

*De Academische Boekengids* 49, maart 2005, pp. 7-9.

**Bioloog Christoph Koch probeert de vraag wat bewustzijn is te beantwoorden aan de hand van reacties die visuele prikkels in de hersenen oproepen. Waarom kan een proefdier dat blind geworden is, denken dat hij toch iets ziet?**

Primaten (mens en aap) zijn echte kijkdieren. Vrijwel overal in hun hersenen vinden we zenuwcellen die reageren op visuele prikkels. De activiteit daarvan hangt samen met de visuele waarneming van objecten en gebeurtenissen, met visueel geleid gedrag of met visuele geheugenprocessen. Bovendien zijn onze lichtzintuigen, samen met de chemische zintuigen, evolutionair de oudste gedragsbepalende systemen. Ze komen in alle stammen van het dierenrijk voor. Als we een mysterieus lijkend verschijnsel als het bewustzijn neurobiologisch willen onderzoeken, ligt het daarom voor de hand te beginnen bij bewuste visuele processen. Dat is een populaire strategie, zodat momenteel het aantal tijdschriftpublicaties over het visuele bewustzijn explosief groeit. Toch zijn er maar weinig monografieën over dit onderwerp verschenen. Christoph Kochs boek *The Quest for Consciousness. A Neurobiological Approach* is dan ook een welkome aanvulling.

Koch is professor in de cognitieve en gedragsbiologie aan het CalTech (California Institute of Technology). Hij publiceert al tien jaar over mogelijke neuronale correlaten van het visuele bewustzijn, vrijwel steeds samen met de pas overleden Nobelprijswinnaar Francis Crick. Koch verwijst dan ook expliciet naar Cricks grote inbreng in de ontwikkeling en rijping van de ideeën die hij in zijn nieuwe boek presenteert. Bij het bespreken van *The quest for consciousness* gebruik ik als achtergrond een boek van Bennett en Hacker dat Herman Philipse eerder in deze gids besprak (*ABG* nr. 42, 2003). Dat is nodig, omdat Bennett en Hacker hun uiterste best hebben gedaan om, zoals Philipse het uitdrukt, de 'Augiasstal van de neurowetenschappen' uit te mesten. Het heeft hen geen citaat in Kochs boek opgeleverd. Ondanks mijn bewondering voor hun werk, zal ik Bennett en Hacker korthedshalve de 'hackers' van de neurobiologische aanpak noemen. Ze leggen namelijk een aantal veiligheidsrisico's bloot voor neurobiologen die zich in dit filosofische mijnenveld begeven.

Het belangrijkste probleem waarop de hackers wijzen, is de zogenaamde meriologische val, het toekennen van eigenschappen van het geheel aan de delen. Veel neurobiologen spreken ten onrechte over het denkende brein of de emotionele hersenen. Alleen mensen en een aantal andere diersoorten worden ervan verdacht te denken, te voelen, te herinneren – hersenen of voeten doen dat niet. Onze hand schrijft niet, de benen lopen niet, de mond eet niet, maar de mens kan wel schrijven, lopen en eten. Dit lijkt misschien muggenzifterij, maar Bennett en Hacker laten zien dat hun bezwaar gewichtiger is. Zo onderscheidde Descartes geestelijke en lichamelijke processen, waarbij de geestelijke processen de lichamelijke bewust in de gaten hielden en stuurden, het zogenaamde interactionistisch dualisme. Fysicalisten (alles is materie/energie, er bestaan geen 'geestelijke' processen) traptten massaal in de meriologische val, door te stellen dat Descartes' geest een hersenproces is; dus dat bewustzijn, gevoelens en herinneringen hersenprocessen zijn. Dan ontstaat het dilemma dat een hersenproces enerzijds een fysisch-chemisch proces is van ionenstromen en elektrische spanningsvariaties, maar anderzijds een bewuste ervaring, zoals een pijn of een emotie. Hoe kan dat? De hackers stellen dat de fysicalisten niet zijn opgeschoten ten opzichte van Descartes, ook al doen ze schamper over hem. De logica van de fysicalisten heeft nog dezelfde dualistische vorm, die rechtstreeks van Descartes is geërfd.

Bennett en Hacker geven aansprekende voorbeelden van die neurobiologische spraakverwarring, zoals bij Crick: 'Wat je ziet is niet wat er echt is, maar wat de hersenen menen dat er is.' Ook een andere neurobioloog, A.R. Damasio (onder andere *Descartes' Error*, 1996), wordt op de korrel genomen, vanwege zijn uitspraak: 'Een bewuste ervaring is het hebben van een beeld of film in de hersenen.'

Het brein ziet natuurlijk geen films of beelden, het heeft immers geen ogen. Zelfs in de primaire visuele cortex (V1), waar een kaartvormige afbeelding van de buitenwereld te vinden is, worden geen films afgespeeld of beelden gepresenteerd. Men spreekt van een corticale 'kaart' als een prikkeling van buurgebieden in de buitenwereld leidt tot activiteit in de buurgebieden van de cortex. Zulke kaarten kun je als enigszins vervormde plaatjes van de buitenwereld opvatten. Maar de volgende hersengebieden kunnen die plaatjesachtige afbeelding in V1 niet 'zien' zoals de neurowetenschapper die ziet als hij met zijn micro-elektroden systematisch de cortex aftast. Hogere visuele cortexgebieden moeten het doen met de activiteitspatronen in de draadjes (axonen) die uit V1 arriveren en met hun onderlinge samenhang in activiteit, de correlatiestructuur.

Dit soort inzichten van onze hackers vinden we overigens al lang in de literatuur. James Gibson (grondlegger van de ecologische psychologie) ventileerde dergelijke kritiek op 'representaties' in de hersenen al sinds het einde van de jaren vijftig van de vorige eeuw. Mijn Utrechtse collega J.J.

Koenderink publiceerde zo'n twintig jaar terug al theorieën over de manier waarop zenuwnetwerken tijdens hun ontwikkeling, uit de correlatiestructuur in de activiteiten, zelf een 'lokaalteken' (positiecode) zouden kunnen opbouwen. Kortom, er bestaan werkbare alternatieven voor de kaart-in-brein, of beeld-in-brein of film-in-brein nonsens. De 'kaartstructuur' lijkt nodig om de verbindingen (axonen) tussen gerelateerde buurgebieden zo kort mogelijk te houden en dus op axonmassa te sparen.

'HET DOEL VAN KOCH IS TE ONTDEKKEN WAT DE KLEINSTE VERZAMELING IS VAN NEURONALE GEBEURTENISSEN EN MECHANISMEN, DIE COLLECTIEF VOLDOENDE ZIJN VOOR EEN BEWUSTE VISUELE ERVARING: DE NEURONALE CORRELATEN VAN DE BEWUSTE WAARNEMING.'

Toch blijkt steeds weer dat werk dat vanuit een verkeerde optiek is gedaan en in hackbare terminologie is beschreven, wel bruikbare inzichten en goede meetgegevens kan opleveren. Vertaling in een vorm die hackers tevredenstelt, blijkt meestal alsnog mogelijk. Hoe zit dat met het werk van Koch (en Crick)? Zij hebben vrijwel vanaf het begin een slimme stap gezet door te spreken over de 'neuronale correlaten van bewustzijn'. Hiermee leggen ze zich niet van tevoren vast op een precieze aard of plaats van de activiteit die met bewustzijn gepaard gaat. Bovendien is impliciet tevens een meetprogramma geformuleerd: kijk welke neuronale activiteiten correleren met bewuste ervaringen (van de eigenaar van die actieve cellen). Met 'neuronaal' wordt bedoeld dat het gaat om neurogroepen, dus zenuwcelgroepen. Koch waagt zich terecht niet aan een definitie van bewustzijn. Er bestaat immers evenmin een sluitende definitie van begrippen als 'gen' of 'leven', maar dat verhindert ons niet die woorden zinvol te gebruiken. Koch bedoelt met bewustzijn de toestand die we allemaal kennen en die we beleven als we wakker zijn en iets ervaren (pijn, liefde, rode appel enzovoorts). Hij staat ons toe eventueel ook dromen mee te tellen, al is daarbij ons geheugen niet 'on line' en correleert ons bewegen (gelukkig) niet met de droominhoud. Het doel is te ontdekken wat de kleinste verzameling is van neuronale gebeurtenissen en mechanismen, die collectief voldoende zijn voor een bewuste (visuele) ervaring: de neuronale correlaten van de bewuste waarneming. Het uitgangspunt is dat iedere ervaren toestandverandering een duidelijke relatie moet hebben met een neuronale toestandverandering.

Koch gaat ervan uit dat evolutionaire continuïteit betekent dat bewustzijn in het dierenrijk wijder verbreid moet zijn. Een neo-behaviorist als de filosoof Daniel Dennett - die door Bennett en Hacker terecht ook uitgebreid wordt gekritiseerd - denkt hier heel anders over. Voor bewustzijn is taal nodig, meent Dennett. Koch heeft zijn eigen, ruimhartiger, standpunt hard nodig, want veel van het werk aan visueel bewustzijn dat hij doet of citeert, is met niet-menselijke primaten gedaan. Daarbij wordt geëist dat het dier een bepaalde ervaring expliciet kan rapporteren, voordat men ervan uitgaat dat dit een *bewuste* ervaring zou kunnen zijn. Dat expliciete rapporteren is ongeveer hetzelfde als een taaluiting bij de mens.

Een fraai voorbeeld hiervan is te vinden in het werk van Alan Cowey en Petra Stoerig, dat Koch beschrijft. Het betreft onderzoek aan blindzien bij aapjes. Iedereen die wel eens een recent filosofisch werkje heeft opgeslagen, zal op de term 'blindzien' zijn gestuit, want filosofen zijn er gek op. Mensen die een primaire visuele cortex (V1) missen (achterhoofdsschors linker- of rechterhersenhelft) zijn in de tegenoverliggende gezichtshelft blind, althans wat bewust zien betreft. Als je ze ertoe brengt te raden of een streep in hun blinde gebied horizontaal of verticaal staat, raden ze dit vrijwel steeds goed. Toch vinden ze de opdracht tamelijk onzinnig, omdat ze immers niks zien. Ze kunnen op dezelfde manier perfect de richting raden van bewegende visuele prikkels, zonder bewust iets te zien bewegen. De verklaring is dat restbanen, om V1 heen, toch nog hogere centra informeren - maar dan zwakjes en alleen over grove prikkels, niet over de fijne details. Die activering is niet sterk genoeg voor bewuste waarneming, maar net wel voor rapportage-op-verzoek, een soort reflexmatige keuze tussen antwoordknoppen.

'HET BREIN ZIET NATUURLIJK GEEN FILMS OF BEELDEN, HET HEEFT IMMERS GEEN OGEN. ZELFS IN DE PRIMAIRE VISUELE CORTEX, WAAR EEN KAARTVORMIGE AFBEELDING VAN DE BUITENWERELD TE VINDEN IS, WORDEN GEEN FILMS AFGESPEELD.'

Bestaat dit verschil in bewuste en niet-bewuste visuele waarneming ook bij aapjes? Om dat na te gaan, verwijderden Cowey en Stoerig eenzijdig V1 bij een aantal aapjes en deden daarna een slim opgezette waarnemingsproef. De aapjes moesten in het midden van een computerscherm blijven staren (fixeren) en drukken op de hoek van het scherm (een 'touch screen') waar telkens kort een prikkel werd getoond. De testmomenten werden met een geluidje aangegeven. De aapjes deden dit uitstekend, ook in hun blinde gezichtshelft. Daarnaast waren ze vóór de operatie getraind om op een aparte knop te drukken als er tijdens de test geen visuele prikkel werd gegeven. De blindzien-aapjes drukten niet alleen steeds in de correcte schermhoeken, maar als de prikkel in hun blinde gezichtshelft lag, drukten ze bovendien nog op de knop 'geen prikkel'. Met andere woorden: hun hand 'wist' waar de prikkels kwamen, maar bewust hadden ze - volgens hun eigen spontane rapportage - niets gezien!

Koch staat uitgebreid stil bij een fenomeen dat hierbij als voorkennis nodig is, namelijk dat we geen bewuste ervaringen hebben van visuele prikkels die bij geoefende handelingen worden gebruikt om deze handelingen te sturen. Als we een veter strikken kijken we ernaar, maar we zijn ons nauwelijks bewust van de diverse aanzichten van de veter terwijl we die strikken. Evenmin weten we bewust hoe we tijdens

het sporten visuele gegevens gebruiken om de vele automatismen (goedgeoefende handelingen) te sturen.

Kochs boek is vlot geschreven en geeft een goed overzicht van de belangrijkste vorderingen en inzichten in de neurobiologie van het zien. Ieder hoofdstuk eindigt met een goede samenvatting, die het gemakkelijk maakt om na te gaan of je wat mist als je het hoofdstuk overslaat. Neurobiologen kunnen dat zeker doen, maar de gemiddelde lezer zal de tekst gewoon van kaptot kapt willen lezen. De beloning is een beter begrip van het zien, en een prikkeling van de geest door Kochs bespiegelingen over neuronale correlaten. Hij gaat ervan uit dat alleen 'expliciete' neuronale representaties tot de bewustzijnsrelaten doordringen, de impliciete niet.

Een kenmerk is expliciet gerepresenteerd als het, na een flink aantal stappen van voorbereiding, een bijeenliggend gespecialiseerd groepje zenuwcellen activeert. Koch noemt de hoeveelheid voorbereiding 'depth of computation', een hackbare terminologie die gelukkig gewoon vermeden kan worden. Het groepje neuronen dat een kenmerk expliciet maakt, moet een essentieel knooppunt zijn. Dat wil zeggen: vernietiging van dat groepje cellen maakt de waarneming van het desbetreffende kenmerk onmogelijk. Voorbeelden zijn objectselectieve neuronengroepen in de temporale cortex. Die ligt aan de zijkant van ons brein, links en rechts achter de slapen. Vernietig daar het juiste clustertje cellen en je bent gezichtsblind, of herkent geen groenten meer. Een aantal essentiële knooppunten kunnen in wisselende coalities samenwerken en zo tot tijdelijke neuronale correlaten van het bewustzijn worden.

Er zijn in de hersenen synchrone activiteitsgolven gevonden die Koch verantwoordelijk houdt voor het ontstaan van steeds weer nieuwe coalities die in onderlinge competitie proberen toegang te krijgen tot de neuronale correlaten van het bewustzijn. Het extremere idee dat synchroon vurende zenuwcellen de neuronale correlaten *zijn*, is nogal omstreden. Koch ziet (tegenwoordig) voor synchroon vuren alleen een rol weggelegd bij het opbouwen van tijdelijke coalities van neuronengroepen. Als er zo'n algemene hausse aan activiteit is (een golftop), worden alle knooppunten die dicht bij drempeloverschrijding zijn net bovendrempelig actief en daarmee tot groep 'gebonden', terwijl de minder actieve knooppunten het niet halen. Zo lezen synchrone golven van moment tot moment af welke knooppunten qua activiteitsniveau bijeenhoren. Bovendrempelig betekent in dit verband dat ze actief genoeg zijn om door de neuronale correlaten van het bewustzijn opgemerkt te worden. Mij overtuigt dit niet, want ik zou zeggen dat de drempel voor activering van die correlaten dan gewoon wat lager had kunnen zijn, zodat de hulp van een golfje activiteit niet nodig was geweest. Dat zou energie sparen, wat overlevingswaarde heeft.

Een ander strijdpunt is of V1, de primaire visuele cortex, neuronale correlaten van bewuste waarneming omvat. Kochs conclusie is negatief. Indirect is V1 natuurlijk van belang, want zonder V1 is er geen bewust zien (alleen blindzien). Maar zonder fotoreceptoren zien we evenmin en daar zoekt niemand visueel bewustzijn. Koch gaat ervan uit dat de primaire rol van bewustzijn bestaat uit samenvatten, beslissen en plannen. Veel onderzoek geeft aan dat de frontale cortex (hersenschors achter het voorhoofd) het neuronale correlaat van die functies bevat. In die buurt moeten we de neuronale correlaten van het bewustzijn volgens Koch dan ook zoeken, en niet in V1. Tijdens wilde visuele dromen van slaapgedepriveerde proefpersonen blijkt V1 bijvoorbeeld geremd en blijken hogere visuele gebieden actief te zijn, zodat die hogere gebieden met de ervaring moeten correleren. Elektrische stimulatie van V1 geeft niet meer te zien dan wat vlekjes en flitsjes, nooit complexe zinvolle visuele ervaringen, terwijl die wel optreden als we hogere visuele cortexgebieden prikkelen. Bij binoculair ruimtelijk zien, zijn we ons er niet van bewust welke informatie door welk oog binnenkomt, terwijl die informatie in V1 aanwezig is. Onderzoek aan de neuronale mechanismen van binoculaire rivaliteit toont aan dat V1-cellen steeds de prikkel in 'hun' oog volgen, terwijl cellen in de temporale cortex de waarneming van de aap weerspiegelen. Dit cruciale experiment bekijken we eerst.

Koch gaat op een aantal gevallen in, waarbij de interpretatie van een prikkel flipt van één mogelijkheid naar een andere. Een bekend voorbeeld is de lijntekening van een kubus, waarbij het voorvlak ineens achtervlak lijkt te worden en dan weer omgekeerd. Binoculaire rivaliteit is onder visuele onderzoekers het meest populaire voorbeeld van zulke bistabiliteit. Hierbij kijkt één oog bijvoorbeeld naar een horizontaal lijnenraster en het andere naar een verticaal lijnenraster. De ogen kunnen gescheiden worden geprikkeld met behulp van een stel spiegels of een kant-en-klare stereoscoop. De proefpersoon ziet dan soms het horizontale, soms het verticale raster of korte overgangsfasen. Doordat de ogen strijdige patronen zien, moeten binoculaire neuronen 'kiezen' voor een interpretatie. Die strijd tussen coalities van cellen die alle mogelijkheden representeren, is het neuronale correlaat van binoculaire rivaliteit. Nikos Logothetis (tegenwoordig verbonden aan het Max Planck Instituut in Tübingen) vond met zijn medewerkers dat cellen in V1 gewoon de prikkel volgen. Dus sommige cellen reageren op horizontale rasters, andere op verticale. En dat doen ze stabiel, de hele tijd, ongeacht wat de aap via een drukknopje rapporteert te zien. In de temporale cortex vallen cellen met horizontale voorkeur stil als de aap het verticale raster ziet en omgekeerd, dus daar correleert de activiteit wel met de bewuste waarneming.

'ER ZIJN VEEL HERSENMODULEN DIE ONGEMERKT, DUS NIET-BEWUST, HUN WERK DOEN. DE TAALMODULEN, BIJVOORBEELD, DIE AUTOMATISCH "VERWOORDEN" WAT WE WILLEN ZEGGEN EN HET ONS MET DE JUISTE INTONATIE LATEN UITSPREKEN, SOMS TOT ONZE EIGEN VERBAZING.'

Het spoor van de neuronale correlaten van bewustzijn voert ons zodoende naar de temporale cortex. Daar ook vinden we gebieden vol 'essentiële knooppunten'. Vooral in de onderste delen van de temporale cortex zijn wonderlijke neuronen gevonden. Zo is er een cel die alleen reageerde op plaatjes van Bill Clinton, ongeacht Bills hoofdoriëntatie, lichaamshouding, kleding of beweging. Ondanks dit soort 'gnostische' (= kennende) neuronen gaat Koch ervan uit - net als vrijwel alle collega's - dat kennis nooit, of vrijwel nooit, is vastgelegd in individuele cellen, maar altijd in groepen van cellen, in coalities. Zijn de coalities in de temporale cortex dan de gezochte correlaten van het bewustzijn? Nee, de interactie met de frontale cortex is onontbeerlijk. Daar zetelt onder andere het werkgeheugen - en zonder werkgeheugen geen bewustzijn, zegt Koch terecht. Zo komen we in het onzekerste deel van de bespiegelingen: Koch stelt dat interactie met de frontale cortex onontbeerlijk is, maar betoogt tevens dat de neuronale correlaten van het bewustzijn niet zomaar in modulen in de frontale cortex gezocht moeten worden. Immers: plannen, denken, beslissen - waarbij de frontale cortex een dominante rol speelt - gebeurt grotendeels niet-bewust. De plannen en besluiten borrelen zomaar op, zonder dat we ons bewust zijn van de onderliggende strategische en op ervaring gebaseerde afwegingen. De frontale cortex is het neuronale correlaat van die niet-bewuste rationaliteit.

Er zijn veel hersenmodulen die ongemerkt, dus niet-bewust, hun werk doen. De taalmodulen, bijvoorbeeld, die automatisch 'verwoorden' wat we willen zeggen en het ons met de juiste intonatie laten uitspreken, soms tot onze eigen verbazing. Deze feiten tonen al aan dat de neuronale correlaten van het bewustzijn in relatief bescheiden delen van de hersenen gezocht moeten worden. Bewustzijn is dus zeker geen holistisch verschijnsel. Koch noemt de hersendelen die voor niet-bewuste activiteiten verantwoordelijk zijn 'zombiemodulen'. Zombies zijn in de filosofische literatuur hypothetische mensachtigen die door hun gedrag niet van andere mensen zijn te onderscheiden, maar die elke vorm van bewustzijn missen. Ze worden meestal ten tonele gevoerd om te betogen dat bewustzijn een epifenomeen is. Maar een mens zonder bewuste pijnwaarneming leeft niet lang, dus pijnbewustzijn is zeker nodig. Blindzien helpt niet om een glas water op te pakken als je dorst hebt, dus bewust zien lijkt ook nodig voor het overleven. Kortom, de zombie als gedachte-experiment is nonsens. Koch vindt dat ook, maar leent de term helaas toch voor niet-bewuste functies van het zenuwstelsel. De planning- en strategiemodulen in de frontale cortex zijn dan zombiemodulen, omdat we geen bewuste beleving van hun werking hebben, alleen van de resultaten van die werking. Ergens in de mêlee van spelers en interacties in de temporale en frontale cortex wordt zo nu en dan een visuele ervaring gescoord, hoewel vele spelers zombies zijn.

In het laatste hoofdstuk van *The Quest for Consciousness* staat een fictief interview, waarin gevraagd wordt of het vinden van de neuronale bewustzijnsrelaten het bewustzijnsprobleem oplost. Koch antwoordt driewerf nee! Toch vergelijkt hij ze vervolgens qua belang met de 'double helix' van het DNA en ontlokt zo aan de fictieve interviewer (dus hemzelf) de opmerking: 'A fond dream.' Ik vind dat wel een mooie samenvatting, want dit is een prettig boek over een prettige droom.

**Wim van de Grind** is emeritus hoogleraar neurobiologie aan de Universiteit Utrecht.

### **Besproken boeken:**

*The Quest for Consciousness. A Neurobiological Approach*  
door **C. Koch**  
Roberts & Company. Englewood, Colorado 2004.  
432 pag., € 50,38

### **Literatuur:**

- M.R. Bennett** en **P.M.S. Hacker** (2003). *Philosophical foundations of neuroscience*. Blackwell Publishers: Oxford.
- H. Philipse** (2003). 'De Augiasstal van de neurowetenschappen', *De Academische Boekengids* nr. 42.