar is, maar ook universeel, dat wil zeggen dat die koppeling in alle culturen voorkomt. Basisemoties vallen dus van het gezicht af te lezen, zelfs over culturele grenzen heen.

Verder is er ook een mogelijke, maar controversiële verklaring voor hoe dat 'aflezen' van de mentale toestand van iemand anders precies tot stand komt in termen van 'spiegelneuronen'. Die neuronen 'vuren' niet alleen als we zelf een handeling uitvoeren, maar ook als we die handeling zien uitvoeren door iemand anders.* Je zou kunnen suggereren dat we op die manier toch een soort onmiddellijke toegang hebben tot het bewustzijn van de anderen.

Maar wat als we toch geen directe toegang tot andere geesten zouden hebben? Vormt indirecte toegang werkelijk een onoverkomelijk probleem? In navolging van John Stuart Mill geloven veel filosofen dat het epistemologisch probleem en het conceptueel probleem kunnen worden opgelost op grond van de analogie tussen enerzijds ons eigen gedrag en de daaraan

* Zie Frith (2007) en Churchland (2002).

gekoppelde mentale toestanden, en anderzijds het gedrag van anderen. Als ik met een vinger een gloeiende plaat aanraak, dan heb ik pijn, roep 'Au!' en trek mijn vinger zo snel mogelijk weg (niet noodzakelijk in die volgorde). Als ik vervolgens iemand anders hetzelfde zie doen (in een min of meer vergelijkbare situatie), dan kan ik afleiden dat hij of zij pijn heeft. Sterker nog, ik kan me zelfs vrij goed voorstellen wat voor soort pijn hij of zij moet hebben. Meer in het algemeen: als ik een persoon handelingen zie verrichten die ik ook verricht wanneer ik in een bepaalde mentale toestand ben, dan mag ik aannemen dat de betrokkene zowel bewustzijn als een vergelijkbare mentale toestand heeft.

In feite is de redenering die ik hierbij maak, gebaseerd op inductie. Bij mij is welbepaald gedrag ('Au!' roepen) gerelateerd aan welbepaalde mentale toestanden (pijn). Als ik vervolgens analoog gedrag van anderen zie, leid ik daaruit af dat zij wel analoge mentale toestanden zullen hebben. Maar erg betrouwbaar is deze vorm van inductie, deze analogieredenering niet, want zij is slechts gebaseerd op een enkel geval, namelijk mijn koppeling tussen welbepaald gedrag en welbepaalde mentale toestanden. Is mijn eigen geval niet een veel te smalle basis voor inductie? Vereist inductie niet dat de koppeling zich in een representatief aantal gevallen toont? En kan de koppeling zich ooit op een andere manier tonen dan louter in dat ene geval dat ik zelf ben, tenzij ik mij stiekem toch beroep op analogieredeneringen en inductie?

Sceptici aanvaarden Mills oplossing doorgaans niet, omdat inductie op basis van één geval, en dan nog een geval dat we beschouwen door middel van introspectie, geen betrouwbare methode is.

4.6 Van solipsisme tot panpsychisme

We hebben nu misschien wel het solipsisme overwonnen door aan te dringen op de wezenlijke, sociale en linguïstische verankering van het individu, maar daar hebben we ook een prijs voor betaald: antropocentrisme. Hierbij is de 'ander' per definitie een ander persoon en niet een ander soort wezen. Mentale toestanden schrijven we echter niet alleen toe aan andere mensen, maar ook aan andere dieren. Het is op zich al interessant dat we voortdurend de neiging hebben om gedrag van dieren te interpreteren in termen van welbepaalde mentale toestanden, die we vervolgens ook nog eens verwoorden in typisch menselijke termen: die hond mist zijn baasje, die kat is ijdel en laat zich niet gek maken, die papegaai vindt dat die man echt wel een loser is en spreekt dat ook uit.

Het functionalisme levert het model waarmee we ook mentale toestanden aan dieren kunnen toekennen. Aangezien een mentale toestand wordt bepaald door zijn functionele rol en niet door zijn concrete, fysische realisatie, kunnen dieren die fysiologisch helemaal anders in elkaar zitten dan

mensen, desalniettemin mentale toestanden hebben die vergelijkbaar zijn met die van mensen. De vraag is of dit functionalisme als alternatief voor het antropocentrisme niet uitloopt op een 'excessief liberalisme', waarbij mentale toestanden worden toegekend aan alles wat maar de juiste functionele organisatie heeft. Je zou welhaast denken dat er een minimum aan gelijkenis nodig is op het niveau van de fysische realisatie, voordat er sprake kan zijn van gelijkenissen op het niveau van functionele organisatie en dus van mentale toestanden. Maar dan dient de vraag zich aan waar we de grens trekken, en waarom precies daar en niet ergens anders.

Het lijkt in ieder geval zo dat we ons bij andere dieren dan de mens gemakkelijker kunnen vergissen in de concrete mentale toestanden die we eraan toeschrijven. Een bekend voorbeeld is het incident waarbij de gorilla Bokito ontsnapte en een trouwe bezoeker van Blijdorp zwaar verwondde. Die vrouw had Bokito vaak bezocht, langdurig oogcontact met hem gehad en zijn ontblote tanden geïnterpreteerd als een liefkozende lach, terwijl dat in 'gorilla-taal' veeleer een teken van dreiging is. De koppeling die de vrouw had gemaakt tussen basisemoties en gelaatsuitdrukkingen, werkte hier volstrekt contraproductief.

Hebben dieren opvattingen? Daar lijkt het wel op als we bestuderen hoe ze met hun omgeving omgaan: ze jagen op prooien, vermijden gevaar, enzovoort. Allemaal voorbeelden van intelligent gedrag, waarbij het organisme een inschatting lijkt te maken van de situatie en erop inspeelt. Maar als een hond een frisbee vangt, heeft hij dan een notie, weliswaar in zijn eigen taal, van een 'frisbee'?

Volgens Wittgenstein delen mensen en dieren alvast geen concepten: 'Als een leeuw zou kunnen spreken, zouden we hem niet kunnen begrijpen.'^{*} Davidson gaat nog een stapje verder: dieren hebben helemaal geen opvattingen.^{**} Om opvattingen te kunnen hebben, heb je een notie van waarheid nodig, en dat vereist op zijn beurt dat je deel uitmaakt van een taalgemeenschap, zo redeneert Davidson. Maar anderzijds: maken dieren dan geen deel uit van hun taalgemeenschap? Hoe kunnen we dat te weten komen? Hoe dienen we dat te begrijpen? Het epistemologisch probleem en het conceptueel probleem lijken opnieuw aan de orde te zijn, zij het dit keer toegepast op andere dieren dan de mens.

Hoe dan ook kan je, afhankelijk van de theorie van bewustzijn die je eropna houdt, aan dieren al dan niet bewustzijn en specifieke mentale toestanden toeschrijven. Maar als je een 'excessief liberale' opvatting hebt van de fysische vereisten voor bewustzijn, bijvoorbeeld dat er al sprake kan zijn van bewustzijn zodra er stofwisseling plaatsvindt, dan zal je ook bewustzijn en mentale toestanden moeten toeschrijven aan veel primitievere vormen van leven. De theorie die alle vormen van leven identificeert met bewustzijnsvormen, wordt *hylemorfisme* genoemd en gaat terug op Aristoteles.

* Zie Wittgenstein (1953).

** Zie Davidson (1982).

Volgens Aristoteles bestaat alles uit twee principes: materie (*hylè*) en vorm (*morphè*). In een concrete substantie zijn materie en vorm onlosmakelijk verbonden, waarbij de vorm de materie organiseert en er structuur aan geeft. Bij levende wezens is de 'ziel' de vorm. Planten hebben een vegetatieve ziel en dieren een sensorische ziel, en de mens heeft een rationele ziel.*

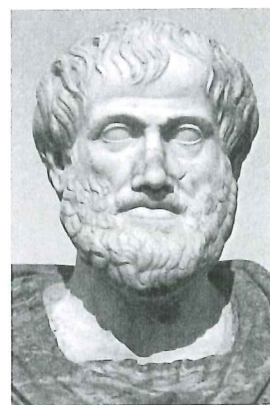
Deze benadering inspireert sommige hedendaagse filosofen, zoals William Jaworski, om een lichtjes 'geactualiseerde' vorm van hylemorfisme naar voren te schuiven als oplossing voor de talrijke problemen waar je in de cognitie- en bewustzijnsfilosofie tegenaan loopt, zodra je een onderscheid maakt tussen mentale en fysische fenomenen.** Het gevolg is wel dat aan alle vormen van leven een vorm van bewustzijn wordt toegeschreven.

Je kan zelfs nog verder gaan dan het hylemorfisme en aan werkelijk alles, dus ook aan objecten, bewustzijn en mentale toestanden toeschrijven. Die opvatting – de tegenhanger van het solipsisme, wat het metafysisch probleem betreft – heet *panpsychisme*. Dat panpsychisme kent een lange en complexe geschiedenis. In de hedendaagse context is het geen specifieke theorie van het mentale, maar veeleer een metatheorie. De redenering van de hedendaagse panpsychist is dat elke manier om het mentale te begrijpen, noodzakelijkerwijs ook de manier is waarop alles – dus ook levenloze objecten – begrepen dient te worden. Het functionalisme bijvoorbeeld mondt volgens de hedendaagse panpsychist noodzakelijkerwijs uit in wat we *panfunctionalisme* zouden kunnen noemen. En hetzelfde geldt voor alle overige theorieën van het mentale: elke theorie van het mentale is meteen een *pantheorie*, een theorie van alles.***

4.7 Andere geesten op een rijtje

Pas met de problematiek van 'andere geesten' lijken de uitdagingen van het scepticisme het scherpst geformuleerd te worden. Als het direct realisme ergens onverdedigbaar lijkt, dan is het wel op het gebied van het bewustzijn en de mentale toestanden van anderen. Met ons eigen bewustzijn en onze eigen mentale toestanden zijn we intiem vertrouwd; we kennen ze 'van binnenuit' door introspectie. Maar het bewustzijn en de mentale toestanden van anderen kennen we slechts 'van buitenaf' door het verbaal of non-verbaal gedrag van anderen. Hoe kunnen we die kloof ooit overbruggen? Hoe kunnen we ooit de koppeling maken tussen wat we als geen ander van binnenuit kennen en beleven, en wat we bij anderen hoogstens kunnen vermoeden?

In paragraaf 4.1 hebben we drie deelproblemen onderscheiden, die we vervolgens hebben geëxploreerd aan de hand van twee sceptische gedachte-experimenten. Nu is het tijd om onze resultaten op een rijtje te zet-



Aristoteles

* Zie Feser (2005), pp. 173-6.

** Zie Jaworski (2011).

*** Zie Skrbina (2005).



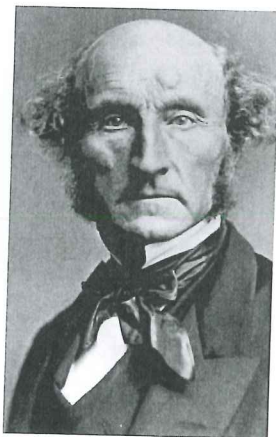
William Jaworski

ten, hoewel we er niet omheen kunnen dat de afzonderlijke deelproblemen en benaderingswijzen in elkaar haken en aldus samen een problematiek vormen.

Het metafysisch probleem

Hoewel geen mens ooit in alle ernst het solipsisme heeft verdedigd, is de cartesische veronderstelling van een geïsoleerd kennistheoretisch subject – een subject dat enkel door middel van introspectie kan doordringen tot de fundamenteën van zijn kennissysteem, met name tot het bestaan en de inhouden van het eigen bewustzijn – intrinsiek solipsistisch te noemen.

Levinas noemt die veronderstelling 'egologisch' en schuift de hele problematiek van 'andere geesten' terzijde, omdat er een direct ethisch appèl uitgaat van 'het gelaat van de Ander'. Dat lijkt een wel erg antropocentrische redenering. Maar zodra we dat antropocentrisme loslaten, dreigen we in een 'excessief liberalisme' te vervallen en staan we weer helemaal opnieuw voor het metafysisch probleem. Geen van de benaderingen van en oplossingen voor het epistemologisch of het conceptueel probleem is beslissend voor het metafysisch probleem: alle opties, met inbegrip van het solipsisme, blijven openstaan. Dat geldt aan het andere eind van het spectrum ook voor het hylemorfisme en het panpsychisme.



John Stuart Mill

Het epistemologisch probleem

De problematiek van 'andere geesten' veronderstelt dat direct realisme op het gebied van het bewustzijn en de mentale toestanden van anderen onmogelijk is. Meer nog dan in het geval van metafysisch scepticisme over de buitenwereld, lijken we aangewezen op een of andere vorm van indirect realisme, bijvoorbeeld Mills voorstel dat we bewustzijn en mentale toestanden toeschrijven aan anderen door middel van analogieredeneringen.

Maar is direct realisme op het gebied van 'andere geesten' werkelijk ondenkbaar en onverdedigbaar? Als er cognitieve mechanismen zijn die ons, zonder tussenstappen, toegang geven tot het bewustzijn en de mentale toestanden van anderen, dan hoeven we geen sceptici ten aanzien van 'andere geesten' te worden en kunnen we volstaan met fallibilisme. Fundamenteeler nog: veronderstelt het epistemologisch probleem niet stiekem dat we wel degelijk het bewustzijn en de mentale toestanden van anderen rechtstreeks kunnen ervaren, zij het dat we daarin niet onfeilbaar zijn? Gaat het begrip van de ander, of van de Ander, niet vooraf aan ons zelfbegrip?

Het conceptueel probleem

Mills oplossing voor het epistemologisch probleem en vergelijkbare pogingen om te komen tot indirect realisme betreffende 'andere geesten', leiden tot het conceptueel probleem: hoe kunnen we op de smalst mogelijk

basis voor inductie en abstractie, namelijk één enkel geval en dan nog een hoogstpersoonlijk geval, begrippen zoals 'pijn' en 'verlangen' vormen die we vervolgens kunnen gebruiken om het bewustzijn en de mentale toestanden van anderen te bevatten?

Volgens Wittgenstein laten we bij het formuleren van dat conceptuele probleem ten onrechte buiten beschouwing hoe psychologische concepten zoals 'pijn' en 'verlangen' concreet gevormd worden. Die begrippen komen niet in isolatie tot stand; het is ondenkbaar dat we een hoogstpersoonlijk privétaalje zouden kunnen ontwikkelen. Psychologische concepten worden veeleer gevormd in een sociale, linguïstische context waarin we – zonder dat we daar erg in hebben, als het ware onbewust – de regels voor het gebruik van die begrippen in samenspraak genereren en vaststellen. Dus ook op het niveau van begripsvorming gaan de anderen in zekere zin vooraf aan onszelf.

Het lijkt er dus sterk op dat ongebreideld scepticisme op het gebied van 'andere geesten' onhoudbaar is. In de woorden van de Engelse dichter John Donne: 'No man is an island.' Of zoals de besproken continentale filosofen benadrukken: de vraag naar 'andere geesten' kan pas geformuleerd worden in een sociale context, aan de hand van een begrippenkader dat men deelt met andere bewuste wezens, met personen die een vergelijkbare binnenwereld en beleavingswereld hebben. Alleen wanneer de probleemstelling volledig wordt opengetrokken en er ook heel andere wezens in het vizier komen, zoals dieren en machines, lijken we nog af te moeten rekenen met metafysische, epistemologische en conceptuele vragen die samen een heuse sceptische problematiek vormen.