

Bine ați venit într-o călătorie fascinantă la bordul unei caravele filosofice, condusă de autor prin „mări niciodată explorate până acum”.

Această odisee prin misterele minții reunește contribuțiile a opt dintre cei mai străluciți experți internaționali în domeniul creierului și conștiinței, inclusiv laureatul Premiului Nobel Sir Roger Penrose.

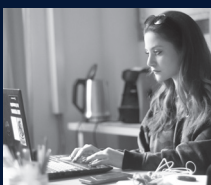
Cartea este structurată în două părți: prima, în care Steven S. Gouveia trasează coordonatele fundamentale ale filosofiei și științei minții; a doua, care prinde viață prin dialogurile sale cu cei opt savanți, oferind cititorului o experiență unică de explorare intelectuală. Pregătiți-vă să porniți într-o aventură captivantă, în care știința și filosofia se întâlnesc pentru a desluși misterele conștiinței.



Steven S. Gouveia este licențiat în Filosofie. La 22 de ani a început un doctorat în Filosofia Minții, cu accent pe Neuroștiință, la Universitatea din Minho. La 30 de ani a fost numit Profesor Onorific la Facultatea de Medicină, Universidad Andrés Bello, Chile. În prezent, este cercetător la Institutul de Filosofie al Universității din Porto, unde coordonează un proiect despre Etica Inteligenței Artificiale în Medicină. A participat la conferințe și emisiuni media internaționale. Mai multe informații: www.stevensgouveia.weebly.com.



Tudor Lucian este psiholog clinician și psihoterapeut acreditat de IPEFE în Terapia Focalizată pe Emoții. A absolvit Facultatea de Psihologie la Universitatea din București și a urmat un master în Psihologie Clinică. Este co-fondatorul unei platforme de gândire critică ce pregătește elevii pentru admiterea la facultățile de psihologie și drept din România.



Gaia Amoroso, studentă ABAPA - Academia de Arte Plastice Arte din Palermo (Italia), absolvent în Design Grafic, specializată în grafică Editorial, ilustrație și fotografie.

STEVEN S. GOUVEIA

ODISEEA MINTII



STEVEN S. GOUVEIA

Susan Blackmore - Sir Roger Penrose
- Nicholas Humphrey - Karl Friston
- Joseph LeDoux - David Chalmers
- Christof Koch - Anil Seth

ODISEEA MINTII

Dialoguri despre creier și conștiință



ODISEEA MINȚII

Dialoguri despre

CREIER

și

CONȘTIINȚĂ

Steven S. Gouveia



2025

Steven S. Gouveia
MLAG | Institute of Philosophy | University of Porto

ISBN: 9798282993554

© Autorul și colaboratorii, sub licență exclusivă.

Această lucrare este supusă dreptului de autor și drepturilor de traducere. Toate drepturile sunt licențiate exclusiv de către autor și colaboratori, fie în întregime, fie în parte din material. Aceasta include drepturile de traducere, retipărire, reutilizare a ilustrațiilor, recitare, diseminare, reproducere pe microfilm sau sub orice altă formă fizică și transmiterea sau stocarea și preluarea informațiilor, adaptarea electronică, software-ul de calculator sau prin metodologie similară sau diferită acum cunoscută sau dezvoltată ulterior.

Utilizarea numelor descriptive generale, a numelor înregistrate, a mărcilor comerciale, a mărcilor de servicii etc. în această publicație nu implică, chiar și în absența unei declarații specifice, că astfel de nume sunt scutite de legile și reglementările relevante de protecție și sunt, prin urmare, gratuite pentru uz general. Autorul și colaboratorii pot presupune cu siguranță că sfaturile și informațiile conținute în această carte sunt considerate a fi adevărate și exacte la data publicării.

Nici autorul, nici colaboratorii nu oferă nicio garanție, expresă sau implicită, cu privire la materialul conținut aici sau pentru orice erori sau omisiuni care ar fi putut fi făcute. Autorul rămâne neutru în ceea ce privește pretențiile juridiciale în hărțile publicate și afilierea instituționale.

Coperta: © Steven S. Gouveia.

Traducător: Lucian Tudor.

Desene originale de pe copertă: © Ana Monteiro.

Transcrieri: © Tássia Vianna.

Editor: Kindle Direct Publishing (Amazon).

Publicație originală: A Odisseia da Mente (portugheză, 2024)

Cuprins

Prefață _____ 7

Partea I | Conștiința

I. Introducere: Conștiința _____ 13

II. Misterul conștiinței _____ 25

III. Conștiința alterată _____ 39

IV. Conștiința evoluată _____ 55

V. Conștiința cuantică _____ 65

Dialoguri I | Conștiința

VI. Dialog cu David Chalmers _____ 83

VII. Dialog cu Susan Blackmore _____ 97

VIII. Dialog cu Nicholas Humphrey _____ 119

IX. Dialog cu Sir Roger Penrose _____ 141

Partea I | Creier

I. Introducere: Creierul _____ 163

II. Creierul predictiv _____ 187

III. Creierul liber _____ 201

IV. Creierul integrat _____ 211

V. Creierul emoțional _____ 227

Dialoguri II | Creier

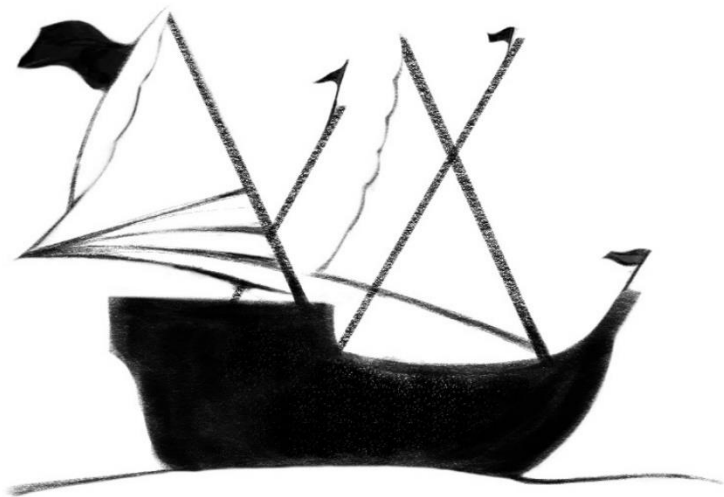
VI. Dialog cu Anil Seth_____	241
VII. Dialog cu Karl Friston_____	261
VIII. Dialog cu Christof Koch_____	279
IX. Dialog cu Joseph LeDoux _____	297
Concluzie _____	315
Mulțumiri _____	318
Biografia autorului_____	323

Prefață

Scopul acestei cărți este de a aduce cititorul la bordul unei caravele filosofice,¹ alături de unii dintre cei mai mari navigatori ai secolului al XXI-lea, într-o călătorie plină de curaj prin mărire învolburate ale cunoașterii problemelor minții umane.

Pe măsură ce ridicăm "pânzele curiozității", vom găsi cu siguranță cel puțin o insulă, unde ne putem opri temporar călătoria. Poate că fiecare sinapsă din creierul nostru poate servi drept busolă, ca cea din centrul copertei acestei cărți, care indică direcții neexplorate și regiuni favorabile unei înțelegeri mai aprofundate a minții umane.

¹ Caravela portugheză s-a remarcat prin mai multe inovații, cum ar fi utilizarea unei vele lateen triunghiulare, care permitea o mai mare capacitate de a naviga împotriva vântului, sau coca sa ușoară și agilă, care îi permitea să fie foarte versatilă. În plus, portughezii au dezvoltat, de asemenea, o tehnologie de navigație mai avansată decât cea utilizată până atunci, prin utilizarea arbaletului și a astrolabului, ceea ce le-a dat posibilitatea de a naviga în locuri din lume care nu fuseseră niciodată explorate.



Imaginea 1: "Caravella" © Gaia Amoroso

În această căutare neîntreruptă ne confruntăm cu o problema centrală: ce îi oferă conștiinței unicitatea ei?

Pe fiecare pagină a acestei cărți, exemple vii de idei și experiențe așteaptă cititorul, încurajându-l să-și creeze propria caravelă și să exploreze „mările ce nu au fost niciodată navigate înainte”.²

² Referință la primul vers din Canto I al operei *Os Lusíadas*, scrisă de Luís Vaz de Camões, poet portughez care a descris aventurile descoperirilor maritime ale portughezilor în secolul al XV-lea. Expresia originală este “Por mares nunca dantes navegados”.

Mintea, ca și oceanul, este vastă și neexplorată, iar căutarea noastră este ghidată de pasiunea descoperirii și de setea de a înțelege un mister neîncheiat, care pare să chinuie fiecare secol al umanității.

Să sperăm că, la finalul acestei călătorii, am reușit să găsim un port bun, care să ne permită măcar o scurtă odihnă binemeritată.

Pentru a încheia această prefață, o notă mai puțin poetică: această carte are o structură specifică, fiind împărțită în două părți: prima parte, dedicată conștiinței, și a doua parte, dedicată creierului.

În fiecare parte, veți găsi două secțiuni: a doua va prezenta dialoguri cu 8 experți de renume internațional pe subiectele abordate, în timp ce prima va include introduceri tematice. Aceste introduceri au scopul de a ajuta fiecare cititor să înțeleagă mai ușor dialogurile, în ciuda complexității temelor prezentate în această carte.

Prin urmare, să înceapă această Odissee a Minții.

Steven S. Gouveia
Porto, Portugalia
05 | 05 | 2025

PARTEA I

Conștiința

PARTEA I

Conștiința

I. INTRODUCERE: CONȘTIINȚA

Imaginează-ți că te afli, chiar acum, pe una dintre cele mai spectaculoase plaje din lume: faimoasa Copacabana din Rio de Janeiro, Brazilia. În timp ce savurezi o *caipirinha* răcoritoare (băutura națională a Braziliei), te bucuri de soarele strălucitor și de priveliștea impunătoare a Muntelui Sugarloaf.

Această experiență este, fără îndoială, magică, dar ceva și mai fascinant se petrece în acest moment. Ce anume? Minunea de a trăi această clipă dintr-o perspectivă unică: propria ta conștiință!

Nimeni nu poate explica pe deplin cum ia naștere experiența conștientă. Cum este posibil ca un ansamblu de neuroni inconștienți să genereze o senzație unificată, subiectivă, a realității și să creeze o minte conștientă?

Cercetători din diverse domenii au folosit metode și abordări variate pentru a contura o imagine clară a modului în care mintea umană este modelată de creier, corp și interacțiunea cu mediul înconjurător. Conștiința rămâne, poate, unul dintre ultimele mari mistere ale lumii moderne.



Imaginea 2: "Architettura interna" © Gaia Amoroso

La fel ca în trecut, când umanitatea a încercat să înțeleagă originea universului, natura timpului, spațiului sau gravitației, conștiința a ridicat întrebări fundamentale la care încă nu avem un răspuns definitiv. Spre deosebire de aceste enigme, studiul conștiinței este marcat de confuzie profundă, dezbateri aprinse și numeroase teorii contradictorii.

Recent, am avut privilegiul de a susține pentru a doua oară un discurs la una dintre cele mai prestigioase conferințe academice dedicate conștiinței: „Știința Conștiinței”. Prima participare a fost în 2019, la Interlaken, Elveția, iar cea mai recentă, în mai 2023, în

Taormina, un oraș pitoresc din Sicilia, cu o priveliște magnifică asupra vulcanului Etna.

În prima zi a conferinței, fiecare participant a primit un cod unic și un link pentru a vota teoria conștiinței pe care o considera cea mai plauzibilă. Spre surprinderea mea, pagina de vot conținea nu mai puțin de 13 teorii diferite! Dacă există atât de multe explicații concurente pentru același fenomen, este clar că încă nu am găsit răspunsul corect.

De ce este atât de dificil, atât pentru știință, cât și pentru filosofie, să formuleze o teorie convingătoare despre conștiință? Cu toții avem experiențe subiective: percepții, emoții, senzații și gânduri. Cum pot niște structuri fizice, cum sunt neuronii, să creeze acest univers interior bogat? Tocmai această dificultate face din conștiință una dintre cele mai profunde și captivante probleme ale științei moderne.

Unii cercetători susțin că misterul conștiinței nu este atât de insondabil pe cât pare. Privind înapoi în istoria științei, descoperim exemple în care concepte odinioară obscure au fost elucidate sub o nouă terminologie, deși ele au fost inițial gândite în mod eronat datorită capcanelor limbajului folosit.³

³ Churchland, P. (1986) *Neurophilosophy: Toward a Unified Science of the Mind/Brain*, Cambridge: The MIT Press.

Un astfel de exemplu este studiul luminii. În Antichitate, teoria corpusculară a luminii, inspirată din atomismul grec, descria lumina ca fiind alcătuită din particule numite „corpuscule”. Această idee a fost susținută de gânditori de renume, inclusiv de Sir Isaac Newton, care credea că fenomenele de reflexie și refracție pot fi explicate doar dacă lumina este compusă din particule. Alternativa – lumina ca undă – părea incompatibilă cu observarea traiectoriilor rectilinii ale luminii.⁴

Newton și-a susținut teoria prin aproximativ 40 de experimente ce păreau să demonstreze natura corpusculară a luminii. Aceasta a fost acceptată timp de secole, până când, în 1810, Thomas Young⁵ a efectuat celebrul său experiment cu fantă dublă, infirmând ipoteza corpusculară și demonstrând că lumina are un comportament ondulatoriu.⁶

Conform acestei noi perspective, lumina era mai apropiată de sunet, fiind compusă din unde electromagnetice.

Teoria undelor a reușit să explice fenomene precum interferența și difracția luminii, aspecte pe care modelul

⁴ Ediție originală: Newton, I. (1704) *Opticks: or, A Treatise of the Reflexions, Refractions, Inflexions and Colours of Light*, Londra.

⁵ P. Rowlands, P. (2017) *Newton and Modern Physics*, London: World Scientific.

⁶ Ediție originală: Young, T. (2007) *Miscellaneous Works of the Late Thomas Young: Including His Scientific Memoirs*, Montana: Kessinger Publishing.

corpuscular nu le putea justifica. Astfel, conceptul inițial de „lumină” a fost redefinit ca radiație electromagnetică, iar teoria corpusculară a fost înlocuită de una mai precisă. Această tranziție ilustrează un principiu important: unele concepte științifice, oricât de logice par inițial, pot fi înlocuite cu altele mai fundamentale, care oferă o înțelegere mai intrinsecă a fenomenului studiat.

S-ar putea să ne gândim la acest exemplu reduționist ca la un caz izolat din istoria științei, un concept considerat că având sens privind existența anumitor fenomene, dar care până la urmă, a fost descoperit ulterior ca fiind doar un pseudoconcept.

Potrivit celor care consideră că vechiul concept de lumină din teoria corpusculară este comparabil cu cel al conștiinței, există numeroase exemple care pot fi aduse în discuție.

Teoria flogisticului, dezvoltată în secolul al XVIII-lea de chimistul Georg Stahl, susținea că toate corpurile combustibile conțineau un element numit „phlogiston”, care era eliberat în aer în timpul arderii. Această teorie își avea rădăcinile în ideile mentorului său, Johann Becher, care a publicat lucrarea *Physica Subterranea*, sugerând că un anumit element se degaja în timpul

procesului de combustie.⁷ Teoria flogisticului reprezenta o evoluție față de concepția antică a celor cinci elemente fundamentale, propusă de Empedocle: foc, pământ, aer, apă și eter.

Stahl și colegii săi credeau că teoria flogisticului avea o bază empirică, bazându-se pe observații din metalurgie. Materialele care ardeau sufereau coroziune, iar gradul acesteia era interpretat ca dovadă a pierderii „phlogiston-ului”. Teoria explica și imposibilitatea arderii în lipsa aerului: fără aer, phlogiston-ul nu putea fi absorbit. Sfârșitul procesului de ardere era atribuit epuizării acestui element din material.

Însă teoria flogisticului a fost treptat pusă sub semnul întrebării. O problemă majoră era că, deși materialele organice își pierdeau masa prin ardere, metalele, în procesul de calcinare (oxigenare), câștigau în greutate – contrazicând așteptările teoriei. Antoine Lavoisier a demonstrat acest fenomen, arătând că metalele se combină cu oxigenul în timpul arderii, ceea ce duce la creșterea masei lor, infirmând astfel ipoteza flogisticului.⁸

Descoperirea oxigenului de către Joseph Priestley în 1774 a fost un pas necesar în această tranziție. Priestley

⁷ Ediție originală: Taylor, S. (2010) *Alchimiștii, fondatori ai chimiei moderne*, Montana: Kessinger Publishing.

⁸ Ediție originală: Bell, M. (2005) *Lavoisier in the Year One*, New York: Atlas Books.

a identificat oxigenul ca element responsabil de ardere, înlocuind astfel phlogiston-ul ca explicație pentru acest proces.⁹

Pe măsură ce chimia elementelor a evoluat, teoria flogisticului a fost abandonată complet, devenind un exemplu clasic de concept științific depășit.

Un proces similar s-a petrecut în cazul teoriei fluidului caloric. În secolul al XVIII-lea, căldura era considerată un fluid invizibil, inodor, prezent în toate corpurile organice. Conform acestei teorii, numită „teoria calorică”, temperatura unui corp era direct proporțională cu cantitatea acestui fluid. Paradoxal, această idee a fost susținută chiar de Lavoisier, cel care contribuise la respingerea teoriei flogisticului.¹⁰ Conform acestei teorii, fluidul caloric ar fi avut o existență finită în cantitatea sa, transferându-se de la corpuri mai fierbinți la corpuri mai reci. Acesta nu putea fi creat sau distrus, așa că teza centrală a acestei teorii a fost conservarea sa constantă, care explica interacțiunile temperaturii dintre diferite corpuri.

Incapacitatea teoriei calorice de a explica diverse fenomene, cum ar fi evaporarea și sublimarea, a determinat apariția unei teorii mai promițătoare –

⁹ Ediție originală: Jackson, J. (2005) *A World on Fire: A Heretic, An Aristocrat and The Race to Discover Oxygen*, New York: Viking.

¹⁰ Ediție originală: Fox, R. (1971) *The Caloric Theory of Gases*, Clarendon Press: Oxford.

teoria cinetică – prin lucrarea lui Benjamin Thompson, cunoscut și sub numele de Conte de Rumford.¹¹

Rumford a observat că ceea ce numim “căldură” nu este o substanță materială cu o formă fluidă, ci mai degrabă ar trebui considerată ca o formă de energie în mișcare. Prin experimentele care au demonstrat căldura ca fiind generată prin frecare, el a contestat astfel ideea predominantă că aceasta este un fluid indestructibil.

Unul dintre cele mai interesante experimente ale sale a implicat frecarea generată de forarea tunurilor, timp în care a observat că de fapt, cantitatea de căldură produsă depășea cu mult ceea ce teoria calorică ar fi prezis.

Descoperirile sale au pus bazele teoriei cinetice a căldurii și au contribuit la dezvoltarea principiului conservării energiei. Astfel, conceptul de fluid caloric a fost eliminat treptat, fiind înlocuit de o înțelegere mai precisă a energiei termice.

Erori conceptuale similare pot fi observate și în istoria psihiatriei. Timp de secole, teoria posedării demonice a

¹¹ Ediție originală: Brown, G. (2001) *Count Rumford: The Extraordinary Life of a Scientific Genius – Scientist, Soldier, Statesman, Spy*, Gloucestershire: Sutton Publishing.

fost acceptată ca explicație pentru diverse tulburări psihice.¹²

În Evul Mediu, credința în demoni a influențat profund percepția asupra bolilor mintale, iar exorcismul era considerat un tratament legitim. Documentele istorice arată o corelare între anumite tipuri de „posesiuni” și simptome ale unor afecțiuni precum epilepsia sau tulburările psihotice.

Filosofi medievali importanți au contribuit la formularea acestor idei. Augustin de Hipona considera că demonii posedau un corp material, în timp ce Toma de Aquino îi descria ca entități non-corporale, separate de lumea fizică. Aceste concepții au influențat viziunea dominantă asupra bolilor mintale timp de secole.

Odată cu avansarea psihiatriei bazate pe dovezi, explicațiile demonologice au fost înlocuite de teorii neurobiologice și psihologice solide. Totuși, la fel ca în cazul flogisticului sau al fluidului caloric, tranziția de la o viziune mistică la una științifică a fost un proces îndelungat și dificil.

Concluzia este că multe concepte odinioară considerate fundamentale, de la phlogiston și fluidul caloric, la posedarea demonică, s-au dovedit a fi erori

¹² Ediție originală: Kemp, S. e Williams, K. (1987) “Demonic possession and mental disorder in medieval and early modern Europe”, *Psychological Medicine*, 17 (1):21-9.

conceptuale, care au fost abandonate pe măsură ce cunoașterea umană a progresat. Aceste exemple ne arată că, de-a lungul istoriei, teoriile dominante pot fi reconsiderate și înlocuite atunci când noile dovezi demonstrează limitele sau incoerențele acestora.



Imaginea 3: "I sensi" © Gaia Amoroso

Atunci când realizăm că anumite concepte nu sunt utile pentru înțelegerea realității, întrebările și problemele formulate pe baza acestora își pierd sensul, iar necesitatea unui răspuns devine irelevantă.

Pentru a înțelege mai bine acest principiu, să analizăm teoria geocentrică, conform căreia Pământul era considerat centrul universului. În acest model, Soarele și Luna ar fi orbitat zilnic în jurul Pământului, în timp ce stelele ar fi rămas fixe pe o „sferă cerească” ce se rotește în jurul axei Pământului. Pe baza acestei perspective, filosofi și gânditori ai vremii se confruntau cu întrebări precum:

- Ce determină sfera cerească să se rotească în jurul Pământului?
- Cum sunt fixate și atașate stelele de această sferă?
- Prin ce mecanism își mențin Luna și Soarele orbitele precise în jurul Pământului?

Aceste întrebări au fost considerate provocări importante timp de secole, însă odată cu discreditarea modelului geocentric, inițiată de Copernic și susținută ulterior de Galileo și Kepler, împreună cu mecanica clasică a lui Newton, ele și-au pierdut relevanța.

În cadrul modelului heliocentric, ideea unei „sfere cerești” a devenit lipsită de sens, iar ceea ce părea odinioară o enigmă majoră nu mai necesita nicio explicație.

Unii filosofi și oameni de știință susțin că o situație similară s-ar putea aplica unor întrebări despre conștiință. Formularea unor astfel de întrebări poate

presupune existența unor concepte fundamentale greșite sau prost definite.

Comparând cu teoriile istorice care s-au dovedit a fi pseudoconcepte, este posibil ca și întrebările noastre actuale despre conștiință să necesite o reevaluare critică a ipotezelor și a coerenței lor.

Astfel, putem reflecta asupra unor întrebări esențiale:

- Cum se face că știm atât de puțin despre ceva intrinsec naturii umane?
- Se poate ca știința contemporană să întâmpine dificultăți în explicarea minții?
- Există o teorie a conștiinței suficient de solidă pentru a explica toate fenomenele sale?

În prima parte a cărții, vom încerca să răspundem acestor întrebări prin dialoguri cu patru gânditori proeminenți în studiul conștiinței, precum David Chalmers, Susan Blackmore, Nicholas Humphrey și Sir Roger Penrose, cu care am avut onoarea de a dezbate aceste teme.

Pentru a facilita înțelegerea dialogurilor din a doua parte a fiecărei secțiuni, următoarele capitole vor oferi o introducere concisă în temele explorate. Vom începe prin a analiza de ce conștiința păstrează această aură enigmatică și în ce măsură poate oferi explicații plauzibile despre sine.

II. MISTERUL CONȘTIINȚEI

Primul dialog explorează ideile filosofului David Chalmers, cel care a adus conștiința în prim-planul discuțiilor filosofice și științifice contemporane. Profesor de filosofie și științe neuronale, co-director al Centrului pentru Minte, Creier și Conștiință de la Universitatea din New York, Chalmers este unul dintre principalii susținători ai ideii că natura conștiinței ridică provocări deosebit de complexe.

După cum vom vedea în discuție, Chalmers a introdus conceptul de „problema grea a conștiinței”. Deși această problemă nu este nouă, ea reprezintă o preocupare principală pentru cei care recunosc problema minte-corp ca fiind una legitimă. Această perspectivă este susținută de cei care cred că mintea posedă proprietăți ce nu pot fi explicate integral prin mecanismele fizice ale lumii.

Multe fenomene mentale, precum învățarea, raționamentul sau memoria, pot fi explicate în termeni de „rol funcțional” deoarece odată ce înțelegem funcția unui sistem, putem explica comportamentul său. Însă, pentru Chalmers, acest principiu nu se aplică în cazul conștiinței:

„Ceea ce face această problemă dificilă și aproape unică este că depășește întrebările legate de

performanța funcțiilor cognitive. Dacă am explica toate funcțiile cognitive și comportamentale asociate experienței – discriminare perceptivă, categorizare, acces intern, raport verbal – rămâne totuși întrebarea fundamentală: de ce aceste funcții sunt însoțite de o experiență conștientă?”¹³

Dacă acceptăm realitatea acestei probleme, una dintre consecințele posibile este că o explicație științifică completă a conștiinței ar putea fi imposibilă.

Un alt gânditor preocupat de analiza problemei conștiinței este filosoful Thomas Nagel. Acesta consideră că problema fundamentală a conștiinței este „subiectivitatea” stărilor mentale conștiente. El susține că faptele despre astfel de stări sunt în mod inerent subiective și pot fi înțelese doar din perspectiva directă a celui care le trăiește.¹⁴

De exemplu, în scenariul de la începutul cărții, doar cititorul are acces direct și subiectiv la experiența de a savura o caipirinha pe plaja Copacabana. Oricine altcineva, privindu-l din exterior, nu poate înțelege pe deplin acea trăire.

Totuși, știința necesită o caracterizare obiectivă a faptelor, independentă de un punct de vedere

¹³ Ediție originală: Chalmers, D. (1996) *The Conscious Mind*, New York: Oxford University Press, p. 202.

¹⁴ Ediție originală: Nagel, T. (1974) “What Is It Like to Be a Bat?”, *The Philosophical Review*, 83 (4): 435-450.

subiectiv. Acest aspect ridică întrebarea dacă faptele despre conștiință pot fi vreodată explicate complet științific, ceea ce ar face „problema minte-corp cu adevărat insolubilă.”¹⁵

Chalmers introduce un argument original pentru a ilustra acest punct de vedere, sub numele de „argumentul filosofic al zombilor”.

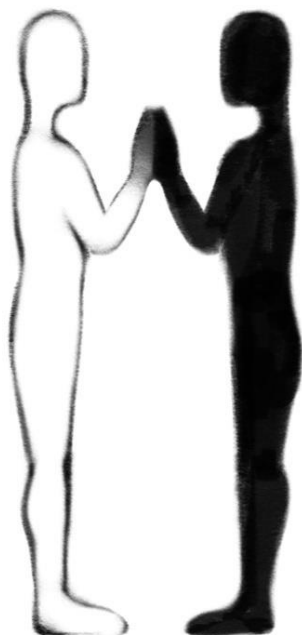
Acest argument explorează posibilitatea logică a unei lumi identice cu a noastră, dar lipsită de conștiință. În această lume, fiecare om ar avea un dublu zombi, un fel de copie identică la nivel molecular, dar care nu experimentează nimic conștient. Deși ar reacționa exact ca o ființă conștientă, acest zombi nu ar avea experiențe interioare.

Gândește-te la acest scenariu ipotetic: te întâlnești pe stradă cu o celebritate pe care o admiri de mult timp. Îți vorbește, îți zâmbește, se comportă exact cum te-ai aștepta. Dar, fără ca tu să știi, această persoană este un zombi filosofic. Nu trăiește nicio experiență conștientă, ci doar manifestă comportamentul unei ființe conștiente. Dacă un astfel de lucru ar fi posibil, ce înseamnă asta pentru natura conștiinței?

În acest scenariu fictiv, deși puțin probabil în realitate (mai ales dacă implică o celebritate), individul respectiv

¹⁵ Ediție originală: Nagel, T. (1974) “What Is It Like to Be a Bat?”, *The Philosophical Review*, 83 (4): 435.

nu are sentimente autentice și nici o experiență conștientă reală. Din perspectiva unui observator, el pare să posede conștiință doar prin comportamentul său exterior.



Imaginea 4: "Gemelli" © Gaia Amoroso

Pornind de la un astfel de exemplu, David Chalmers argumentează că existența „zombilor filosofici” este logic posibilă, chiar dacă poate fi imposibilă din punct de vedere natural. Cu alte cuvinte, chiar dacă am înțelege perfect organizarea funcțională a ființei umane, aceasta nu ar implica automat și prezența

conștiinței. Astfel, un izomorf din silicon al cititorului, cum ar fi un sistem artificial în care cipurile înlocuiesc neuronii, ar putea funcționa identic cu un om, dar fără să aibă experiență subiectivă.

Dacă această idee este coerentă din punct de vedere conceptual și posibilă logic, rezultă că stările conștiente nu sunt identice cu stările fizice ale creierului.¹⁶

Această concluzie ne conduce la un alt filosof, Joseph Levine, care susține că existența conștiinței în lumea fizică ridică problema unui „decalaj explicativ”.¹⁷

Levine argumentează că între conștiință și procesele fizice ale creierului există un decalaj explicativ ireductibil. O explicație științifică riguroasă ar trebui să permită deducerea fenomenului studiat dintr-un set de legi și condiții inițiale. De exemplu, dacă știm că fulgerul este o descărcare electrică, atunci, cunoscând toate condițiile necesare unei astfel de descărcări la un moment dat, putem deduce cu certitudine că fulgerul a avut loc.

Cu toate acestea, susțin atât Levine, cât și Chalmers, acest tip de deducție nu pare posibil în cazul conștiinței.

¹⁶ Din punct de vedere istoric, experimentul de gândire al zombilor filosofici are scopul de a respinge teoria fizicalistă a minții, care susține că stările conștiente pot fi reduse la stări fizice specifice, în relații de identitate strictă.

¹⁷ Ediție originală: Levine, J. (1983) "Materialism and qualia: The explanatory gap", *Pacific Philosophical Quarterly*, 64: 354-361.

Indiferent cât de bine înțelegem mecanismele creierului și legile fizicii, rămâne mereu o întrebare deschisă: este conștiința prezentă în aceste procese sau nu?

Desigur, mulți fiziciști, din cei care consideră că tot ce există trebuie explicat prin legile fizicii și chimiei, resping această idee. Ei susțin că o cunoaștere exhaustivă a mecanismelor fizice ar include, inevitabil, și o explicație completă a conștiinței. Pentru ei, conștiința nu ar fi altceva decât un produs emergent al unor procese fizice pe care, deocamdată, nu le înțelegem pe deplin, dar care ar putea fi elucidate în viitor.

Pentru a ilustra mai bine acest „mister al conștiinței”, putem apela la un alt experiment de gândire intitulat “celebra poveste a doctorului Mary”.

Mary este cel mai strălucit specialist în neuroștiințe pe care l-a cunoscut vreodată lumea. Ea deține cunoștințe absolute despre toate proprietățile fizice ale creierului, în special despre cele implicate în percepția vizuală a culorii. În plus, este expertă în fizica luminii și a culorilor.

Prin urmare, putem afirma că Maria știe tot ce este de știut despre felul în care ființele umane experimentează vizual culorile, cuprinzând atât procesele neuronale implicate, cât și procesele fizice.

Există, totuși, o mică problemă. Mary este prizonieră încă de la naștere într-o casă în care are acces doar la

camera ei. Mai rău decât atât, această cameră nu are nicio lumină. Toate cunoștințele pe care Mary le-a învățat despre creier și minte au avut loc într-un mediu monocromatic, folosind materiale și instrumente alb-negru.

Ce încearcă să demonstreze acest scenariu despre natura conștiinței? Imaginează-ți că Mary, după ani de izolare într-o cameră monocromatică, este în sfârșit eliberată. Așteptată cu nerăbdare de iubitul ei, cu care a schimbat nenumărate scrisori de dragoste, are loc întâlnirea mult visată.

Într-un gest romantic, acesta îi oferă un trandafir roșu. Pentru prima dată, Mary privește un obiect colorat pe care nu l-a mai văzut niciodată, deși cunoaște în detaliu toate mecanismele fizice și neurologice ale percepției culorii, inclusiv cele legate de trandafirii roșii.

Și totuși, știe Mary dinainte la ce să se aștepte? Sau experimentarea directă a culorii roșii îi revelează ceva nou, un aspect al realității care i-a scăpat în ciuda cunoștințelor sale exhaustive despre creier și culori?

Acest experiment de gândire, formulat inițial de filosoful Frank Jackson, ilustrează două idei esențiale.¹⁸ În primul rând, experiențele subiective, numite qualia, sunt reale și nu simple construcții teoretice ale

¹⁸ Ediție originală: Jackson, F. (1982) "Epiphenomenal Qualia", *Philosophical Quarterly*, 32 (127): 127–136.

filosofilor și oamenilor de știință. În al doilea rând, ipoteza conform căreia conștiința poate fi explicată exhaustiv prin proprietăți fizico-chimice este problematică, deoarece exemplul lui Mary sugerează existența unor adevăruri despre culori care par a fi non-fizice.

Deși acest argument are o puternică forță intuitivă, unii gânditori contestă concluziile sale.¹⁹ Un contraargument este că Mary ar putea, de fapt, anticipa experiența culorii roșii chiar înainte de a o percepe direct, ceea ce ar contrazice ideea că dobândește o cunoaștere cu totul nouă în momentul observării trandafirului.

În esență, acceptarea sau respingerea acestor idei depinde de poziția filosofică a cititorului asupra conștiinței. Dacă experimentul lui Mary te convinge, este posibil să consideri că există un aspect fundamental al conștiinței pe care știința actuală nu îl poate explica. Dacă, însă, crezi că Mary știa deja la ce să se aștepte, este probabil să susții că nu există nimic misterios în conștiință care să nu poată fi înțeles în termeni pur fizici.

O altă modalitate de a explora natura conștiinței este prin prisma dezvoltărilor tehnologice și a ipotezei că am

¹⁹ Ediție originală: Churchland, P. M. (1985) "Reduction, Qualia, and the Direct Introspection of Brain States", *The Journal of Philosophy*, 82 (1): 8-28.

putea trăi într-o simulare, un scenariu asemănător celui din filmul *The Matrix*. Filosoful Nick Bostrom a formulat această idee în celebrul său articol „Trăiești într-o simulare computerizată?”,²⁰ argumentând că viitorul umanității se poate încadra într-unul dintre următoarele trei scenarii:

(i) Specia umană este extrem de probabil să dispară înainte de a ajunge la un stadiu „post-uman”, unde ar fi capabilă să creeze simulări complexe.

(ii) Este foarte puțin probabil ca o civilizație post-umană să folosească resursele necesare pentru a rula un număr mare de simulări ale propriei sale istorii evolutive.

(iii) Aproape sigur trăim într-o simulare computerizată.

Bostrom susține că cel puțin una dintre aceste propoziții trebuie să fie adevărată. Dacă excludem prima ipoteză și admitem că o civilizație avansată ar putea ajunge la maturitate tehnologică, sau dacă respingem și a doua ipoteză, presupunând că astfel de civilizații ar avea interesul și capacitatea de a rula un număr masiv de simulări, atunci rezultă că este mult mai probabil ca noi să trăim într-o simulare decât într-o lume reală. Dacă fiecare civilizație reală creează milioane de lumi simulate, numărul ființelor conștiente care trăiesc într-o simulare ar fi incomparabil mai mare

²⁰ Ediție originală: Bostrom, N. (2023) “Are you living in a computer simulation?”, *Philosophical Quarterly*, 53 (211): 243-255.

decât cel al ființelor care trăiesc într-o realitate „de bază”.

Pe baza acestui raționament, probabilitatea ca noi să trăim într-o simulare este semnificativă, iar David Chalmers acceptă această posibilitate, avansând următoarele premise:

(iv) Este mai probabil ca simulările umane conștiente să fie posibile.

(v) Dacă simulările umane conștiente sunt posibile, multe civilizații avansate vor crea astfel de simulări.

(vi) Prin urmare, există o probabilitate ridicată (25% sau mai mult) ca noi să trăim într-o simulare computerizată.²¹

Dar cât de plauzibilă este ipoteza simulării? Dacă ne amintim de filmul *The Matrix*, există multiple scene în care Neo, protagonistul, trăiește o stare de incertitudine totală, neștiind dacă se află în realitate sau într-o lume simulată.

Dacă și noi trăim într-o simulare, descoperirea acestui fapt ar fi extrem de dificilă din punct de vedere subiectiv. Cu toate acestea, unii filosofi și oameni de știință consideră că argumentul probabilistic al simulării

²¹ Ediție originală: Chalmers, D. (2022) *Reality+: Virtual Worlds and the Problems of Philosophy*, New York: W. W. Norton.

eșuează, deoarece unele dintre premisele sale sunt discutabile.

Într-adevăr, este posibil ca într-o zi computerele să dezvolte conștiință, însă este puțin probabil ca această conștiință să fie identică cu cea umană, având în vedere diferențele fundamentale dintre mecanismele fizice ale calculatoarelor și cele neurobiologice ale creierului.

Ipoteza conform căreia o conștiință artificială ar putea funcționa în mod similar celei umane se bazează pe ideea de independență de substrat, adică pe presupunerea că stările mentale pot apărea în orice sistem fizic suficient de complex, indiferent de materialul din care este alcătuit.²²

Totuși, această presupunere poate fi pusă sub semnul întrebării. Este posibil ca fenomenul conștiinței să fie intrinsec dependent de un substrat biologic, dezvoltat de evoluție prin corpul și creierul uman. Dacă acesta este cazul, atunci niciun sistem artificial nu ar putea reproduce cu adevărat conștiința umană în forma sa subiectivă.

Un alt punct vulnerabil al argumentului simulării privește validitatea celei de-a doua premise (ii), care

²² Secțiune bazată pe: Gouveia, S. & Neiva, D. (2017) "The Problem of Consciousness on the Mind Uploading Hypothesis" In *Philosophy of Mind: Contemporary Perspectives*, Cambridge: Cambridge Scholars Publishing.

afirmă că o civilizație avansată ar avea tendința să ruleze un număr masiv de simulări. De ce ar investi o civilizație postumană resurse uriașe pentru a crea mii sau milioane de universuri simulate, în loc să le folosească pentru alte priorități existențiale? Nu este deloc evident că un astfel de proiect ar fi justificat din punct de vedere practic.

Mai mult decât atât, o astfel de civilizație ar putea considera simularea ființelor conștiente ca fiind profund imorală. Dacă simularea unui univers include și simularea ființelor umane și, dacă ipoteza independenței de substrat este corectă, ceea ce rămâne o întrebare deschisă, atunci aceste ființe ar avea capacitatea de a simți durerea și suferința.

În acest caz, rularea unui număr infinit de astfel de simulări ar putea genera un volum incontrollabil de suferință, ceea ce ar constitui o greșeală etică de proporții colosale. De fapt, am putea vorbi despre una dintre cele mai mari erori morale din întreaga istorie a umanității.

În cele din urmă, acceptarea sau respingerea ipotezei simulării rămâne o alegere individuală. Dar, oricare ar fi poziția cititorului, merită să reflectăm la o întrebare simbolică: dacă ai fi pus în fața deciziei, ai alege pastila roșie, care presupune acceptarea realității așa cum este ea, sau pastila albastră, pentru a rămâne în continuare într-o lume simulată?

Apropo de alterarea stărilor de conștiință, în continuare vom explora, alături de psihologul și profesorul Susan Blackmore, diverse stări modificate ale conștiinței și lecțiile pe care acestea ni le pot oferi despre natura minții conștiente.

III. CONȘTIINȚA ALTERATĂ

Al doilea dialog o include pe Susan Blackmore, scriitoare, lector și profesor invitat la Universitatea din Plymouth, Regatul Unit. Cu un doctorat în psihologie, Blackmore explorează natura conștiinței prin metodologii neconvenționale.

De obicei, când discutăm despre „conștiință”, avem tendința de a o aborda dintr-o perspectivă normalizată, considerând-o în starea sa obișnuită, atunci când suntem treji și experimentăm realitatea de zi cu zi mai degrabă decât în timpul visării sau al altor stări modificate.

Totuși, de-a lungul istoriei, s-a recunoscut existența unor stări de conștiință distincte de cele normale. Acestea pot fi induse prin consumul de substanțe psihoactive, meditație sau, așa cum vom explora mai târziu, prin vise lucide. Dar de ce ar merita să ne concentrăm asupra acestor stări neobișnuite? Pentru că ele ne oferă o perspectivă unică asupra mecanismelor conștiinței, permițându-ne să comparăm modul în care experiența subiectivă se transformă în condiții alterate, ceea ce numim „Stare Alterată de Conștiință” (SAC).

Cum putem defini aceste stări? O metodă obiectivă este clasificarea lor în funcție de modul de inducere. De exemplu, o stare modificată prin consumul de

Ayahuasca diferă fundamental de una indusă prin hipnoză sau meditație.



Imaginea 5: "Essere" © Gaia Amoroso

Deși această abordare este utilă, nu este suficientă pentru a înțelege pe deplin conștiința. Motivul? Aceeași substanță poate produce efecte diferite de la o persoană la alta. Dacă eu, autorul acestei cărți, și cititorul am consuma aceeași substanță psihoactivă, experiențele noastre ar fi foarte diferite. Nu există o relație cauzală strictă între o substanță și o anumită stare alterată de conștiință.

O altă metodă de definire a SAC se bazează pe măsurători fiziologice – temperatura corpului, ritmul cardiac, consumul de oxigen cortical etc. Totuși, și această abordare are limitări fiindcă acești parametri variază semnificativ între indivizi, atât în ceea ce privește reacțiile fiziologice, cât și percepția subiectivă a stării respective.

Știm că modificări minore ale fiziologiei pot corespunde unor schimbări majore ale stării subiectiv, dar și invers. Din acest motiv, nu există o corelare directă între indicatorii fiziologici și experiențele conștiinței. Astfel, încercarea de a defini aceste stări extrem de volatile pe baza unor parametri obiectivi necesită o abordare atentă.

În literatura de specialitate, psihologia, neuroștiințele și filosofia propun mai multe definiții pentru SAC. Iată câteva dintre ele:

- „(...) alterare calitativă a tiparelor generale de funcționare mentală, astfel încât individul percepe că procesele conștiinței sale sunt radical diferite de cele obișnuite.”;²³
- „(...) schimbare temporară a modelului general al experienței subiective, astfel încât individul are senzația că funcționarea sa mentală este

²³ Ediție originală: Tart, C. (1972) “States of consciousness and state-specific sciences”, *Science*, 176: 1203-1210.

semnificativ diferită de starea normală de veghe”;²⁴

- “(...) stare care apare ori de câte ori conștiința se îndepărtează de un model obișnuit de funcționare, trecând într-o formă percepută ca fiind diferită de către persoana care o trăiește”.²⁵

Deși aceste definiții sunt interesante, ele sunt limitative într-un anumit sens, deoarece definesc stările alterate de conștiință doar prin opoziție cu cele „normale”. Problema fundamentală este că nu avem o definiție clară și universal acceptată a conștiinței „normale”, ceea ce face ca aceste abordări comparative să fie insuficiente pentru a identifica trăsăturile comune ale stărilor modificate.

O alternativă ar fi explorarea unor exemple concrete care pot ajuta la clarificarea acestor dificultăți conceptuale. Să începem, așadar, cu o categorie de stări alterate pe care cititorul le-ar putea cunoaște deja: cele induse de droguri.

Drogurile psihoactive sunt substanțe care afectează funcționarea psihică și conștiința, acționând asupra neurotransmițătorilor sau neuromodulatorilor din creier. Aceste substanțe pot amplifica efectele unui

²⁴ Ediție originală: Farthing, G. (1992), *The Psychology of Consciousness*, Englewood Cliffs, NJ: Prentice Hall.

²⁵ Ediție originală: Nolen-Hoeksema, et al. (2014) *Atkinson & Hilgard's Introduction to Psychology*, Andover: Cengage Learning.

neurotransmițător fie stimulând eliberarea acestuia, fie împiedicând recaptarea sa, prelungindu-i astfel acțiunea. Pe de altă parte, ele pot reduce activitatea neurotransmițătorilor prin inhibarea sau blocarea receptorilor postsinaptici. Motivul pentru care aceste substanțe au un impact profund asupra conștiinței este că un singur neurotransmițător poate fi activ în multiple regiuni ale creierului și chiar în întregul organism.

Un exemplu bine cunoscut este amfetamina, unul dintre cei mai studiați stimulanți, care are trei efecte majore asupra creierului:

- stimulează eliberarea de serotonină;
- stimulează eliberarea de dopamină;
- inhibă recaptarea serotoninei.

Serotonina joacă un rol esențial în reglarea stării de spirit și a somnului, în timp ce dopamina mediază comportamentul motivat de recompensă și influențează percepția sinelui, a celorlalți și a mediului înconjurător.

Efectele consumului de amfetamine includ creșterea energiei, intensificarea senzațiilor tactile și amplificarea sentimentelor de iubire și empatie.²⁶ Aceste experiențe sunt puternic influențate de mediul în care este consumată substanța. Utilizarea frecventă poate duce

²⁶ Ediție originală: Holland, J. (ed.) (2001) *Ecstasy: The complete guide: A comprehensive look at the risks and benefits of MDMA*, Rochester, VT: Park Street Press.

la dezvoltarea toleranței și, implicit, la dependență pe termen lung.

Totuși, utilizată în contexte terapeutice și în doze controlate, amfetamina s-a dovedit eficientă în tratarea unor afecțiuni precum tulburarea de stres post-traumatic și anxietatea socială, având uneori rezultate mai bune decât tratamentele convenționale.²⁷

Cel mai comun psihedelic este canabisul, care conține aproximativ 85 de compuși canabinoizi. Descrierea efectelor sale subiective este dificilă, deoarece variază considerabil de la o persoană la alta. Cu toate acestea, cercetările științifice au identificat câteva efecte comune. La doze mici, canabisul poate induce euforie și relaxare, în timp ce la doze mari poate provoca anxietate și paranoia.

Efectele senzoriale includ o percepție accentuată a profunzimii, intensificarea simțurilor, creșterea plăcerii sexuale, o senzație de încetinire a timpului și de expansiune a spațiului, precum și o concentrare sporită asupra momentului prezent. Unii consumatori raportează trăiri mistice sau sacre, iar alții afirmă că le stimulează creativitatea. Totodată, poate afecta memoria pe termen scurt și încetini gândirea. La doze mari, pot apărea fenomene de sinestezie, în care

²⁷ Ediție originală: Danforth, A., et al. (2016) *MDMA-assisted Therapy: a new treatment model for social anxiety in autistic adults*, *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, 64: 237–249.

senzații diferite se combină, de exemplu, perceperea culorilor ca având sunete.

Un alt psihedelic din ce în ce mai popular este ayahuasca, considerată o substanță cu efecte vindecătoare. Consumul său provoacă inițial episoade de greață și vărsături, urmate de o serie intensă de senzații, viziuni și transformări perceptuale.²⁸ Experiențele includ frecvent contemplarea morții și trăiri mistice legate de probleme personale sau întrebări existențiale profunde.

Din punct de vedere neurologic, consumul de ayahuasca reduce conectivitatea dintre parahipocamp și cortexul retrosplenal, un fenomen corelat cu senzația de dizolvare a sinelui.²⁹

Acesta poate explica de ce utilizatorii raportează adesea sentimentul că eul lor se contopește cu mediul înconjurător.

Una dintre cele mai puternice substanțe halucinogene este LSD-ul (dietilamida acidului lisergic), ale cărui efecte variază de la stări de euforie și extaz la sentimente de teroare și dezintegrare a sinelui. Un

²⁸ Ediție originală: Luna, L. e White, S. (2016) *Ayahuasca Reader: Encounters with the Amazon's Sacred Vine*, Santa Fe: Synergetic Press.

²⁹ Ediție originală: Uthaug, M. et al. (2018) "Sub-acute and long-term effects of ayahuasca on affect and cognitive thinking style and their association with ego dissolution", *Psychopharmacology*, 235: 2979-2989.

studiu controlat cu placebo, în care participanții au primit 75 de micrograme de LSD intravenos, a arătat o creștere a conectivității funcționale în întregul creier, cu efecte locale care corespundeau schimbărilor subiective ale experienței.

De exemplu, halucinațiile vizuale au fost corelate pozitiv cu un flux sanguin crescut și o conectivitate funcțională sporită în cortexul vizual primar (V1).

Toate aceste substanțe oferă consumatorilor experiențe conștiente diferite de cele „normale”, ceea ce le face valoroase pentru înțelegerea conștiinței. Dar de ce?

Pentru că ne permit să comparăm activitatea neuronală a unei persoane aflate sub influența drogurilor cu cea a unei persoane într-o stare obișnuită, oferindu-ne informații despre dinamica rețelelor neuronale și relațiile dintre diferite regiuni ale creierului.



Imaginea 7: "Siamo uguali" © Gaia Amoroso

Acum, să abordăm meditația, ca o altă perspectivă asupra Stărilor alterate de conștiință. Poate fi considerată un ASC? Unele definiții sugerează că:

„(...) meditația poate fi privită ca un proces gradual, cumulativ și de lungă durată, menit să inducă o stare modificată de conștiință.”³⁰

Practicanții budismului raportează adesea experiențe de kenshō (trezire), un concept care desemnează o stare de „iluminare” în care se presupune că persoana accesează natura profundă a minții. De asemenea, majoritatea celor care meditează afirmă că mintea lor suferă o transformare profundă, ceea ce justifică încadrarea meditației ca ASC.

Cu toate acestea, unii cercetători susțin că meditația nu este altceva decât o formă particulară de somn. Neuropsihiatrul Peter Fenwick³¹ a demonstrat că profilurile EEG ale meditației nu sunt identice cu cele ale somnului, dar că mulți practicanți intră în microsomn în timpul meditației. Un alt studiu a constatat că subiecții

³⁰ Ediție originală: Wallace, B. e Fisher, L. (1991) *Consciousness and Behavior*, Boston, MA: Allyn and Bacon.

³¹ Ediție originală: Fenwick, P. (1987) "Meditation and the EEG" In M. West (ed.) *The Psychology of Meditation*, Oxford: Clarendon Press.

au dormit aproximativ o treime din timpul petrecut în meditație.³²

Beneficiile meditației, cum ar fi reducerea anxietății și depresiei sau îmbunătățirea funcțiilor cognitive, ar putea fi explicate parțial prin efectele microsomnului. Cu toate acestea, mulți practicanți susțin că pot distinge clar între meditația profundă și somn, deși această diferență rămâne dificil de măsurat obiectiv.

O altă modalitate de a explora stările alterate de conștiință este prin intermediul viselor lucide. Un vis lucid se caracterizează prin conștientizarea, în timpul somnului, că visezi.

În fiecare noapte, somnul uman urmează un ciclu format din trei stări principale: veghe, somn REM (mișcare rapidă a ochilor) și somn non-REM. Un ciclu complet de somn include patru sau cinci alternanțe între fazele REM și non-REM, adesea însoțite de micro-treziri inconștiente. Stările de veghe și somn pot fi caracterizate prin mai mulți indicatori comportamentali, precum viteza de trezire, mișcările oculare, tonusul muscular și activitatea cerebrală, măsurabile prin tehnici precum electroencefalograma (EEG).

³² Ediție originală: Austin, J. (1998) *Zen and the brain: Toward an Understanding of Meditation and Consciousness*, Cambridge, MA: MIT Press.

Se știe că, în timpul somnului REM, creierul este extrem de activ, iar EEG-ul său seamănă cu cel din starea de veghe. În schimb, în somnul non-REM, rata generală de activitate neuronală rămâne ridicată, dar urmează un model diferit, dominat de unde lente și ample, spre deosebire de modelele complexe și rapide din starea de veghe.

În timpul somnului, creierul se izolează de mediul extern prin mai multe mecanisme. În faza non-REM, există o inhibare a fluxului senzorial la nivel talamocortical, în timp ce în REM această inhibare este mai periferică. Studii EEG și fMRI au arătat că, în timpul stimulilor auditivi din faza REM, cortexul auditiv rămâne activ. De asemenea, în fazele intermitente ale REM, care includ mișcări oculare și spasme musculare, creierul operează într-un circuit închis, fiind funcțional izolat de lumea exterioară.³³

Un alt mecanism specific somnului REM este blocarea transmiterii comenzilor motorii de către trunchiul cerebral spre măduva spinării, prevenind astfel ca activitățile mentale din vis să fie transpuse în mișcări reale. Totuși, eșecuri ale acestui mecanism pot explica fenomene precum somnambulismul, care poate duce,

³³ Publicație originală: Wehrle, R. et al. (2007) "Functional microstates within human REM sleep: First evidence from fMRI of a thalamocortical network specific for phasic REM periods", *European Journal of Neuroscience*, 25: 863-871.

de exemplu, la treziri nocturne și consum alimentar inconștient.

În timpul somnului REM, activitatea cerebrală crește semnificativ în amigdala, hipocamp și cortexul cingulat anterior, precum și în anumite regiuni ale sistemului vizual și ale ariilor de procesare vizuală. În schimb, cortexul prefrontal dorsolateral, asociat cu funcții executive precum memoria de lucru, rezolvarea problemelor și planificarea acțiunilor, prezintă o activitate redusă comparativ cu starea de veghe.

Cercetarea viselor lucide oferă un cadru interesant pentru explorarea corelațiilor neuronale ale conștiinței. Posibilitatea de a lega datele fiziologice, neurochimice și comportamentale de relatările subiective ale viselor permite o mai bună înțelegere a dimensiunilor experienței conștiente.

Un alt mod de a aborda conștiința este prin experiențele în afara corpului (Out of Body Experiences – OBE), cum a fost cazul profesoarei Susan Blackmore, care, la vârsta de 19 ani, a trăit un episod intens de acest tip (vezi dialogul cu Susan Blackmore).



Imaginea 8: "Esperienza" © Gaia Amoroso

Un OBE este o experiență în care individul are senzația că percepe realitatea dintr-o poziție externă corpului său fizic. Cei care au astfel de episoade raportează frecvent trăiri paranormale și o credință mai puternică în fenomene supranaturale, comparativ cu persoanele care nu au trecut prin astfel de experiențe. De asemenea, OBE-urile sunt corelate cu o memorie mai bună a viselor și o frecvență mai mare a viselor lucide.³⁴

Unii interpretează aceste experiențe ca dovezi ale existenței unui suflet imaterial, separat de corp și creier. Deși această perspectivă este firească pentru cineva

³⁴ Ediție originală: Blackmore, S. (2017) *Seeing Myself: The New Science of Out-of-body Experiences*, London: Robinson.

fără cunoștințe științifice aprofundate, există explicații alternative care merită luate în considerare.

Dacă „sufletul” ar putea percepe lumea fizică în timpul unui OBE, ar trebui să existe o interacțiune între acesta și mediul fizic, ceea ce ar contrazice natura sa non-fizică și ar implica detectabilitatea sa prin metode științifice. Pe de altă parte, dacă sufletul este cu adevărat non-fizic, atunci nu ar putea interacționa cu realitatea materială pentru a percepe corpul din exterior. Acest paradox ridică întrebări asupra validității interpretărilor dualiste.

Teoriile psihanalitice au propus că OBE-urile ar putea reflecta frica de moarte, regresia ego-ului sau reexperimentarea traumei nașterii. Cu toate acestea, aceste teorii sunt dificil de testat științific și au avut un impact limitat asupra dezvoltării înțelegerii fenomenului OBE.

Totuși, dacă ne uităm la cele mai actuale studii din neuroștiințe, aflăm că lobul temporal joacă un rol semnificativ asupra OBE deoarece, la pacienții epileptici, stimularea lobului temporal provoacă episoade de acest fel, precum și experiențe mistice.

Omul de știință Michael Persinger a propus că aceste credințe și experiențe mistice sunt „creații” ale funcției lobului temporal al creierului, reușind să inducă mai multe OBE-uri și distorsiuni corporale atunci când a folosit stimularea magnetică transcraniană.

Zona specifică a creierului implicată în stimulare este joncțiunea temporoparietală dreaptă, având în vedere că această zonă este responsabilă de procesarea informațiilor vizuale, tactile, proprioceptive și vestibulare. Aceste semnale, atunci când sunt combinate, formează schema corporală pentru fiecare dintre noi. În acest fel, putem oferi o explicație științifică pentru un fenomen care poate părea mistic, dar este, în esență, un fenomen natural.

În continuare, vom introduce și alte idei din domeniul psihologiei, de data aceasta, dintr-o perspectivă influențată de evoluție. Scopul este de a explica de ce noi, ca ființe umane și unele animale, avem experiențe subiective conștiente.

IV. CONȘTIINȚA EVOLUATĂ

Al treilea dialog îl aduce în prim-plan pe renumitul psiholog și teoretician Nicholas Humphrey, profesor emerit de psihologie la London School of Economics.

Călătoria lui Humphrey în studiul conștiinței a început la vârsta de 23 de ani, pe când era doctorand într-un laborator de psihologie la Universitatea din Cambridge. În cadrul cercetărilor sale, a studiat inițial o maimuță asupra căreia s-a intervenit neurochirurgical: în timp ce aceasta era anesteziată, i-a fost introdus un electrod în coliculul superior, o regiune primitivă a creierului responsabilă de procesarea vizuală. Acest centru vizual subcortical precede, din punct de vedere evolutiv, cortexul vizual, care permite percepția conștientă la oameni și mamifere. Aspectul intrigant al acestui experiment a fost că, deși maimuța era inconștientă, activitatea neuronală din coliculul superior era intensă, sugerând că procesarea vizuală avea loc, dar fără o experiență conștientă asociată.

Câteva luni mai târziu, Humphrey a studiat o altă maimuță, numită Helen, al cărei cortex vizual fusese complet îndepărtat, lăsând intact doar coliculul superior. Conform așteptărilor, Helen ar fi trebuit să fie complet oarbă. Cu toate acestea, după multiple

interacțiuni, a început să manifeste un comportament care indica o anumită formă de percepție vizuală: reușea să culeagă fructele preferate, să-l recunoască pe cercetător și să interacționeze cu mediul său. Aceste observații l-au condus pe Humphrey la o ipoteză surprinzătoare: Helen putea procesa informații vizuale fără a avea experiențe conștiente ale acestor informații.

Această descoperire l-a determinat pe coordonatorul său doctoral, profesorul Larry Weiskrantz, să analizeze un caz uman similar. Un pacient al cărui cortex vizual fusese grav afectat în urma unui accident a fost supus unor teste. Deși susținea că nu vede nimic în partea afectată a câmpului său vizual, el putea identifica cu precizie obiecte aflate acolo.



Imaginea 9: "Everything is worth it" © Gaia Amoroso

Această abilitate paradoxală a fost denumită vedere oarbă (blindsight) și a deschis noi perspective asupra modului în care procesarea vizuală se poate desfășura și în absența conștiinței fenomenale. Deși acest fenomen a ajutat la înțelegerea mecanismelor vizuale, a ridicat și întrebări fundamentale despre natura conștiinței.

Aceste observații l-au inspirat pe Humphrey să exploreze un aspect cheie: de ce și cum a evoluat conștiința? Principala contribuție a lui Humphrey este o teorie care explică apariția conștiinței prin mecanisme darwiniene. În cuvintele sale:

“(...) Voi susține că adevărul despre conștiință, dacă o vedem din perspectiva corectă, este că este produsul unui fragment extrem de improbabil de inginerie biologică: o operă de artă a naturii, care creează în mintea noastră impresii misterioase, dar care are o explicație fizică relativ simplă.”³⁵

Primul pas este să argumentăm că fenomenul conștiinței are un impact asupra comportamentului nostru. Și de ce comportamentul? Cunosând despre fenomenul de conștiință, așa cum o înțelegem, că este o trăsătură inerentă vieții de pe Pământ, putem presupune că, precum orice altă trăsătură specializată

³⁵ Ediție originală: Humphrey, N. (2011) *Soul Dust: the Magic of Consciousness*, Princeton: Princeton University Press.

a organismelor vii, ea a evoluat deoarece conferă un anumit avantaj selectiv.

Într-un fel sau altul, conștiința trebuie să ajute organismul să supraviețuiască și să se reproducă. Acest lucru se poate întâmpla numai dacă, într-un fel, modifică modul în care organismul interacționează cu lumea exterioară, un proces care implică de obicei un comportament.

Dar cum se manifestă conștiința în comportamentul unui organism? Humphrey observă că ființele conștiente nu prezintă trăsături fizice distinctive comparativ cu cele non-conștiente. Conștiința nu aduce mai multă forță, sănătate sau frumusețe; în schimb, ea influențează psihologia organismului, modelându-i gândurile, dorințele și convingerile. Această modificare psihologică determină un comportament adaptativ, care oferă un avantaj evolutiv.

Totuși, dacă conștiința oferă un astfel de avantaj, de ce a fost ea selectată prin evoluție? Aici intervine punctul central al teoriei lui Humphrey: dacă selecția naturală poate „percepe” impactul psihologic al conștiinței asupra comportamentului, atunci și alți observatori externi pot recunoaște aceste avantaje.

Astfel, suntem pe cale să construim o explicație plauzibilă pentru evoluția conștiinței. Pe lângă faptul că aduce un beneficiu fiecărui individ, acest avantaj se reflectă în comportamentul său, făcându-l observabil și

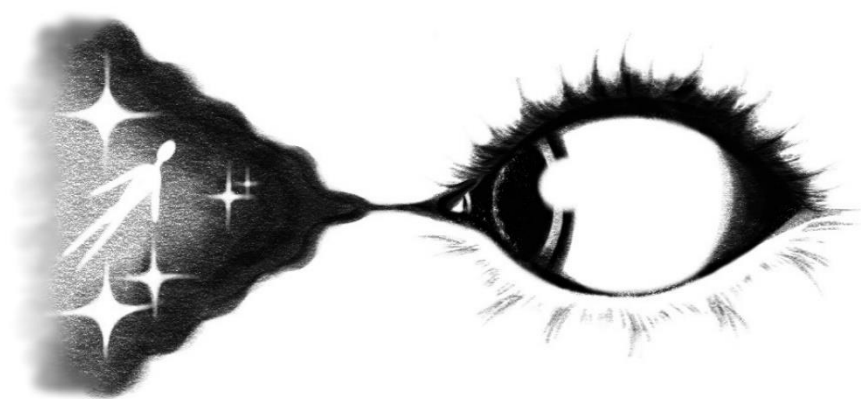
selectabil evolutiv. Prin urmare, conștiința nu este doar un produs pasiv al activității cerebrale, ci un element esențial al strategiei adaptative a organismelor vii.

Cum a apărut conștiința din punct de vedere evolutiv și cum a devenit o trăsătură fundamentală a ființelor umane de astăzi? Nicholas Humphrey consideră că apariția conștiinței a fost un fenomen brusc, de tip „totul sau nimic”, survenit relativ târziu în istoria evoluției, când strămoșii noștri hominizi au început să dezvolte abilități sociale complexe, precum imitația, înșelăciunea și limbajul. Astfel, conștiința ar putea fi o proprietate emergentă, evoluată în principal pentru funcțiile sale sociale – capacitatea de a înțelege, anticipa și influența comportamentul altor indivizi. Selecția naturală ar fi favorizat indivizii capabili să interpreteze și să prezică acțiunile celorlalți, conducând astfel la dezvoltarea și persistența conștiinței.

Așa cum se întâmplă și în cazul primatelor mari din prezent, oamenii au trăit dintotdeauna în grupuri sociale complexe. În acest context, capacitatea de a înțelege intențiile altora este esențială pentru supraviețuire – de la stabilirea ierarhiilor sociale și identificarea persoanelor de încredere, până la formarea alianțelor strategice.

Potrivit lui Humphrey, strămoșii care puteau interpreta, prezice și manipula comportamentul altora aveau un avantaj clar. S-ar putea susține că aceste abilități s-ar fi

putut dezvolta doar prin observarea comportamentului celorlalți și a consecințelor acestuia. Totuși, el propune o ipoteză alternativă: indivizii ar fi dobândit capacitatea de introspecție, de a se pune în locul altora și de a-și „observa” propria minte.



Imaginea 10: "Percezione" © Gaia Amoroso

Humphrey compară această capacitate cu un „ochi interior”, îndreptat nu spre lumea exterioară, precum simțurile noastre obișnuite, ci spre lumea interioară a individului. Acest „ochi” nu percepe direct activitatea neuronală a creierului, ci o interpretare psihologică accesibilă a acesteia – ceea ce experimentăm ca stări conștiente subiective.

Dar acest lucru ridică o întrebare fundamentală: ar putea conștiința să fie o „invenție” evolutivă, în sensul

că ea nu a existat într-o anumită etapă a dezvoltării umane și a apărut mai târziu? Humphrey oferă un răspuns dual:

- conștiința este o facultate cognitivă selectată evolutiv pentru a ne ajuta să ne înțelegem pe noi înșine și mediul nostru;
- în același timp, conștiința este și o fantezie creată de creier, care modifică valoarea pe care o atribuim propriei existențe.

Această abordare ridică însă riscul unei erori conceptuale cunoscute sub numele de „teatrul cartezian”, un concept introdus de Daniel Dennett. Această idee presupune că în creier există un loc specific unde are loc „proiecția” lumii exterioare, similar unui ecran pe care conștiința percepe realitatea.

Humphrey nu susține o poziție iluzionistă radicală asupra conștiinței – ideea că experiența conștientă ar fi o simplă iluzie generată de creier. În schimb, el sugerează că ceea ce se întâmplă în „teatrul conștiinței” nu este o replică exactă a lumii externe, ci o reconstrucție a lumii interioare, domeniul subiectivității și al qualiei.

Totuși, el respinge realismul qualiei, așa cum este susținut de filosofi precum David Chalmers sau Thomas Nagel, care consideră qualia drept entități fundamentale și independente de procesele fizice. Din

această perspectivă, experiențele subiective nu sunt „lucruri” distincte în sine, ci construcții ale creierului.

Această abordare sugerează că, deși conștiința pare să posedă proprietăți misterioase și non-fizice, acest efect ar putea fi rezultatul unui truc al creierului. Deoarece creierul este, în esență, un sistem de procesare a informației bazat pe simboluri, el poate construi reprezentări mentale care reflectă realitatea fizică, dar care pot, de asemenea, să creeze impresii care nu corespund direct lumii obiective.

După cum subliniază Daniel Dennett, considerat unul dintre cei mai importanți susținători ai teoriei iluzioniste: „Conștiința este o iluzie a creierului, pentru creier, de către creier”.³⁶

Atunci când avem o experiență conștientă subiectivă, creierul nostru joacă un veritabil spectacol de magie, făcându-ne să credem că trăim ceva dincolo de propriul nostru sistem biologic. Într-un fel, această perspectivă contrazice ideea că experiența umană este unică, misterioasă și excepțională în univers.³⁷

Pentru Nicholas Humphrey, însă, aceasta nu este neapărat o veste rea. Dimpotrivă, înțelegerea conștiinței ca un produs al proceselor naturale poate fi

³⁶ Ediție originală: Dennett, D. (1991) *Consciousness Explained*, New York: Little Brown.

³⁷ Ediție originală: Humphrey, N. (2020) “The Invention of Consciousness”, *Topoi*, 39: 13-21.

un atu, fiind o oportunitate de a ne îmbogăți ca ființe umane. Această îmbogățire poate avea loc prin artă, știință și prin transcenderea unei existențe pur pământești.

Pentru a încheia această prezentare a teoriei sale, să aruncăm o privire asupra ultimelor cuvinte din capitolul 23 al cărții sale recente, *Sentience*, unde reflectează asupra rarității conștiinței în univers:

„Chiar dacă ideea că o trăsătură evoluată natural este „intenționată” este greșită, îmi imaginez că Darwin însuși ar fi putut vedea conștiința fenomenală ca pe o realizare „ultimă” – gloria încununată a procesului evolutiv care a început odată cu Big Bang. Este o invenție atât de sublimă încât, dacă ar înceta să mai existe, ar diminua într-adevăr întreaga creație.”

Această viziune evoluționistă a conștiinței oferă o abordare interesantă care poate fi studiată empiric. Cu toate acestea, nu clarifică modul în care creierul nostru produce proprietatea emergentă a subiectivității conștiente.³⁸

Pentru a încheia această primă parte a cărții dedicată conștiinței, voi prezenta ideile ultimului invitat cu care

³⁸ Secțiunea bazată pe: Humphrey, N. (2023) *Sentience: the Invention of Consciousness*, Cambridge, MA: MIT Press.

am dezbătut acest subiect, profesorul Sir Roger Penrose.

V. CONȘTIINȚA CUANTICĂ

Pentru a pregăti terenul pentru una dintre cele mai stimulante conversații din această carte, voi introduce câteva dintre conceptele dezvoltate de laureatul Nobel pentru fizică, Sir Roger Penrose, profesor emerit la Universitatea din Oxford.

Cum ar trebui să ne gândim la conștiință? Mulți oameni de știință și filosofi o consideră o proprietate emergentă a activității neuronale, rezultată din calculele efectuate de rețelele de neuroni interconectați prin sinapse. În această viziune, creierul este văzut ca un computer digital biologic care procesează informații din mediu, iar conștiința joacă un rol important în acest proces.

Problema centrală, în acest context, este să înțelegem tipul de calcul realizat de creier și dacă acesta poate fi replicat artificial. Totuși, Penrose contestă această analogie computațională, susținând că activitatea creierului nu poate fi redusă la simple operațiuni algoritmice. Spre deosebire de computerele digitale, creierul operează la o scară de procesare complet diferită, iar explicația conștiinței trebuie căutată dincolo de paradigma clasică a neuroștiinței computaționale.

Abordarea lui Penrose presupune extinderea cadrului explicativ dincolo de fizica clasică și modelele tradiționale ale neuroștiinței, explorând un nivel mai

profund de realitate – cel cuantic. Pentru aceasta, el contestă ideea că proprietățile cuantice încetează să fie relevante pe măsură ce sistemele biologice devin mai complexe.

Criticii susțin că, la nivelul neuronilor și interacțiunilor cerebrale, fizica cuantică nu joacă un rol semnificativ și că procesele cognitive ar trebui interpretate exclusiv prin legile fizicii clasice.

Opunându-se acestei teze, un grup de oameni de știință³⁹ a arătat că mecanica cuantică are un rol fundamental în fotosinteză, descoperind că în celulele clorofilei, transferul luminii are aproape 100% eficiență, ceea ce depășește cu mult orice tip de tehnologie umană modernă.

Acest lucru indică faptul că trebuie să existe un mecanism care protejează coerența cuantică și permite acest nivel excepțional de eficiență.

Dacă fenomenele cuantice pot avea un impact atât de profund în biologie, este plauzibil să ne întrebăm dacă acestea ar putea fi implicate și în funcționarea creierului și în emergența conștiinței.

³⁹ Ediție originală: Engel, G. et al. (2007) "Evidence for wavelike energy transfer through quantum coherence in photosynthetic systems", *Nature*, 446 (2007): 782-786.

De ce trebuie să trecem dincolo de computerele digitale pentru a ajunge la o teorie viabilă a conștiinței?

Penrose susține că fizica clasică are trei presupuneri care exclud o explicație a conștiinței umane:

- determinism causal: presupune că putem cunoaște toate condițiile inițiale ale unui sistem fizic și, prin urmare, putem prezice comportamentul său viitor pe baza unui lanț causal;
- localitate/independență: două sisteme separate în spațiu nu pot interacționa instantaneu;
- realism obiectivist: obiectele din realitatea externă există cu proprietăți bine definite care sunt independente de orice observator.

Penrose susține că conștiința este non-computațională, adică nu poate fi explicată printr-un set de operații algoritmice. Această idee derivă direct din teorema incompletitudinii lui Gödel, conform căreia există adevăruri matematice care nu pot fi demonstrate de un sistem formal de calcul.⁴⁰

Mai exact, el argumentează că ceea ce numim „calitatea mentală a înțelegerii” nu poate fi redusă la o serie de operații algoritmice, deci trebuie să existe un proces non-calculabil, care depășește limitele modelului computațional standard. Aceasta înseamnă că mintea

⁴⁰ Ediție originală: Gödel, K. (1931) “Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I”, *Monatshefte für Mathematik und Physik*, 37: 349–360.

umană are acces la un tip de procesare care nu poate fi realizat de un computer digital.

Penrose sugerează că ingredientul non-calculabil pentru conștiință trebuie să se regăsească într-un domeniu unde teoriile noastre fizice actuale sunt incomplete, dar care este relevant pentru scalele de funcționare ale creierului.

Singura posibilitate „serioasă” de a găsi acest mecanism este, în opinia sa, incompletitudinea teoriei cuantice – o problemă pe care mari fizicieni precum Albert Einstein și Erwin Schrödinger o recunoscuseră deja sub numele de Problema Măsurării.

Pentru a aborda această problemă, Penrose propune o extindere a mecanicii cuantice, introducând o teorie denumită “reducere obiectivă” (Objective Reduction – OR).

Această teorie sugerează că starea cuantică nu se prăbușește pur și simplu în momentul observării (așa cum afirmă interpretarea clasică de la Copenhaga), ci că există un proces fizic fundamental care determină această tranziție într-un mod obiectiv și independent de observator.

Se reduce conștiința la un proces de colaps al stărilor cuantice? În interpretarea clasică a mecanicii cuantice, actul de observare determină colapsul funcției de undă, încălcând principiul superpoziției. În celebrul paradox al

Pisicii lui Schrödinger, pisica este atât vie, cât și moartă până când un observator intervine, forțând sistemul să adopte o stare definită.⁴¹

Pentru a înțelege această idee într-un mod intuitiv, imaginează-ți că arunci o monedă pentru a decide ce vei mânca la cină: dacă pică pajură, vei alege seitan pane; dacă pică cap, raclette elvețiană. În momentul în care moneda este în aer, aceasta se află într-o suprapunere de stări, fiind simultan cap și pajură. Doar atunci când cade și o privești, starea devine definită.



Imaginea 11: "Riflessione" © Gaia Amoroso

⁴¹ Ediție originală: Schrödinger, E. (1935) "Die gegenwärtigen situation in der quantenmechanik", *Naturwissenschaften*, 23: 807-812, 823-8, 844-9.

În mod similar, particulele subatomice există în stări suprapuse până când o observație externă determină colapsul într-o stare concretă. Însă Penrose respinge această abordare, evidențiind un conflict între mecanica cuantică și teoria relativității.⁴²

Spre deosebire de interpretarea standard, în care observația determină reducerea stării cuantice, el susține că acest colaps are loc spontan și că tocmai acest proces dă naștere conștiinței.

O critică majoră la adresa teoriei lui Penrose este că mecanica cuantică funcționează doar la nivel subatomic, în timp ce procesele neuronale au loc la o scară mult mai mare. Se argumentează că, odată ce fenomenele cuantice se extind la nivel macroscopic, ele sunt afectate de mediul înconjurător – un mediu „fierbinte”, „umed” și „zgomotos” din punct de vedere termodinamic, ceea ce ar duce inevitabil la prăbușirea stărilor cuantice și la pierderea coerenței cuantice.

Această lipsă de coerență poate fi explicată printr-o analogie: imaginează-ți că ai o pisică cuantică, care, conform principiului superpoziției, este simultan vie și moartă. Dacă pisica interacționează cu particulele din mediu (de exemplu, cu fotonii), se produce incoerența cuantică. Întrucât mediul „observă” pisica indirect,

⁴² Ediție originală: Einstein, A. (1916) “Die Grundlage der allgemeinen Relativitätstheorie”, *Annalen der Physik*, 49 (7): 769-822.

această interacțiune determină prăbușirea superpoziției, forțând sistemul să adopte o stare definită (sperăm, desigur, că este cea în care pisica rămâne vie).

Aplicând acest raționament la creier, pare implauzibil ca principiile mecanicii cuantice să fie implicate în schimbul de informații între neuroni, deoarece coerența cuantică ar trebui să fie instabilă într-un mediu biologic. Dar reprezintă aceasta o problemă insurmontabilă pentru teoria lui Penrose?

Potrivit lui Penrose, această critică nu este validă, deoarece biologia a evoluat pentru a gestiona problema incoerenței cuantice. Acest lucru ar fi fost posibil prin dezvoltarea unei structuri subatomice în neuroni, cunoscută sub numele de microtubuli - structuri proteice cu o formațiune tubulară în interiorul celulelor eucariote (parte a citoscheletului). Ei joacă diverse roluri, cum ar fi determinarea formei celulei, coordonarea mișcării și supravegherea diviziunii celulare.

Acești microtubuli ar putea păstra coerența cuantică, similar procesului extrem de eficient observat în fotosinteză.

Ideea microtubulilor nu îi aparține în mod original lui Penrose, ci provine de la colegul său, anesteziștii american Stuart ff. Acesta i-a propus lui Penrose că în

creier coerența cuantică ar putea avea loc în acești microtubuli.

Hameroff sugerează că microtubulii sunt dispozitivul cuantic pe care Penrose îl căuta în teoria sa, deoarece ajută la reglarea conexiunilor sinaptice, iar forma lor tubulară îi poate proteja de zgomotul înconjurător al neuronului. Pentru a vizualiza rolul microtubulilor, imaginează-ți că neuronii funcționează ca niște fabrici complexe, iar microtubulii sunt rețele de transport care organizează și coordonează circulația materialelor în interiorul acestor fabrici.

Cu toate acestea, este nevoie de mai mult decât un lanț continuu de momente aleatorii de coerență cuantică pentru a avea vreun impact asupra conștiinței. Aceste momente de conștientizare sunt cele care, orchestrate de microtubulii din creierul nostru, au capacitatea de a stoca amintiri și procesa informații.

Pentru Hameroff și Penrose, microtubulii pot păstra în mod adecvat coerența cuantică până când ajung la nivelul neuronal. Pentru conștiință, este necesar ca mulți microtubuli din mai mulți neuroni diferiți să acționeze într-un mod „orchestrat”.

Care este, atunci, diferența dintre stările cuantice normale și stările cuantice care conduc la conștiință? Pentru acești oameni de știință, cheia constă în coerența globală, ei susținând astfel reducerea

obiectivă (OR), în care stările cuantice se prăbușesc într-o opțiune sau alta.

Partea „orchestrată” este propusă pentru a se asigura că eforturile de colaborare între microtubuli sunt necesare pentru a influența nivelul neuronal.

Pentru a rezuma ideile de până acum, avem următoarea configurație teoretico-conceptuală:

- este propus un model în care conștiința apare prin efecte cuantice care apar în structurile subcelulare interne neuronilor cunoscute sub numele de microtubuli;
- acest model postulează așa-numitele „colapsuri obiective” care implică tranziția sistemului cuantic de la o superpoziție a mai multor stări posibile la o singură stare definită, dar fără intervenția unui observator sau a unei măsurători, ca în majoritatea modelelor de mecanică cuantică;
- conform lui Penrose și Hameroff, mediul intern al microtubulilor este deosebit de potrivit pentru astfel de colapsuri obiective. Auto-colapsurile rezultate produc un flux coerent care reglează activitatea neuronală și permite procese mentale non-algoritmice.

Pentru a înțelege pe deplin teoria conștiinței propusă de Roger Penrose, este important să analizăm Interpretarea Penrose a mecanicii cuantice. Aceasta este o speculație asupra relației dintre mecanica cuantică (QM) și relativitatea generală a lui Einstein,

având implicații profunde asupra naturii realității și a emergenței conștiinței.

În această interpretare, o stare cuantică rămâne în superpoziție până când diferența de curbură a spațiu-timp atinge un anumit prag critic, moment în care superpoziția colapsează într-o stare definită. Acest fenomen este numit auto-colaps (Objective Reduction – OR).

Această perspectivă se distanțează atât de:

- Interpretarea de la Copenhaga, conform căreia superpoziția eșuează în momentul în care un observator efectuează o măsurătoare (dar fără un mecanism obiectiv clar pentru colaps);
- Interpretarea „Many Worlds”, care afirmă că toate stările posibile ale unei superpoziții sunt la fel de reale, doar că incoerența dintre ele face imposibilă interacțiunea ulterioară între lumi.

În contrast cu aceste două interpretări dependente de observator, Penrose propune o teorie obiectivă a colapsului funcției de undă, în care prăbușirea funcției de undă este un proces fizic în sine, independent de observator.

Penrose argumentează că o funcție de undă nu poate fi susținută într-o stare de superpoziție dincolo de un anumit prag de energie, iar această limită este

determinată de influențele gravitaționale dintre particulele implicate.

Această ipoteză duce la conceptul de reducere obiectivă gravitațională (OR), un mecanism care ar putea explica tranziția dintre lumea cuantică și lumea clasică. Colapsul nu este produs de o măsurătoare externă, ci apare spontan, atunci când superpoziția creează o diferență suficientă în curbura spațiu-timp.

Punctul de plecare al acestei teorii este un criteriu denumit „un graviton”, inspirat din propunerea Diósi-Penrose. Aceasta oferă un prag fizic obiectiv pentru durata de viață a stărilor cuantice suprapuse, sugerând că fiecare eveniment OR – adică fiecare colaps spontan – ar putea reprezenta un moment de „experiență protoconștientă”.⁴³

Pentru a realiza acest lucru, este necesar ca superpoziția să evite incoerența imediată a mediului și să persiste până la atingerea unui anumit timp limitat.

Acest lucru se realizează argumentând că o superpoziție cuantică este:

– „orchestrată”, adică organizată corespunzător, impregnată de informații cognitive și capabilă de integrare și calcul;

⁴³ Ediție originală: Hameroff, S. & Penrose, R. (2014) “Consciousness in the universe: A review of the ‘Orch OR’ theory” *Physics of Life Review*, 11 (1): 39-78.

- izolată de un mediu aleatoriu insuficient de orchestrat, pentru ca superpoziția să evolueze spre colaps și apoi să creeze un moment de conștiință.

Un obstacol major al teoriei lui Penrose este că ecuația Schrödinger descrie formalismul mecanicii cuantice pentru un sistem aflat la temperatură zero absolut (0 K). Este absurd să presupunem că un creier conștient ar putea funcționa într-un astfel de mediu - dimpotrivă, temperatura creierului uman este considerabil mai mare.

De asemenea, știm că computerele cuantice moderne necesită temperaturi extrem de scăzute, apropiate de zero Kelvin, pentru a menține coerența cuantică și a preveni incoerența indusă de zgomotul termic. Astfel, se poate argumenta că menținerea fenomenelor cuantice relevante la temperaturi biologice ridicate este extrem de problematică și reprezintă un obstacol serios pentru întreaga teorie a conștiinței propusă de Penrose.

Cu toate acestea, biologia pare să fi găsit deja modalități de a stabiliza coerența cuantică în medii termice mult mai puțin controlate decât cele din laboratoarele de fizică. Așa cum am discutat anterior, există dovezi că plantele utilizează transportul de electroni în fotosinteză, păstrând coerența cuantică la temperatura

camerei.⁴⁴ Dar cum rămâne cu creierul uman? Există dovezi că procese cuantice pot avea loc la temperaturi biologice relevante?

În 2009, Anirban Bandyopadhyay și colegii săi de la Institutul Național de Știința Materialelor din Japonia au folosit nanotehnologia pentru a studia proprietățile electronice și optice ale microtubulilor individuali. Descoperirile lor au arătat că efectele cuantice pot apărea în microtubuli chiar la temperaturi biologice, sugerând că stări cuantice coerente ar putea exista în microtubuli la temperatura creierului uman indicând o posibilitate care validează, cel puțin parțial, ipoteza lui Penrose și Hameroff.⁴⁵

Un alt indiciu biologic provine nu de la un organism complex, ci de la un organism unicelular simplu: parameciul. Deși acest microorganism nu are niciun neuron, el poate naviga prin mediu, se poate hrăni și se poate reproduce. Cum reușește să realizeze aceste funcții fără un sistem nervos central? Prin utilizarea unor structuri intracelulare – microtubulii. Acest exemplu sugerează că microtubulii nu sunt doar structuri de susținere celulară, ci ar putea avea un rol

⁴⁴ Ediție originală: Hildner, R., Brinks, D. et al. (2013) "Quantum coherent energy transfer over varying pathways in single light harvesting complexes", *Science*, 340 (639): 1448-1451.

⁴⁵ Ediție originală: S. Sahu, S. Ghosh, K. Hirata, D. Fujita, A. Bandyopadhyay (2013) "Multi-level memory-switching properties of a single brain microtubule", *Applied Physics Letters*, 102: 123701.

fundamental în procesarea informației la nivel biologic.⁴⁶

Deși teoria conștiinței bazată pe reducerea obiectivă (OR) este complexă și controversată, ea își propune să abordeze două dintre cele mai profunde probleme științifice și filosofice ale secolului XXI:

- *Elucidarea naturii conștiinței umane*, una dintre cele mai mari enigme ale științei.
- *Împăcarea a două teorii aparent ireconciliabile într-o viziune unificată* – mecanica cuantică și teoria relativității generale a lui Einstein. Aceasta este aspirația unei „Teorii a totului” (Theory of Everything), o idee fundamentală în fizica modernă.



Imagen 12: "Generativo" © Gaia Amoroso

⁴⁶ Original publication: Nakagaki, T., Yamada, H. e Toht, Á. (2000) "Maze – solving by an amoeboid organism", *Nature*, 407: 470.

Cu această propunere, Penrose oferă informații despre rezolvarea acestor două provocări simultan.

După această introducere, sper că îți va fi mai ușor să urmărești dialogurile cu cei patru gânditori excepționali cu care am avut onoarea de a explora natura conștiinței. Indiferent dacă această călătorie te lasă mai lămurit sau mai confuz, cred că putem fi de acord asupra unui lucru: o confuzie clară este întotdeauna mai valoroasă decât o certitudine neclară.

DIALOGURI I

Conștiința

DIALOGURI I

Conștiința

VI. Dialogul cu David Chalmers



David Chalmers este profesor de filosofie și științe neuronale și co-director al Centrului pentru Minte, Creier și Conștiință de la Universitatea din New York. De asemenea, este profesor distins de filosofie la Universitatea Națională din Australia. Și-a obținut doctoratul în filosofie și științe cognitive la Universitatea Indiana, în cadrul grupului de cercetare condus de Douglas Hofstadter, un pionier în domeniul inteligenței artificiale.

Chalmers este renumit pentru formularea conceptului de „problemă grea” a conștiinței și pentru contribuțiile sale asupra ideii de „minte extinsă”, conform căreia tehnologia pe care o utilizăm poate deveni efectiv parte din mintea noastră. De asemenea, a influențat dezbaterile filosofice despre limbaj, metafizică, tehnologie și inteligență artificială. Este co-fondator și fost președinte al Asociației pentru Studiul Științific al Conștiinței și co-director al Fundației PhilPapers.

Este autorul Reality+: Virtual Worlds and the Problems of Philosophy (2022), The Conscious Mind (1996), The

Character of Consciousness (2010) și Constructing the World (2014). Mai multe informații: <https://consc.net/>

În cartea dumneavoastră *The Conscious Mind* (1996), ați introdus „problema dificilă a conștiinței” în contrast cu „problema ușoară”, care privește corelațiile neuronale ale conștiinței și implicațiile lor asupra comportamentului. După toți acești ani, mai considerați că problema conștiinței este la fel de „grea” ca atunci?

David Chalmers: Aș spune că am progresat în înțelegerea problemei în diverse moduri, deci există cu siguranță o mișcare înainte. Cu toate acestea, nu cred că cineva a reușit încă să rezolve „problema grea”, iar în principiu, ea rămâne la fel de dificilă.

Distincția de bază pe care am făcut-o este între problemele „ușoare” ale conștiinței privind explicarea diverselor funcții comportamentale și cognitive, pentru care avem deja un cadru științific solid, și problema grea, care privește experiența subiectivă. În acest caz, metodele standard din științele cognitive nu oferă o paradigmă clară pentru a explica de ce toate aceste procese dau naștere unor experiențe subiective.

În cele din urmă, este posibil ca singura cale de a aborda problema grea să fie schițarea unor principii fundamentale care să lege conștiința de procesele fizice

- iar teorii precum IIT încearcă exact acest lucru. Cu toate acestea, nicio teorie nu a ajuns la un nivel de validare empirică suficient de puternic pentru a fi acceptată ca o explicație definitivă a conștiinței.

Această separare între problema "ușoară" și problema "greă" persistă. Dacă ne uităm la modelele obișnuite din științele cognitive, vedem că ele nu abordează esențialul "problemei dificile". Prin urmare, avem nevoie de ceva nou.

Între timp, oamenii au explorat o mulțime de idei diferite care sunt mult mai bine înțelese decât în urmă cu 30 de ani, fie că este vorba de panpsihism, fie de abordarea mecanicii cuantice, fie de teoria iluzionismului. S-au făcut progrese semnificative în fiecare dintre aceste domenii, însă, fundamental vorbind, problema greă rămâne la fel de "dificilă" ca întotdeauna.

Cum vă imaginați o abordare științifică riguroasă pentru înțelegerea naturii conștiinței? Ar implica acest lucru o bază solidă în matematică și validare formală, sau credeți că este esențial să integrați datele empirice cu dovezi matematice? Cu alte cuvinte, ce criterii considerați relevante pentru dezvoltarea unei științe serioase a conștiinței?

David Chalmers: Văd știința conștiinței ca o disciplină care trebuie să integreze datele obiective (persoana a III-a), provenite din descrierile sistemelor cognitive, cu datele subiective (persoana I), obținute prin introspecție și experiență directă.

Datele obiective sunt accesibile prin metodele standard ale științei – măsurători, observații ale comportamentului și analize ale activității cerebrale.

Pe de altă parte, datele subiective provin din experiența conștientă și introspecția individuală. Obiectivul final al științei conștiinței este de a formula principii riguroase care să conecteze aceste două domenii, integrând perspectiva subiectivă cu cea obiectivă într-un cadru coerent.

Acest lucru va necesita mai mulți pași relevanți. În timp ce metodele pentru studierea conștiinței dintr-o perspectivă obiectivă (persoana a III-a) sunt deja bine dezvoltate, metodele pentru investigarea experienței subiective (persoana I) sunt încă primitive. Avem nevoie de tehnici mai precise pentru a cartografia trăirile conștiente într-un mod reproductibil și validat științific.

De asemenea, vom avea nevoie de metode de formulare a structurii conștiinței în termeni riguroși. Ați menționat matematica: aici este un loc în care cred că matematica poate interveni deoarece putem lucra pentru a găsi descrieri matematice ale structurii conștiinței. Și s-ar putea ca, cel puțin parțial, principiile

matematice să conecteze structura proceselor fizice (adică, datele la persoana a III-a) cu structurile experienței subiective (adică, datele la persoana I).

Un exemplu în această direcție este Teoria Informației Integrate (IIT), care încearcă să cuantifice gradul de conștiință al unui sistem prin măsurători matematice ale integrării informaționale (ϕ - φ). Cred că este o formă foarte promițătoare pentru o știință a conștiinței. Adică, o descriere matematică a conștiinței nu va epuiza conștiința. Acest lucru poate fi ilustrat de experimentul de gândire „Mary în camera alb-negru”. Deși Mary ar putea cunoaște toate informațiile matematice despre culori înainte de a ieși din cameră, prima experiență directă a culorii roșii îi oferă o informație nouă, inaccesibilă prin cunoașterea pur formală.

Cu toate acestea, o formulare matematică a conștiinței poate reprezenta cel puțin o descriere foarte utilă a structurii conștiente. Aceasta ar putea duce la dezvoltarea a ceea ce unii cercetători numesc „fenomenologie obiectivă” – un mod de a descrie experiențele subiective într-un cadru științific riguros.

În cea mai recentă carte a dumneavoastră, *Reality+: Virtual Worlds and the Problems of Philosophy (2022)*, explorați în profunzime „ipoteza simulării”. În esență, susțineți că scenariile science-fiction, precum cel din *The Matrix*, ar putea fi mai plauzibile

decât sunt percepute în mod obișnuit. Această ipoteză sugerează că realitatea noastră actuală ar putea fi de fapt o lume simulată, creată de o civilizație avansată. În plus, susțineți că ființele din astfel de simulări ar putea avea o conștiință asemănătoare cu a noastră. Ați putea detalia raționamentul din spatele acestei posibilități?

David Chalmers: Această discuție se leagă, parțial, de vechea întrebare: este conștiința dependentă de biologie sau este mai degrabă o chestiune de procesare a informației, de calcul și organizare funcțională?

Întotdeauna am fost de partea celor care susțin că nu biologia este esențială pentru conștiință, ci mai degrabă organizarea funcțională și procesarea informației.

Un mod în care am ilustrat această idee, încă din anii '90, în *The Conscious Mind*, a fost printr-un experiment de gândire: ce s-ar întâmpla dacă am înlocui progresiv fiecare neuron cu un cip de siliciu, păstrând cu fidelitate organizarea funcțională? Argumentul meu este că, atâta timp cât această tranziție se face într-un mod care conservă suficient de bine structura și procesarea informației, conștiința ar fi păstrată.

Desigur, acest scenariu nu ne duce direct la simulări, ci mai degrabă la ideea unor izomorfi din siliciu care ar fi conștienți la fel ca noi. Dar, odată ce am ajuns până aici,

cred că nu este decât un pas până la concluzia că și ființele simulate ar putea fi conștiente.

De ce? În termeni simpli, o simulare a creierului meu ar fi, în esență, analogă cu un izomorf din siliciu al creierului meu: ar avea multe componente care interacționează și care procesează aceleași informații. Dacă ceea ce contează este structura și procesarea informației, atunci, în principiu, toate acestea ar putea fi reproduse într-o simulare suficient de fidelă.

Sigur, problema dificilă a conștiinței rămâne nerezolvată deoarece nu înțelegem încă cum o simulare ar putea da naștere conștiinței. Dar, în egală măsură, nici nu înțelegem cum un creier biologic generează conștiință. Argumentul meu este că o simulare bine realizată se află pe același nivel cu un creier biologic în această privință.

Interesant. Totuși, aș argumenta că și procesarea informației este inerent legată de substratul material al sistemului. Cu alte cuvinte, există ceva fundamental unic în biologia organică, modelată de procesul evolutiv, care servește drept fundație pentru experiența conștientă la oameni. Totuși, presupun că nu împărtășiți această perspectivă, susținând că compoziția materială nu este

esențială, atâta timp cât structura funcțională rămâne intactă. Este corect?

David Chalmers: Aș fi de acord că substratul poate influența semnificativ modul în care informația este procesată. De exemplu, structura neuronilor joacă un rol fundamental în procesarea informației în creier. Cu toate acestea, nu văd de ce acest lucru nu ar putea fi simulat, cel puțin în principiu. Oricare ar fi particularitățile neuronilor, nu există, din câte îmi dau seama, nimic care să fie fundamental necalculabil sau nesimulabil.

Acum, dacă Sir Roger Penrose are dreptate și există procese mecanico-cuantice speciale în neuroni, atunci ar putea exista limitări ale simulării pe un computer clasic. În acest caz, ar trebui să explorăm mai multe direcții.

Dacă teoria lui Penrose este corectă, este posibil ca gravitația cuantică să implice procese care nu pot fi simulate nici măcar pe un computer cuantic obișnuit.

Însă chiar și atunci, am putea dezvolta computere gravitaționale cuantice, care să utilizeze acest nou tip de calcul. În principiu, nu văd de ce nu am putea construi o simulare a unui creier pe un astfel de sistem. Acest tip de simulare ar merge dincolo de simularea pe un computer clasic, dar cred că ar fi totuși o abordare interesantă a simulării.

Problema conștiinței a fost dezbătută în filosofie timp de sute, chiar mii de ani. Totuși, unii susțin că filosofia nu a adus contribuții semnificative în această direcție. Cum pot filosofii să contribuie activ la găsirea unei soluții pentru problema grea a conștiinței? Care este, în opinia dumneavoastră, rolul real al filosofiei în această dezbatere?

David Chalmers: Ce poate face filosofia? O întrebare foarte bună. Cred că filosofii au explorat multiple direcții pentru a aborda problema grea, iar în ultimele două decenii am asistat la multe contribuții importante. De exemplu, există o revigorare a panpsihismului, dar și noi încercări de a aborda problema combinației (adică modul în care conștiința individuală a particulelor ar putea genera conștiința unificată a unui organism). Filosofi precum Philip Goff, Hedda Mørch și Galen Strawson au propus soluții interesante în această direcție.

Avem și teoria iluzionismului despre conștiință, dezvoltată de oamenii de știință cognitivă, dar a fost cu adevărat împinsă de o mulțime de filosofi, printre care Dan Dennet, Keith Frankish și alții.

Așadar, filosofii au contribuit la dezvoltarea unor posibile soluții pentru anumite aspecte ale problemei grele. Personal, nu mă preocupă prea mult dacă o soluție vine de la un filosof sau de la un om de știință –

Însă cred că orice progres va necesita o interacțiune între filosofie și știință.

Pornind de la ideile lui Sir Roger Penrose, credeți că avem nevoie de un nou „tip” de fizică pentru a face progrese în rezolvarea problemei grele a conștiinței? Ce teorii credeți că ar putea avea succes în abordarea naturii experienței subiective?

David Chalmers: Ce fel de teorie ar putea explica conștiința? Îmi place ideea „matematicii conștiinței” sau cartografierea structurii conștiinței și corelarea acesteia cu procesele fizice din creier.

Părerea mea este că o teorie completă a conștiinței ar putea necesita legi psihofizice fundamentale care leagă procesele fizice de experiența conștientă.

Panpsihismul este o posibilitate, dar nu este singura. Un dualist ar putea, de asemenea, să susțină existența unor legi psihofizice care nu reduc conștiința la materie, dar explică legătura dintre conștiință și fizică. Teoria Informației Integrate (IIT), propusă de Giulio Tononi, este o altă abordare posibilă, care încearcă să cuantifice conștiința pe baza integrării informaționale.

Depinde foarte mult de dezvoltarea științei. O obiecție majoră la o teorie pur matematică a conștiinței este că aceasta ar fi încă supusă problemei lui Mary. Așadar, aș susține că ar trebui să facem distincție între caracterul

calitativ al conștiinței de caracterul său structural, cel care ar putea, cel puțin, să ajungă la legi obiective matematice psihofizice care pot explica caracterul structural al conștiinței. S-ar putea să mai fie ceva de făcut cu privire la caracterul calitativ, cum ar fi “roșeața roșului”.

Unii neurocercetători și filosofi consideră că conștiința este o eroare de categorie, un concept care nu mai are loc în vocabularul științific și filosofic modern. De ce ar trebui să ne pese de problema conștiinței? Ce argumente pot fi aduse în favoarea relevanței acestei dezbateri?

David Chalmers: Nu pot forța pe nimeni să fie interesat de conștiință – într-un anumit sens, asta depinde de fiecare om. Există multe aspecte ale lumii care pot fi fascinante, iar nu toată lumea trebuie să fie captivată de problema grea a conștiinței. Este perfect legitim.

Dacă scopul tău este pur și simplu să prezici comportamentul altor oameni, s-ar putea să poți ignora conștiința. Însă preocupările noastre merg dincolo de simpla observare a comportamentului – ne pasă de alți oameni nu doar pentru ce fac, ci și pentru ceea ce experimentează. Acest lucru are implicații directe asupra eticii și politicii. Luați, de exemplu, întrebarea „Cum ar trebui să tratăm animalele?”. Pentru a răspunde la această întrebare, este necesar să știm

dacă animalele sunt conștiente și ce fel de stări de conștiință au.

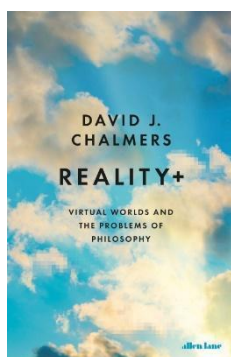
De exemplu: „suferă ele?” și puteți spune „pot să știu totul din comportamentul animalului”. Grozav, dar asta nu îmi spune lucrul cel mai relevant pe care trebuie să-l știu, și anume cum ar trebui să tratez moral animalele, adică să știu ceva despre experiența lor subiectivă.

Dacă considerați că experiența subiectivă este într-adevăr la baza valorii și a sensului în viața noastră, atunci acesta este ceva cu care ne vom ocupa pentru a găsi răspunsul la unele dintre aceste întrebări practice importante. Acesta este cel puțin un motiv să-ți pese de conștiință.

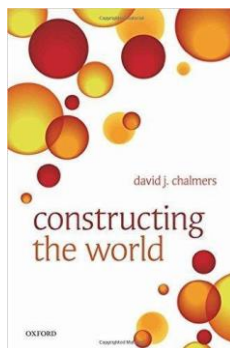
Pe lângă implicațiile morale, există și motive intelectuale pentru care conștiința este importantă. Este una dintre cele mai mari anomalii din viziunea noastră asupra universului. Dacă nu avem un cadru teoretic care să o includă, atunci nu avem o teorie completă a realității. Conștiința este un fenomen fundamental al existenței noastre, iar ignoranța în această privință lasă un gol în înțelegerea noastră asupra universului.

Oricum, sunt două motive pe care le pot oferi pentru a argumenta despre importanța conștiinței.

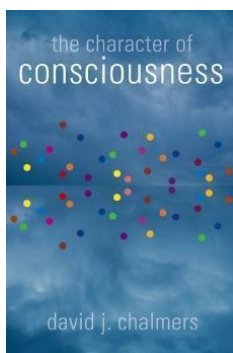
Cărți scrise de David Chalmers



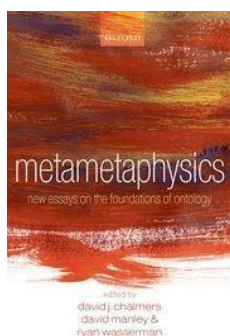
2022



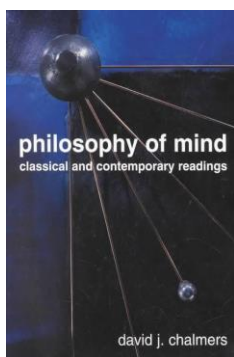
2012



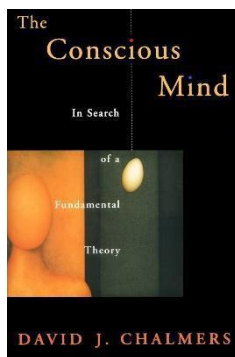
2010



2009



2002



1996

VII. Dialogul cu Susan Blackmore



Susan Blackmore este scriitoare, lector și profesor invitat la Universitatea din Plymouth, Regatul Unit. A absolvit psihologia și fiziologia la Universitatea din Oxford în 1973 și și-a obținut masterul și doctoratul la Universitatea din Surrey în 1980. Activitatea sa de cercetare se concentrează pe conștiință, teoria evoluționistă, memorii și meditație.

Pe lângă contribuțiile sale academice, Blackmore scrie articole pentru numeroase reviste și ziare, menține un blog pe Psychology Today și participă frecvent la emisiuni de radio și televiziune, unde abordează teme legate de conștiință și experiențe neobișnuite. Este autoarea a peste cincisprezece cărți, șaiszeci de articole academice și aproximativ optzeci de capitole în volume colective.

Unele dintre cărțile ei sunt: *Dying to Live: Near Death Experiences* (1993), *The Meme Machine* (1999), *Conversations on Consciousness* (2005), *Zen and the Art of Consciousness* (2011), *Seeing Myself: What Out-of-body Experiences tell us about life and consciousness: an introduction* (2018).

Mai multe informații la: www.susanblackmore.uk

Ați fost unul dintre primii savanți care a acordat o importanță filosofică majoră stărilor modificate de conștiință. Să începem cu experiența neobișnuită pe care ați avut-o în tinerețe – o experiență extracorporală care v-a influențat profund interesul pentru conștiință. Ne puteți povesti despre această experiență și despre impactul pe care l-a avut asupra direcției dvs. de cercetare?

Susan Blackmore: Întreaga experiență a durat aproximativ două ore și jumătate. Eram în camera unui prieten, în primul meu semestru la Oxford, aveam 19 ani și eram extrem de obosită, lipsită de somn și fumam puțin cannabis, dar nu suficient pentru a explica complet ceea ce s-a întâmplat – deși probabil a contribuit.

Am început să cobor printr-un tunel – un tunel de copaci, cu frunze de jur împrejur, îndreptându-mă spre o lumină. Acest lucru s-a întâmplat înainte ca termenul „experiență din apropierea morții” să fie inventat, așa că nu aveam nicio referință pentru ceea ce trăiam. Nu știam nimic despre aceste fenomene.

Tunelul a dus la o experiență în afara corpului. Simțeam că mă află în afara corpului meu și puteam să mă uit în jos și să-mi văd propriul corp, în timp ce încă vorbeam cu un prieten. Acesta mă întreba constant: „Ce poți vedea acum?”, iar eu continuam să călătoresc – sau cel puțin așa mi se părea – în jurul lumii, văzând tot felul de lucruri.

Pentru a scurta povestea, la un moment dat am încercat să mă întorc în corpul meu, dar a fost extrem de dificil. M-am simțit micșorându-mă tot mai mult, până când aproape am dispărut, ceea ce m-a speriat teribil. Apoi, brusc, m-am extins, din ce în ce mai mult, într-o experiență mistică clasică - despre care, din nou, nu știam nimic la acel moment. Am devenit una cu totul. Eu nu mai existam, dar exista ceva ce mi se părea că este totul.

S-au întâmplat multe alte lucruri în acest proces, dar în cele din urmă am fost epuizată. Prietenul meu m-a întrebat: „Ei bine, nu mai există altceva?“, iar eu m-am gândit: „Nu, pentru că sunt totul. Cum ar putea exista altceva?“ Gândurile continuau, dar nu mai erau sub formă de cuvinte.

Mi-a luat două zile să mă readaptez și să simt că sunt din nou în corpul meu.

Această experiență m-a făcut inițial să cred în tot felul de fenomene paranormale, ceea ce, retrospectiv, era destul de illogic, dar de înțeles pentru o tânără de 19 ani, fără cunoștințe în acest domeniu. La începutul anilor '70, neuroștiința nu oferea încă explicații solide pentru astfel de experiențe.

Din acest motiv, am ajuns să cred în telepatie, clarviziune, psihokineză, fantome, spirite, poltergeist, viață după moarte și existența sufletului. Această

experiență mi-a influențat cariera. Am refuzat un doctorat convențional într-o instituție de prestigiu și am ales să-mi creez propriul doctorat despre paranormal.

Mi-au luat aproximativ cinci ani de cercetare intensivă a fenomenelor paranormale pentru a ajunge la concluzia că, aproape sigur, ele nu există. Aceasta a fost transformarea mea: de la a crede orbește în fenomene misterioase, la a pune întrebările fundamentale despre conștiință.

Am realizat că întrebările esențiale nu sunt „Există un spirit, un suflet sau o conștiință separată de creier?” sau „Există o viață de după moarte?” și nici „Există lumi de dincolo sau spirite?”. Adevărata întrebare este „Ce este această experiență? Cum creează creierul aceste fenomene? Cum interacționează creierul, corpul și lumea pentru a genera conștiință?”. Aceste întrebări sunt mult mai interesante decât preocupările despre viața de apoi sau fenomene supranaturale. De atunci, acestea au rămas în centrul cercetării mele.

În urma acestei experiențe, știm că diverse medicamente pot induce stări alterate de conștiință. Credeți că putem folosi psihedelice precum DMT sau LSD, de exemplu, pentru a studia natura minții conștiente? Considerați că aceste metodologii sunt utile în înțelegerea fenomenului

subiectivității, având în vedere dificultățile metodologice actuale în investigarea științifică a conștiinței?

Susan Blackmore: Depinde ce înțelegeți prin „a studia natura minții”. Dacă vă așteptați la răspunsuri imediate, atunci nu. Dar, într-un sens mai general, da. Totul depinde de ce vreți să spuneți: dacă vorbim despre experiența personală, despre felul în care cineva poate începe să înțeleagă conștiința trăind direct aceste stări, atunci sigur, psihedelicile pot oferi o perspectivă diferită. Dacă însă ne referim la cercetarea științifică propriu-zisă, la posibilitatea ca aceste substanțe să ne ofere o teorie solidă despre conștiință, asta rămâne încă de văzut.

Motivul pentru care spun asta e simplu: mintea este atât de radical schimbată într-un „trip”, încât îți poate arăta câteva lucruri imediat. De exemplu, îți poate demonstra că starea obișnuită de conștiință nu este singura posibilă. Gândiți-vă la ce spunea William James în 1890 – că dincolo de realitatea obișnuită există nenumărate alte moduri de a fi conștient, doar că sunt foarte aproape de noi, iar pentru a le accesa avem nevoie de un declanșator, cum ar fi un drog psihedelic sau o experiență mistică spontană.

Una dintre cele mai interesante cercetări recente pe acest subiect este studiul lui Timmermann și

colaboratorii săi din 2019⁴⁷, despre DMT. După cum probabil știți, DMT este principalul compus psihoactiv din ayahuasca și are un efect profund asupra rețelelor funcționale ale creierului.

Creierul uman funcționează prin mai multe rețele interconectate, dar una dintre cele mai importante este rețeaua de mod implicit (DMN). Aceasta este responsabilă pentru sentimentul de sine – e acea rețea de conexiuni cu rază lungă de acțiune care adună toate informațiile despre corpul tău, amintirile, opiniile despre tine însuși și modul în care te vezi în lume. Se leagă de joncțiunea temporoparietală dreaptă, unde schema corporală interacționează cu sistemele de control din cortexul frontal și cu memoria din lobii temporali.

Ce efect are DMT asupra DMN? Această rețea a sinelui este profund perturbată, iar în alte părți ale creierului, cum ar fi cortexul vizual, activitatea locală crește, ceea ce poate explica halucinațiile vizuale intense pe care le provoacă. Când sistemul de „selfing” este destabilizat, sentimentul unui sine separat începe să dispară.

Acum, ce înseamnă asta? Aș spune că înseamnă că putem vedea dincolo de iluzia sinelui conștient – sinele

⁴⁷ Timmermann, C., Roseman, L., Schartner, M. et al. (2019) “Neural correlates of the DMT experience assessed with multivariate EEG”, *Sci Rep*, 9: 16324.

pe care îl percepem ca fiind centrul experienței noastre, cel care pare să aibă conștiință și liber arbitru. În opinia mea, ambele sunt iluzii. Iar faptul că DMT le poate descompune pentru câteva ore – mai mult, dacă este ayahuasca, sau doar 15 minute în cazul DMT pur – îți permite să trăiești fără această percepție obișnuită a sinelui.

Asta, cel puțin, te poate ajuta să vezi începuturile unui aspect al iluziilor conștiinței. Problema este că majoritatea teoriilor despre conștiință sunt încă blocate în aceste iluzii și nu ne duc nicăieri. Poate că psihedelicele ne pot ajuta să ieșim din acest impas. Cred că cercetările recente încep cu adevărat să dezvăluie multe despre cum funcționează mintea.

Și mai avem multe de descoperit. Acum, că restricțiile legale nu mai blochează cercetarea psihedelică așa cum s-a întâmplat decenii întregi, cred că urmează să învățăm lucruri cu adevărat fascinante.

În domeniul explorării conștiinței, în special prin fenomene precum visele lucide, unde experiența oglindește îndeaproape starea de veghe, mai degrabă decât somnul profund, cum percepeți contribuțiile potențiale ale investigațiilor empirice asupra acestor vise? Pot astfel de studii să ofere perspective valoroase și să ajute la cartografierea și

Înțelegerea cuprinzătoare a diverselor stări de conștiință?

Susan Blackmore: Când spuneți că visele lucide sunt „aproape de starea de veghe”, nu sunt sigură la ce vă referiți exact. Din ceea ce știm până acum – și, desigur, există variații – visele lucide au loc, în cea mai mare parte, în timpul somnului REM. Asta înseamnă că trebuie să fi adormit profund și să fi trecut prin mai multe cicluri de somn pentru ca un vis lucid să apară. Sigur, există excepții, cum ar fi atunci când cineva este extrem de epuizat și intră direct în REM, dar în general, visele lucide apar în fazele REM mai avansate ale somnului.

Dacă prin „aproape de starea de veghe” vă referiți la faptul că visele lucide apar în perioadele mai active ale REM, unde activitatea neuronală este mai intensă decât în etapele inițiale ale somnului REM, atunci da, sunt de acord. Dar ce este cu adevărat interesant aici – și asta ne duce din nou la rețeaua de mod implicit – este că în timpul viselor lucide, conexiunile dintre joncțiunea temporoparietală și lobiile frontali devin mai puternice. Această modificare neuronală pare să fie baza psihologică a momentului în care cineva își spune „oh, eu sunt aici acum și pot controla visul!”.

Îmi place să aflu aceste detalii pentru că ele distrug multe dintre teoriile alternative despre visele lucide – idei despre suflete care se trezesc, călătorii astrale și tot

felul de interpretări mistice pe care oamenii tind să și le imagineze. Pe de altă parte, este puțin frustrant că psihologia standard nu acordă prea multă atenție viselor lucide. Nu sunt suficient de mulți cercetători care le studiază cu adevărat, dar cred că sunt o resursă incredibilă pentru a înțelege mai bine cum funcționează creierul și mintea noastră.

Cu experiența dvs. vastă de aproximativ 30 de ani în meditația zen, cum v-a influențat această practică perspectiva asupra minții conștiente? În ce moduri credeți că meditația, cu accent pe atenție și conștientizare de sine, oferă perspective asupra naturii conștiinței? În plus, luând în considerare aspectele subiective și introspective ale meditației, cum abordați provocările încorporării unor astfel de perspective în înțelegerea noastră a unui studiu științific mai larg al conștiinței?

Susan Blackmore: Depinde ce înțelegeți prin „înțelegere”, pentru că nu sunt sigură că meditația are acest scop. Cred că, în primul rând, este menită să clarifice mintea – și acesta este, de departe, cel mai bun motiv pentru a vrea să meditezi. Sigur, într-un fel, m-a ajutat să înțeleg lucruri despre minte, dar practica principală este despre a renunța la iluzii, la credința că „eu sunt aici” și „lumea este acolo”, la dorința de a deveni cineva mai important.

Și dorința (mea) de a fi importantă este tot acolo, chiar și după 40 de ani de practică (așa că poate că nu funcționează atât de bine, haha!). Dar, încetul cu încetul, îți schimbă fundamental modul de a gândi. Te face mai deschis la ceea ce este, mai acceptant, mai puțin atașat – și asta este, în esență, ceea ce urmărește meditația.

Poate să ne ajute să înțelegem mintea? Da, bineînțeles. Gândiți-vă la meditorii cu experiență, cei care sunt foarte bine antrenați să intre în stări modificate ale conștiinței. Ei pot deveni subiecți de studiu pentru cercetări care să ne ofere o perspectivă mai clară asupra modului în care funcționează conștiința.

În ultimii ani, am devenit foarte interesată de meditația Jhanas, unde se spune că există opt stări discrete de conștiință care pot fi accesate prin concentrare profundă. Am auzit prima oară despre asta cu zeci de ani în urmă și m-am gândit cât de uimitor ar fi dacă aceste stări ar exista cu adevărat și ar putea fi accesate doar printr-un set precis de instrucțiuni.

Am crezut că nu voi fi niciodată capabilă să le experimentez. Dar apoi un mare profesor Jhanas a venit în Anglia, l-am întâlnit și imediat am vrut să particip la retragerile lui. Am fost la mai multe și am practicat acest tip specific de meditație. Spre surprinderea mea, am constatat că este posibil, doar urmând aceste instrucțiuni destul de intense, să ajungi în aceste stări

modificate clar demarcate, și mai mult decât atât, să treci între ele.

Această practică începe cu o concentrare profundă, urmată de alte tehnici pentru a intra în aceste stări distincte. Până acum, pot accesa destul de bine primele trei și poate și a patra, dar continui să exersez și poate într-o zi voi reuși să ating și celelalte.

Din ce am experimentat eu, sunt absolut convinsă că acești practicanți care au dezvoltat această formă de meditație chiar au descoperit aceste stări și au învățat cum să ajungă la ele. Asta ne spune ceva despre ce este capabilă mintea umană. Când poți face asta – sau măcar când citești și înțelegi ce spun cei care o practică – îți dai seama că mintea poate intra într-o relație complet diferită cu lumea și cu propriul concept despre sine.

Așa că, da, răspunsul este în general da.

Se pare că diferitele stări modificate ale conștiinței pot fi identificate prin tehnici de imagistică neuronală. Având în vedere experiența dumneavoastră în această meditație specifică, credeți că este plauzibil ca printr-o practică suficientă, modele distincte de activitate cerebrală corespunzătoare celor opt faze ale acestei meditații particulare să poată fi observate prin

neuroimagini? Cum ar putea integrarea neuroștiinței și a practicilor de meditație să ne îmbunătățească înțelegerea conștiinței, în special în dezlegarea corelațiilor neuronale asociate cu stările specifice ale conștiinței modificate?

Susan Blackmore: Din câte știu eu, au existat doar două experimente în acest sens, ambele realizate cu Leigh Brasington. Este foarte costisitor să faci experimente pe creier cu fMRI sau cu alte tipuri de tehnologie imagistică.

Problema este că subiectul trebuie să fie capabil, în această stare de concentrare foarte profundă, să le indice cercetătorilor când trece la următoarea stare și să facă legătura cu activitatea sa cerebrală. Cu siguranță, cercetătorii au constatat schimbări în momentul în care acesta a spus că trece de la starea “doi” la “trei”, de exemplu.

Dar nu avem suficiente date și nu cred că Leigh a reușit să parcurgă cele opt stări, deoarece ultimele două sunt ceea ce am putea numi „inconștiente”.

Adică, sunt atât de departe, încât ar fi foarte greu să le comunicăm în acel cadru experimental - dar ar fi cu siguranță incredibil. Este foarte, foarte greu să meditezi cum trebuie într-un scanner fMRI, cu tot zgomotul și toate celelalte. Așadar, cred că sunteți puțin cam optimist în această privință.

Cercetarea conștiinței ridică provocări semnificative, iar persoane precum Deepak Chopra promovează adesea idei spirituale îndoielnice despre mintea conștientă. În cadrul discuțiilor dvs. cu el pe astfel de subiecte, care este perspectiva dvs. asupra acestor idei spirituale? Cum abordați echilibrul dintre explorarea misterelor conștiinței și menținerea rigurozității științifice, în special atunci când abordați concepte care pot fi lipsite de suport empiric sau de o bază științifică?

Susan Blackmore: Dacă cineva este interesat de Deepak Chopra, vă recomand să căutați pe YouTube dezbateră mea cu el și să urmăriți videoclipul⁴⁸. O să vedeți un guru spiritual autoproclamat care, sincer, se comportă surprinzător de prost.

Nu pot răspunde pentru toate ideile din această categorie, dar pot vorbi despre Deepak Chopra. El susține că conștiința este primară și că materia nu există. Problema este că această afirmație nu înseamnă absolut nimic – e doar o serie de cuvinte goale. Majoritatea lucrurilor pe care le spune sunt lipsite de sens.

Acum, să fiu corectă, unele dintre sfaturile sale practice despre minte, despre antrenarea atenției și despre meditație sunt foarte bune. Chiar înțelege aceste

⁴⁸ Video original: youtube.com/watch?v=_ZFGkqhNhgM.

aspecte și poate fi un ghid util pentru cineva care vrea să înceapă să mediteze. Dar când trece la teoria lui generală despre conștiință, devine complet nesatisfăcător.

Luați o problemă filosofică simplă, problema minte-corp. Dacă ești materialist, susții că totul este materie, dar atunci ai dificultăți în explicarea experienței subiective – și aici intrăm în ceea ce David Chalmers a numit „problema grea a conștiinței”. Poate că, de fapt, punem întrebarea greșită, poate că nici măcar nu este o problemă reală, dar totuși, este o întrebare deschisă.

Pe de altă parte, dacă ești idealist, spui că totul este conștiință, dar atunci nu poți explica materia. Deci, evident, nici una dintre aceste extreme nu funcționează prea bine. Dualismul? Nici el nu funcționează, pentru că ai două tipuri de „substanțe” complet diferite care trebuie să interacționeze, iar asta creează mai multe probleme decât soluții.

Așadar, avem o problemă fundamentală aici – motiv pentru care eu mă consider un iluzionist când vine vorba de conștiință. Dar Deepak Chopra? El e pur și simplu un idealist, doar că nu vrea să recunoască asta.

Ce face el? Spune doar că „totul este conștiință” și că „materia nu există”. Minunat! Și acum ce? Această idee nu face nicio predicție, nu are nicio bază teoretică, nu

oferă nicio concluzie practică, dar cu toate acestea, oamenii o adoră!

De ce? Pentru că sună bine. Oamenii vor să creadă că propria lor conștiință este magică și importantă, că au puterea de a modela realitatea doar prin gândurile lor. Deepak le oferă această iluzie: „Oh, el vorbește despre conștiință, deci trebuie să fie profund, trebuie să fie un om spiritual!”.

Bineînțeles, îl caricaturizez un pic, dar în esență, ideile sale sunt complet lipsite de conținut științific. Și mai rău decât atât, apelează la cercetarea științifică și o denaturează în mod oribil pentru a o face să pară că îi susține teoriile.

Acum, ca să fie clar, nu vorbesc în numele tuturor profesorilor și liderilor spirituali. Sunt mulți și sunt extrem de variați. Unii chiar au perspective interesante, alții sunt doar vânzători de iluzii. Deepak Chopra? Din punct de vedere științific, nu este decât un exemplu foarte bun despre cum să spui multe fără să spui nimic.

Respingeți atât materialismul, cât și idealismul în înțelegerea conștiinței. Poate fi văzut iluzionismul ca o sinteză, propunând că creierul, ca entitate materială, generează iluzia conștiinței?

Susan Blackmore: Asta e o întrebare foarte bună. Cineva mi-a pus aceeași întrebare la o conferință despre psihedelize și tot ce pot să spun este: majoritatea iluzioniștilor sunt materialişti, asta e sigur. Dar, sincer, nu am cu ce să înlocuiesc materialismul.

Iluzionismul, așa cum îl văd eu, este destul de larg: toate modurile existente de a gândi conștiința sunt greșite și încă nu am găsit o cale de ieșire din această dilemă. Iar faptul că o numim „problema grea” și încercăm să explicăm cum apare conștiința din creier – exact așa cum o descrie Chalmers – este, în esență, un mod de a gândi dualist.

Și totuși, toți acești materialişti – cu excepția lui Chalmers, care a inventat termenul „problema grea” – continuă să fie materialişti fără să rezolve cu adevărat problema. Ei spun, la fel ca și mine: „Dacă credem că conștiința este ceva ce avem, care are putere, care face ceva și care a evoluat cu un scop, atunci înțelegem greșit realitatea.”

Dennett și Frankish, de exemplu, ar spune: „Da, există un creier material, iar din el se construiesc iluziile.” Astfel, ei încearcă să rezolve problema modului în care apar aceste iluzii. În loc să se ocupe de problema grea, o înlocuiesc cu problema iluziei sau cu ceea ce Chalmers numește „meta-problema conștiinței”.

Și, sincer, sunt foarte mulțumită de această schimbare. Nu mă consider filosof, dar dacă cineva mă întreabă, aș spune că sunt monistă neutră. Monistă, pentru că dualismul pur și simplu nu funcționează. Dar, în același timp, nu cred că conceptele noastre actuale de „materie”, „minte” sau „material” sunt utile.

Nu știu din ce este făcut universul. Poate că fizica fundamentală ne va ajuta să înțelegem asta. În prezent, fizicienii sunt încă blocați încercând să combine mecanica cuantică cu teoriile einsteiniene.

Ne putem gândi și la noile teorii informaționale, care susțin că tot ceea ce există este informație. Sau putem lua în considerare termodinamica și entropia. Poate că aici se află cheia pentru a rezolva problema conștiinței.

Momentan, încerc să învăț și eu mai multe despre toate aceste lucruri, dar e puțin dificil.

Având în vedere poziția dumneavoastră cu privire la stările de conștiință, în care susțineți că conștiința este în esență o iluzie creată de creier, ați putea detalia dacă percepeți o potențială legătură între iluzia conștiinței și iluzia sinelui? Cu alte cuvinte, cum ați descrie relația complexă, dacă există, dintre aceste două iluzii?

Susan Blackmore: Cred că cel mai bun răspuns ar fi că sunt extrem de apropiate. Iluzia conștiinței începe odată cu separarea sinelui de ceilalți. Asta ne duce la ceea ce numim „problema altor minți” – sentimentul că aceasta este conștiința mea și că e diferită de a ta, pentru că nu pot ști cum este pentru tine să vezi albastru, dacă albastrul tău este același cu albastrul meu și tot felul de întrebări de genul acesta.

Este aceasta întreaga iluzie? Nu. Ce mă interesează cu adevărat în privința iluziilor sunt capcanele mentale în care cădem atunci când încercăm să explicăm conștiința.

Dacă sunteți familiarizați cu lucrarea lui Dan Dennett, în special *Consciousness Explained*, veți vedea că el vorbește despre aceste capcane și continuă să le exploreze în cărți ulterioare, cum ar fi *Intuition Pumps*. Teoria sa, „a ciornelor multiple”, duce la o concluzie care multora li se pare bizară: nu există o distincție clară între procesele cerebrale care sunt conștiente și cele care nu sunt.

În mod tradițional, oamenii caută corelațiile neuronale ale conștiinței – adică încearcă să descopere ce procese din creier „generează” conștiința. Dar, conform teoriei lui Dennett, asta este o abordare fundamental greșită, pentru că nu există un moment sau o zonă din creier unde să „apară” conștiința.

Îmi place să pun această întrebare: „Ești conștient acum?”. Dacă cineva își pune întrebarea serios, poate avea un moment foarte ciudat. Probabil că se gândește: „Bineînțeles că sunt conștient! Dar... stați puțin... cu o clipă în urmă doar ascultam, fără să mă gândesc la asta. Acum ceva s-a schimbat. Am devenit mai conștient!”

Acest experiment mental demonstrează ceva fascinant: dacă nu știi sigur dacă ai fost conștient acum o clipă, atunci nimeni nu știe. Nu putem atașa un contor de conștiință pe creier ca să măsurăm „unde” era conștiința ta într-un anumit moment. Tot ceea ce facem este să reconstruim retrospectiv: „Da, eram conștient de asta și nu de aia”, dar acest proces este întotdeauna o atribuire a conștiinței, nu o măsurare a unui fenomen real.

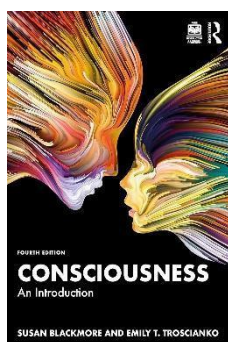
Acum, ca să răspund direct la întrebarea ta: sunt sinele și conștiința bazate pe aceeași iluzie? Într-un fel, da, pentru că „eu” sunt cel care este conștient, ceea ce deja presupune o separare între mine și această conștiință a mea și lucrurile de care sunt conștient.

Așa că, presupun că răspunsul meu este „Da” și „Nu”. Da, pentru că sunt strâns legate. Dar nu, pentru că pot fi separate în diferite niveluri ale iluziei.

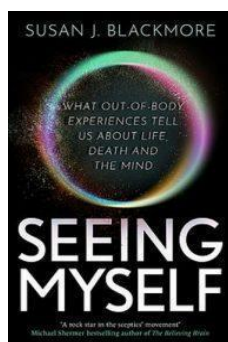
Cred că domeniul conștiinței este împotmolit în iluzii. Nu vom face niciun progres real până când nu vom

Înțelege cum apar aceste iluzii și nu vom învăța să vedem prin ele.

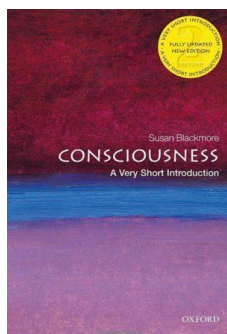
Cărți scrise de Susan Blackmore



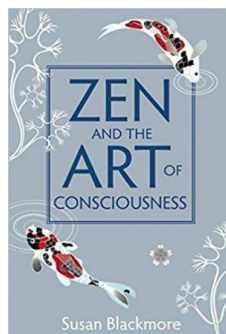
2024



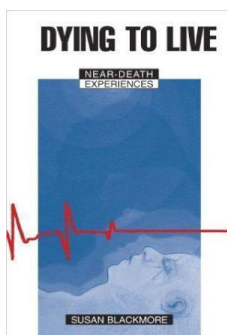
2020



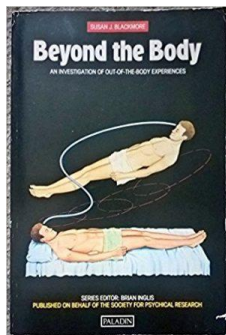
2017



2011



1993



1992

VIII. Dialogul cu Nicholas Humphrey



Nicholas Humphrey este profesor emerit de psihologie la London School of Economics, profesor invitat de filosofie la New College of the Humanities și Senior Fellow la Darwin College, Cambridge.

Este un psiholog teoretician, cunoscut la nivel internațional pentru activitatea sa privind evoluția inteligenței și a conștiinței umane. Interesele sale sunt variate: a studiat gorilele de munte cu Dian Fossey în Rwanda; a fost primul care a demonstrat existența „vederii oarbe” în urma leziunilor cerebrale la maimuțe; a propus celebra teorie a „funcției sociale a intelectului”.

Este un autor prolific și a publicat mai multe cărți: *Consciousness Regained* (1983), *The Inner Eye* (2003), *Seeing Red: A Study in Consciousness* (2009), *A History of the Mind* (2012), *The Mind Made Flesh* (2012) și *Soul Dust: the Magic of Consciousness* (2012).

A publicat recent noua sa carte *Sentience: The Invention of Consciousness* (2023).

Mai multe informații: <https://www.humphrey.org.uk>

În vara anului 1966, ați avut o întâlnire importantă cu Helen, o maimuță oarbă al cărei cortex vizual fusese îndepărtat pentru cercetări științifice. Ați putea aprofunda impactul pe care această întâlnire l-a avut asupra înțelegerii dumneavoastră despre conștiință? Mai exact, ce lecții sau idei ați tras din această experiență și cum a influențat ea perspectivele dumneavoastră ulterioare asupra naturii conștiinței?⁴⁹

Nicholas Humphrey: A fost o experiență extrem de atipică când am dat prima dată peste această maimuță. Eram student cercetător la Universitatea Cambridge și în laborator era o maimuță care fusese operată pentru a se descoperi funcția cortexului vizual. Un coleg de-al meu îndepărtase întregul cortex vizual din creierul acestei maimuțe și, deloc surprinzător, părea complet oarbă.

Am întâlnit-o pe Helen la un an după operație. Stătea acolo, uitându-se în gol, fără niciun interes să își folosească ochii. Dar am fost intrigat, pentru că, deși cortexul vizual fusese îndepărtat, o mare parte a sistemului vizual primitiv era încă intactă – coliculul superior din mijlocul creierului, aceeași zonă pe care o folosesc peștii, broaștele și alte vertebrate nemamifere pentru a vedea. Această parte a creierului le permite să

⁴⁹ Video cu Helen: youtube.com/watch?v=rDIswQHwt8.

prindă muște, să vâneze, să navigheze prin mediul lor – fără probleme.

Mi-am pus întrebarea: ar putea Helen să-și recapete vederea folosind acest sistem? Maimuțele normale nu-l utilizează, dar poate că era posibil să-l „reactivez”.

Când supervizorul meu plecase pentru o săptămână la o conferință în Zurich, am profitat de ocazie și am început să petrec timp cu Helen, încercând să interacționez cu ea în orice mod posibil. Spre surprinderea mea, în doar câteva ore am reușit să o fac să-și folosească ochii. Era clar că mă urmărea – dacă îi fluturam un fruct în față, întindea mâna să-l ia.

La sfârșitul săptămânii, deja atingea luminițe, urmărind obiecte pe care le mișcam la capătul unui băț. Eram uluit! Am trimis imediat o telegramă supervizorului meu: „Am învățat-o pe Helen să vadă. Nu o să vă vină să credeți”. Aveam 20 de ani, eram un simplu student, iar el era un profesor senior important. Nu a fost deloc încântat de mesajul meu.

Când s-a întors la Cambridge, l-am convins să vină să o vadă pe Helen. A trebuit să accepte că s-a întâmplat ceva extraordinar – Helen putea vedea! Mi-a permis să continui să lucrez cu ea, iar asta a dus la o relație de șapte ani în care am urmărit dezvoltarea ei.

După acești ani, Helen se comporta ca o maimuță cu vedere normală. Alerga prin cameră, ridica obiecte de

pe podea fără să se lovească de obstacole, își folosea ochii pentru a prinde muște în aer.

Dar ceva nu era în regulă. Ceva lipsea.

Când era relaxată, încrezătoare, putea vedea perfect. Dar dacă se simțea speriată sau amenințată, vederea ei dispărea. Se împiedica, părea din nou oarbă. Era ca și cum putea vedea doar atâta timp cât nu se gândea prea mult la asta.

Mi-am dat seama că era un tip extraordinar de vedere – o vedere în care maimuța însăși nu credea. Am scris o lucrare despre asta și am numit fenomenul „a vedea și a nu vedea nimic”, inspirându-mă din Sartre, pentru că lipsea ceva esențial.

Mai târziu, supervisorul meu a testat pacienți umani cu leziuni similare și a descoperit ceea ce azi numim „vedere oarbă”. Pacienții credeau că sunt orbi, dar puteau ghici cu precizie forme și poziții ale obiectelor din câmpul lor vizual. Cu toate acestea, ei spuneau: „Nu e nimic acolo”, „Nu înțeleg”, „Vederea mea nu are nimic de-a face cu mine”.

Acest fenomen m-a condus spre întrebarea cheie: dacă poți vedea fără să ai senzații vizuale, atunci la ce servește conștiința vizuală?

Dacă broaștele, peștii, oamenii și maimuțele pot folosi un sistem vizual fără a avea senzații conștiente, atunci de ce am evoluat pentru a avea experiențe vizuale?

Asta m-a urmărit toată viața mea științifică.

În ultimii 50 de ani, am încercat să înțeleg rolul funcțional al senzațiilor conștiente. Și asta m-a condus spre o altă întrebare majoră:

De ce experiențele vizuale au calități fenomenale atât de ciudate? Cum este să vezi roșu? Cum este să simți gustul zahărului? Cum este să auzi un țipăt sau sunetul clopotelor?

Nu primim doar informații despre obiecte, ci și informații despre cum este pentru noi să experimentăm acele obiecte.

Afirmați că fenomenul conștiinței este un produs al evoluției, deoarece conferă beneficii specifice supraviețuirii oamenilor...

Nicholas Humphrey: Eu nu spun că este un produs al evoluției pentru că ar avantaja supraviețuirea. Ceea ce întreb este: ce altceva poate fi, decât un produs al evoluției? Dacă acceptăm că este un rezultat al evoluției, atunci, cel puțin în acest caz, trebuie să aibă avantaje de

supraviețuire. Așadar, trebuie să descoperim care sunt aceste avantaje, mai exact.

Dar ați mai afirmat în trecut că este o formă de iluzie generată de creierul nostru. Această perspectivă caracterizează conștiința ca fiind o construcție mentală, o reprezentare internă a realității pe care o experimentăm. Ați putea detalia mai mult despre modul în care conceptualizați această iluzie și rolul ei în modelarea experiențelor noastre subiective? În plus, ce implicații are această perspectivă pentru înțelegerea naturii realității și a locului nostru în cadrul acesteia?

Nicholas Humphrey: Obișnuiam să susțin asta, da. Dar am citit mai multe lucrări și cărți care descriu conștiința ca fiind o iluzie, iar acest punct de vedere a devenit acum popular printre anumiți filosofi, precum Keith Frankish, Dan Dennett și alții, pe care îi numim „iluzioniști”.

Însă eu m-am retras de la această etichetă. Nu cred că este modul corect de a descrie conștiința.

A numi ceva o iluzie sugerează că este o greșeală, că suntem în eroare atunci când atribuim experienței calitățile pe care le trăim. Nu cred că este corect.

Când vedem roșu sau mirosim un trandafir, este exact așa cum pare a fi. Așa simțim despre el, așa este. A spune că această experiență este o iluzie înseamnă a subestima rolul pe care îl joacă în psihologia și fenomenologia noastră.

Conștiința fenomenală este o descriere veridică a ceea ce înseamnă pentru o ființă umană să trăiască aceste experiențe. Întrebarea este: „De ce o reprezentăm în acest fel?”

Senzațiile sunt reprezentări, iar acestea ar fi putut fi doar reprezentări ale faptelor, adică niște fapte fizice seci și goale. Dar nu sunt.

În opiniile dumneavoastră actuale, vă descrieți ca fiind „suprarealist” în ceea ce privește qualia sau conștiința. Ați putea detalia aspectele specifice ale suprarealismului pe care le considerați potrivite pentru a caracteriza perspectiva dumneavoastră asupra acestor fenomene? Cum surprinde această etichetă esența ideilor dumneavoastră despre qualia și conștiință și ce implicații are pentru înțelegerea noastră mai largă a minții și a experienței subiective?

Nicholas Humphrey: Căutam un alt termen care să nu fie „iluzionist”. Keith Frankish a scris o lucrare despre iluzionism, iar eu i-am răspuns spunând: „Nu, eu cred

că nu este o iluzie; nu este ireal; dacă este ceva, este super real.” Roșeața fenomenală este mai roșie decât roșul, durerea fenomenală este mai dureroasă decât durerea.

Și am preluat această formulare de la Pablo Picasso, care a fost unul dintre primii suprarealiști, deși el însuși nu a acceptat termenul. A creat o sculptură celebră a unei capre și a spus: „Capra mea este mai capră decât orice capră reală.”

Cu alte cuvinte, el încerca să exprime în arta sa esența obiectului pe care îl crea, susținând că aceasta depășea realitatea fizică.

Și cred că poate că acesta este modul corect de a vorbi și despre senzațiile fenomenale. Într-un sens, ele par să meargă mai adânc decât faptele superficiale ale cazului.

Postulați o relație strânsă între qualia și ceea ce numiți „sinele fenomenal”. Cum contribuie qualia la construcția sau experiența sinelui fenomenal și ce implicații are această legătură pentru înțelegerea experienței subiective?

Nicholas Humphrey: Cred că qualia, experiența fenomenală, este fundamentul sinelui. Este punctul de plecare pentru economia noastră psihologică și, de fapt, acesta este motivul pentru care a evoluat.

Ceea ce fac senzațiile este să dea o realitate substanțială sentimentului nostru despre cine suntem și despre existența noastră în lume.

Celebrul filosof David Hume spunea că, atunci când încercă să-și examineze propria minte și să descopere ce înseamnă să fie el însuși, nu găsește nimic altceva decât senzații. Și a fost dezamăgit de acest lucru, deoarece senzațiile sunt temporare, nu par să aibă o realitate continuă și nu oferă o bază substanțială pentru sine.

Eu cred că acest lucru este complet greșit. Cred că senzațiile ne oferă cea mai solidă bază pe care am putea-o avea pentru existența noastră în lume. Ele sunt dovezi mereu prezente ale modului în care ne trăim viața, ale importanței noastre, ale modului în care ne simțim noi înșine și, ceea ce este mai important, ale individualității noastre.

Pentru că, din câte știm, experiențele pe care le trăim nu seamănă cu nimic altceva în lume.

Sigur, continuăm să presupunem că și alți oameni au experiențe ca și noi, dar dovezi pentru asta nu există. Tot ce știm este că senzația mea de roșu seamănă cu senzația de roșu pe care am avut-o înainte. Pare să aibă ceva în comun cu celelalte senzații ale mele, dar totul este făcut în stilul meu. Acesta este semnul realității mele continue.

Îmi place să compar continuitatea senzațiilor cu continuitatea operelor de artă ale unui pictor. Așa cum toate picturile lui Cézanne sunt ale lui Cézanne, așa cum toate tablourile lui Vermeer sunt ale lui Vermeer, toate senzațiile mele sunt ale mele. Și acesta este fundamentul realității mele.

Pe măsură ce creștem, descoperim că acesta este centrul nostru psihologic și presupunem că și ceilalți oameni au un centru al sinelui echivalent, în care filmul senzațiilor este similar cu al nostru. Pentru fiecare dintre ei, acest „film” va fi individual, privat și important. Odată ce adoptăm această viziune asupra altor persoane, începem să schimbăm parametrii vieții sociale. Ajungem să trăim în ceea ce am numit „Societatea Eurilor”, a Eurilor Fenomenale. Și aceasta este, cu siguranță, baza culturii și civilizației umane.

Marea întrebare este: merge acest lucru dincolo de oameni? Se gândesc și alte animale la ele însele în acest fel?

Urmând această idee, susțineți că rolul conștiinței este de a oferi un sens vieții și existenței umane. Ați putea detalia acest aspect, adesea neglijat de cercetătorii și intelectualii care studiază fenomenul conștiinței?

Nicholas Humphrey: Aceasta este o întrebare importantă. Pentru a da un sens vieții, da, acesta este un punct de plecare important. În primul rând, conștiința ne oferă un sentiment al propriei noastre importanțe—nu suntem doar fenomene fizice, ci existăm pe un alt plan, cu o dimensiune spirituală. Într-un fel, trăim dincolo de limitele timpului și spațiului, într-un tărâm ce nu poate fi redus la simplele proprietăți materiale ale materiei.

Aceasta este o descoperire fundamentală despre natura umană și nu se rezumă doar la un simplu „da, eu contez”. Cultura și limbajul au rafinat această idee, ducând la convingerea că avem un suflet nemuritor.

S-ar putea să pară surprinzător ca un om de știință să ia în serios această perspectivă, dar eu o fac. Credința în suflet a fost una dintre forțele motrice ale istoriei umane și, în mare măsură, responsabilă pentru cele mai importante realizări ale omenirii.

Odată ce acceptăm ideea sufletului, atât al nostru, cât și al celorlalți, dobândim ambiții mai mari pentru ceea ce vrem să lăsăm în urmă și pentru ceea ce ne propunem să realizăm în viață.

Desigur, aceasta este o perspectivă specific umană. Nu cred, de exemplu, că un câine are conștiința unui suflet, deși și el se percepe ca un individ cu semnificație. Pentru animale, povestea este, probabil, diferită.

În cazul oamenilor, însă, conștiința a produs rezultate extraordinare și neașteptate. În noua mea carte, o descriu ca fiind bijuteria coroanei evoluției biologice— un fenomen pe care ar trebui să-l luăm foarte în serios.

În noua dvs. carte, *Sentience: The Invention of Consciousness* (Invenția conștiinței), susțineți că aceasta este limitată la mamifere și păsări datorită unei trăsături fiziologice foarte specifice: sângele cald. Este o idee inovatoare și surprinzătoare, care până acum nu a fost luată în considerare ca explicație plauzibilă pentru conștiință în dimensiunea sa senzorială. Ce v-a determinat să formulați această ipoteză?

Nicholas Humphrey: Nu este o afirmație științifică definitivă, dar cred că trebuie să acceptăm că sensibilitatea — conștiința fenomenală — este o dezvoltare relativ târzie în evoluție și nu se regăsește la formele de viață primitive.

Mulți cercetători, inclusiv prietenul și colegul meu Daniel Dennett, consideră că sensibilitatea există într-un continuum: unele animale ar avea mai multă, altele mai puțină. Eu nu sunt de acord. Cred că sensibilitatea apare ca un prag evolutiv — ea se manifestă doar la un anumit punct, deoarece implică un mecanism cerebral specific, bazat pe bucle de feedback. Și nu apare decât atunci când devine necesară.

Pentru un animal care nu își conștientizează propria existență, nu interacționează complex cu altele și nu trebuie să fie, cum spun eu, un „psiholog”, conștiința nu are utilitate.

Când am început să mă întreb „Când anume a apărut conștiința?”, mi-am dat seama că există o tranziție evolutivă pe care atât filosofii, cât și biologii au ignorat-o în mare parte: momentul în care vertebratele au devenit animale cu sânge cald.

Acum aproximativ 200 de milioane de ani, unii dinozauri au dobândit această trăsătură, iar mai târziu și păsările și mamiferele. Acest lucru a schimbat complet relația lor cu mediul: au devenit ființe autonome, independente de condițiile externe, capabile să fie active zi și noapte, să exploreze și să își regleze intern temperatura.

Această autonomie le-a oferit un nou sentiment de individualitate, de sine. Însă a avut loc și o altă schimbare: creierul lor s-a încălzit, iar viteza transmiterii impulsurilor nervoase a crescut semnificativ.

Dacă ridici temperatura unui creier cu 12-15 grade, ajungând la 37-38°C la mamifere și chiar la 40°C la păsări, viteza semnalelor neuronale se triplează. Această accelerare a proceselor cerebrale a permis o reorganizare a creierului, favorizând apariția unor mecanisme de feedback care altfel nu ar fi putut exista.

Astfel, cred că această combinație, un nou stil de viață, o autonomie mai mare față de mediu și un creier mai rapid, a condus la dezvoltarea unui nou mod de a percepe sinele. Acest proces a făcut posibilă apariția unei imagini psihologice mai sofisticate asupra propriei existențe, un element bazal fenomenului de conștiință, așa cum îl înțelegem astăzi.

Sunteți familiarizat cu opera filosoafei Susan Schneider, care a propus un test de conștiință pentru inteligența artificială, menit să determine dacă o mașină poate fi considerată conștientă. Credeți că aceste sugestii oferă o bază solidă pentru dezvoltarea unui astfel de test? Dacă natura a „inventat” conștiința umană, credeți că vom putea crea conștiința artificială?

Nicholas Humphrey: Ei bine, acestea sunt două întrebări diferite. Adică, Susan Schneider sugerează câteva teste pe care le-am putea folosi. Sunt unele destul de bune, destul de apropiate de genul de teste pe care le-aș sugera. Doar că ea nu ia în serios latura socială a conștiinței. Avem nevoie de această abilitate, deoarece conștiința ne permite să înțelegem mințile altor ființe asemenea nouă. Dacă ar fi inclus acest criteriu, testul ei s-ar fi apropiat mai mult de un instrument de diagnosticare a sensibilității într-o mașină.

În ceea ce privește inteligența artificială, nu există nicio dovadă că mașinile actuale sunt conștiente. Unii au speculat că modelele de limbaj avansate, cum ar fi ChatGPT, ar putea fi sensibile. De exemplu, David Chalmers a estimat, ceea ce mi se pare incredibil, că există o șansă de 10% ca ChatGPT să fie deja conștient. Însă această afirmație nu are sens: nu poți fi doar „10% conștient”. În plus, nu există niciun motiv să credem că aceste modele de limbaj posedă sensibilitate, deoarece nu au nici nevoia, nici mecanismele necesare pentru acest lucru și nu îndeplinesc criteriile de diagnosticare.

Desigur, asta nu înseamnă că nu vom putea, la un moment dat, construi o mașină care să îndeplinească aceste criterii. Dar acest lucru nu se va întâmpla accidental, doar prin creșterea capacității de procesare a datelor sau a nivelului de inteligență artificială. Inteligența nu este echivalentă cu sensibilitatea. Dacă vrem mașini conștiente, va trebui să le proiectăm deliberat în acest scop, dar deocamdată nu știm cum să facem asta.

În cartea mea, explorez tipul de mecanisme care ar trebui integrate într-o mașină pentru ca aceasta să devină sensibilă. Dar înainte de a încerca să creăm astfel de entități, ar trebui să ne întrebăm: de ce am face-o?

Un posibil motiv ar fi dorința de a conferi mașinilor capacitatea de a înțelege mințile altora, de a-și dezvolta

un sentiment al propriei importanțe psihologice, asemenea oamenilor. Dacă urmăm exemplul naturii, am putea proiecta mașini cu adevărat conștiente, poate chiar capabile să îndeplinească sarcini dincolo de limitele Pământului.

Oamenii nu vor putea niciodată să trăiască într-o galaxie îndepărtată, dar mașinile ar putea. Am putea crea entități care să exploreze spațiul în locul nostru. Însă, pentru a stabili un stil de viață autonom, pentru a dezvolta știință și cultură în medii noi, aceste mașini ar trebui să creadă în propria lor importanță existențială.

Așadar, întrebarea nu este doar dacă putem crea conștiință artificială, ci dacă dorim cu adevărat să o facem.

Susțineți că anumite animale sunt sensibile, în timp ce altele nu au deloc această capacitate. Știm că „sensibilitatea” este, în general, considerată o proprietate morală pentru ca un organism să aibă statut moral și să fie recunoscut ca posesor de drepturi. În acest context, ce implicații etice credeți că are cercetarea dumneavoastră, în special în raport cu etica și drepturile animalelor?

Nicholas Humphrey: Dacă ipoteza mea este corectă, ar trebui să fim precauți înainte de a atribui ideea de „sensibilitate” altor creaturi. Nu ar trebui să

presupunem automat că aceasta este prezentă doar pentru că un organism are un sistem nervos complex sau manifestă inteligență, o tendință destul de răspândită în rândul oamenilor.

Majoritatea dezbaterilor despre sensibilitate și drepturile animalelor se concentrează pe cât de inteligente sunt aceste ființe, nu pe cât de conștiente sunt în mod fenomenal. În cartea mea, abordez într-o anumită măsură această problemă, dar nu insist prea mult asupra criteriilor precise pentru acordarea statutului moral animalelor, pe baza unei conștiințe similare cu a noastră.

Dintr-o perspectivă științifică, nu neapărat din postura de cetățean, consider că majoritatea animalelor nu sunt conștiente. Nu cred că homarii sunt conștienți, nici caracatițele, cu atât mai puțin viermii. Aceasta este o concluzie care ne poate face să reevaluăm unele dintre legile recente privind protecția animalelor conștiente.

De exemplu, în Marea Britanie, anul trecut a fost adoptată o lege care recunoaște, prin definiție legală, homarii ca fiind sensibili. Drept urmare, este acum ilegal să fierbi un homar de viu. Există multe motive pentru care cineva ar putea considera această practică inacceptabilă, însă, în opinia mea, conștiința nu este unul dintre ele.

Când credeți că a apărut pentru prima dată conștiința la ființele umane? Credeți că se dezvoltă încă din perioada intrauterină, sau mai degrabă în prima copilărie? Un răspuns la această întrebare ar putea avea implicații profunde asupra înțelegerii dezvoltării umane și a drepturilor etice privind viața și bunăstarea.

Nicholas Humphrey: Nu cred că conștiința este prezentă de la început, iar un motiv pentru această afirmație este stadiul de dezvoltare al creierului la naștere. Când se nasc, copiii umani au un creier incomplet mielinizat, învelișul de mielină nu acoperă încă celulele nervoase, ceea ce înseamnă că multe dintre circuitele cerebrale nu sunt pe deplin funcționale. De exemplu, cortexul vizual al unui nou-născut nu este activ în mod normal.

Totuși, știm că bebelușii pot percepe vizual mediul înconjurător iar comportamentul lor ne indică acest lucru. Cu toate acestea, în acel stadiu, cred că nu au o experiență fenomenală propriu-zisă. Cu alte cuvinte, pot „vedea”, dar fără a avea conștiința deplină a ceea ce percep.

Majoritatea oamenilor nu concep posibilitatea de a percepe vizual fără o experiență conștientă a imaginii. Chiar și atunci când există cazuri clinice care sugerează acest lucru, ele nu sunt luate în considerare suficient de serios.

De exemplu, dacă un nou-născut reacționează la zâmbetul mamei sale, tindem să presupunem că percepe fața mamei în același mod în care o facem noi. Eu cred, însă, că acest lucru nu este adevărat. Estimez că abia în jurul vârstei de trei-patru luni, când mielinizarea este suficient de avansată, circuitele cerebrale devin capabile să susțină o experiență vizuală conștientă.

Neurologul portughez António Damásio, autor a numeroase lucrări pe acest subiect, introduce conceptul de sine stratificat, incluzând proto-sinele, eul nuclear și eul autobiografic, pe care îl consideră o formă de conștiință. Cum vă poziționați în raport cu acest cadru conceptual?

Nicholas Humphrey: Sunt de acord cu multe dintre ideile lui Damásio și am discutat despre ele în diverse contexte. Totuși, eu sunt mai specific decât el în ceea ce privește fundamentul sinelui, pe care îl consider bazat pe senzații. Toate celelalte forme ale sinelui sunt construite pornind de la această experiență senzorială de bază.

Subiectul stărilor noastre mentale este sinele, iar acesta se dezvoltă pe baza senzațiilor. Pe măsură ce evoluăm, sinele nostru devine mai complex, manifestându-se în forme precum eul social sau autobiografic. Totuși, consider că Damásio este prea grăbit să afirme că toate

aceste niveluri ale sinelui sunt echivalente și relevante în mod egal. Cred că evită întrebarea cheie: de ce ar trebui ca un anumit nivel al sinelui să fie mai important decât celelalte?

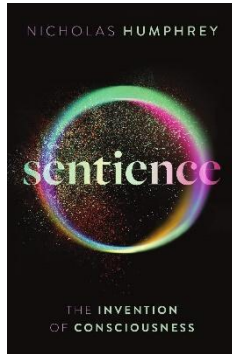
Din perspectiva mea, dacă pierzi sinele fenomenal, nucleul experienței conștiente, nu mai rămâne nimic. Nu poți păstra un eu autobiografic sau un eu social dacă pierzi însăși capacitatea de a avea experiențe conștiente.

Un aspect interesant este că, în aproape toate cazurile de disociere severă sau de perturbare a sinelui, indivizii par să-și păstreze totuși un nucleu fenomenal de bază. Acesta este centrat pe percepții fundamentale, precum durerea, culorile, lumina. Aceste aspecte nu dispar; dacă ar dispărea, persoana însăși ar înceta să existe ca subiect conștient.

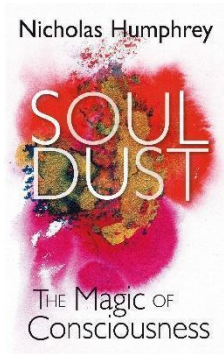
Există chiar cazuri clinice în care pacienții susțin că nu mai există deloc. De exemplu, în sindromul Cotard, pacienții insistă că sunt morți.

Dacă un medic le spune: „Nu ați murit, sunteți aici, vorbim chiar acum”, pacientul poate răspunde: „Nu, am murit. Eu nu mai exist.” Cred că ceea ce încearcă acești pacienți să exprime este, de fapt, faptul că și-au pierdut sinele fenomenal, fundația experienței conștiente.

Cărți scrise de Nicholas Humphrey



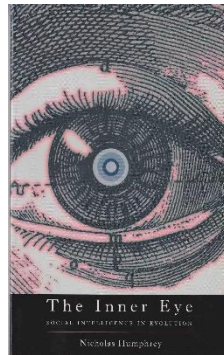
2023



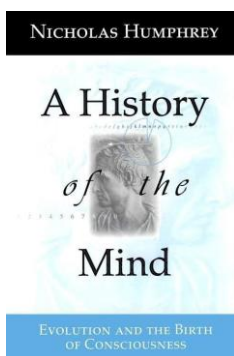
2011



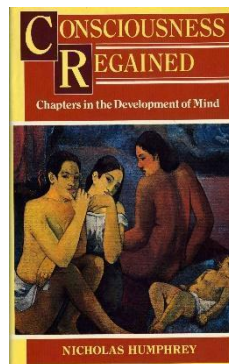
2006



2002



1998



1983

IX. Dialogul cu Sir Roger Penrose



Sir Roger Penrose este profesor emerit de matematică la Universitatea din Oxford și membru emerit al Wadham College, Oxford, precum și membru de onoare al St. John's College, Cambridge și University College London.

În 1994, a fost numit Cavaler pentru serviciile aduse științei de către Regina Elisabeta a II-a.

În 1969, împreună cu Stephen Hawking, Penrose a demonstrat că toată materia din interiorul unei găuri negre se prăbușește într-o singularitate.

A primit numeroase premii și distincții, inclusiv Premiul Wolf pentru Fizică din 1988 și Premiul Nobel pentru Fizică din 2020.

Este autorul mai multor cărți axate pe natura conștiinței, cum ar fi *The Emperor's New Mind* (1989), *Shadows of the Mind* (1994), *The Large, the Small and Human Mind* (1997), *The Road to Reality: Un ghid complet al legilor universului* (2007) sau *Cycles of Time: An Extraordinary New View of the Universe* (2013).

Mai multe informații:
maths.ox.ac.uk/people/roger.penrose

Sunteți unul dintre cei mai proeminenți oameni de știință ai secolului XXI și ați fost distins cu Premiul Nobel pentru Fizică în 2020 pentru demonstrarea faptului că găurile negre sunt o consecință inevitabilă a teoriei relativității generale a lui Einstein, contrazicând astfel scepticismul inițial al acestuia. Pe lângă această descoperire, v-ați concentrat și asupra naturii conștiinței. Când ați început să fiți interesat de acest subiect?

Roger Penrose: Pot să mă întorc la momentul în care eram doctorand, deoarece unul dintre ingredientele cheie ale modului în care mă gândesc la conștiință a apărut datorită unui șir de prelegeri la care am participat la Cambridge (cu mult timp în urmă, după cum vă puteți imagina). La acea vreme, studiam matematică pură, mai exact geometrie algebrică, dar am ales să urmez și trei cursuri care nu aveau legătură directă cu domeniul meu. Unul dintre ele era un curs despre relativitatea generală, predat de Hermann Bondi, care a avut o influență majoră asupra mea. Un altul era un curs de mecanică cuantică susținut de Paul Dirac. Cel de-al treilea era o prelegere susținută de S.W.P. Steen, un logician care predă logică matematică și teoria mașinilor Turing. Acolo am aflat pentru prima dată ce înseamnă, din punct de vedere tehnic, ideea de calculabilitate, dar și despre cele două teoreme ale lui Gödel, care m-au uimit.

Prima dată când am auzit despre teorema lui Gödel, ideea că în matematică există afirmații care nu pot fi demonstrate mi s-a părut profund deranjantă. Am participat la acest curs pentru a înțelege mai bine fenomenul și am realizat că teorema nu spune pur și simplu că există adevăruri nedemonstrabile, ci că, dacă urmezi un set de reguli formale, poți construi o propoziție matematică care afirmă: „Nu sunt demonstrabilă prin aceste reguli.” Apare atunci dilema: această afirmație este adevărată sau falsă? Dacă ar fi falsă, ar trebui să fie demonstrabilă prin reguli, dar asta ar însemna că regulile produc afirmații false, ceea ce le-ar compromite validitatea. Dacă este adevărată, înseamnă că nu poate fi demonstrată prin regulile respective, dar noi putem totuși să-i recunoaștem adevărul, chiar dacă sistemul formal nu o poate demonstra.

Această concluzie m-a fascinat: înțelegerea este mai puternică decât simpla respectare a regulilor. Pentru a realiza că o propoziție este adevărată, nu este suficient să urmezi reguli mecanice, trebuie să înțelegi sensul acestora. Acest proces de înțelegere implică conștiința. De aici am ajuns la ideea că gândirea conștientă transcende algoritmi și că mintea umană nu este un computer. Aceasta este concluzia la care am ajuns și pe care am încercat să o expun în cartea mea *The Emperor's New Mind*.

Scopul cărții a fost dublu: să explic această perspectivă asupra conștiinței și să ofer cititorilor o introducere accesibilă în subiecte precum matematica și fizica, într-un mod captivant și ușor de urmărit. Am crezut că, până la finalizarea cărții, voi înțelege suficient de multe despre neurofiziologie pentru a identifica ce procese cerebrale nu pot fi explicate prin calcul. Însă nu am reușit să găsesc un răspuns clar.

Totuși, am ajuns la o ipoteză: colapsul funcției de undă din mecanica cuantică ar putea juca un rol cheie. Dacă există un proces fizic care transcende calculul algoritmic, trebuie să-l căutăm în domeniul fizicii fundamentale. Aceasta este o provocare deschisă. Fizicienii sunt de acord că procesele din creier sunt guvernate de legile fizicii, dar întrebarea rămâne: care parte a fizicii este responsabilă pentru conștiință? Cred că răspunsul se află dincolo de fizica convențională.

Când am scris *The Emperor's New Mind*, nu știam exact care ar putea fi mecanismul fizic responsabil, iar la sfârșitul cărții am rămas cu această întrebare nerezolvată.

Putem considera că teza dumneavoastră despre conștiință este o consecință directă a teoremei lui Gödel? Având în vedere afirmația dumneavoastră că înțelegerea matematică este o caracteristică a

conștiinței și că aceasta nu este computațională, ceea ce înseamnă că nu poate fi redusă la un proces algoritmic, credeți că această perspectivă este fundamentată pe teorema lui Gödel?

Roger Penrose: Da, exact.

În acest sens, de ce credeți că mulți experți în inteligența artificială consideră că mintea funcționează ca un computer digital și că ingineria inversă a creierului uman este o activitate fezabilă? Ce motivează, în opinia dumneavoastră, aceste convingeri filosofice despre minte și creier, în special în rândul lor?

Roger Penrose: Cred că, pur și simplu, nu au acordat suficientă atenție argumentului meu. Există diverse obiecții care pot fi ridicate împotriva poziției mele, iar cea mai puternică dintre ele este următoarea: dacă nu cunoaștem algoritmi pe care îi folosim în creierul nostru, atunci nu putem aplica rezultatele lui Gödel pentru a demonstra că mintea umană transcende calculul algoritmic. Această critică este validă până la un punct.

În cartea mea *Shadows of the Mind*, am abordat această obiecție printr-un contra-argument. Chiar dacă am presupune că în creierul nostru există un algoritm, ceea ce, de altfel, nu știm dacă este adevărat, rămâne

Întrebarea: cum ar fi putut apărea un astfel de algoritm? Trebuie să presupunem că a evoluat prin selecție naturală, ca orice altceva, ceea ce înseamnă că ar trebui să fi conferit un avantaj selectiv.

Am ilustrat această idee printr-o imagine descrisă în carte: strămoșii noștri desfășurau activități necesare pentru supraviețuire, cum ar fi construirea de adăposturi, domesticirea animalelor sau cultivarea plantelor. În prim-plan, însă, se află cineva care are o idee despre cum să construiască o capcană pentru mamuți, o idee care oferă un avantaj clar. Dar, tot în prim-plan, se află și un matematician care lucrează la o teoremă, complet absorbit de calculele sale, în timp ce un tigrău cu dinți de sabie se apropie de el.

Morala acestui scenariu este că gândirea matematică avansată nu pare să fi oferit un avantaj selectiv evident. Și totuși, oamenii sunt capabili de raționamente matematice care depășesc orice nevoie evolutivă directă. Un exemplu fascinant este teorema lui Goodstein, pe care o descriu în *Shadows of the Mind*. Aceasta demonstrează că, urmând anumite reguli matematice simple, se ajunge inevitabil la un rezultat finit, deși procesul implică un număr uriaș de pași. Problema este că această teoremă nu are niciun fel de avantaj selectiv.

Cum știm că este adevărată? O știm datorită unor rezultate extraordinare dezvoltate de matematicianul

Georg Cantor. Dar ceea ce este cu adevărat remarcabil este că nu trebuie să fii Cantor pentru a înțelege această demonstrație. Este suficient să înțelegi raționamentul său, iar acest lucru nu presupune decât o ușoară schimbare de perspectivă.

Cum am fi putut evolua astfel încât să fim capabili să înțelegem astfel de adevăruri matematice, dacă am fi fost doar niște algoritmi? Nu există nicio posibilitate ca un algoritm atât de sofisticat să fi apărut prin selecție naturală, deoarece capacitatea de a înțelege astfel de concepte nu oferă un avantaj clar în lupta pentru supraviețuire.

Aici intervine argumentul meu: indiferent ce este această calitate generală a înțelegerii, ea trebuie să fie o caracteristică a conștiinței. Desigur, conștiința nu se limitează la gândirea matematică; implică și alte aspecte, cum ar fi percepția culorilor sau experiența durerii. Dar dacă în matematică vedem exemple clare de raționamente care transcend calculul algoritmic, atunci înțelegerea umană în sine trebuie să fie dincolo de computație.

Desigur, nu spun că teorema lui Goodstein este în sine „dincolo de calcul”, pentru că, odată ce ai datele de intrare potrivite, un computer ar putea verifica anumite versiuni ale acesteia. Însă din perspectiva selecției naturale, nu există absolut niciun avantaj evolutiv care să fi condus la apariția acestui tip de raționament. Cu

toate acestea, calitatea generală de a înțelege este necesară pentru existența noastră.

Ceea ce încerc să argumentez este că această capacitate generală de a înțelege nu este algoritmică. Nu mă refer aici la experiențele senzoriale, cum ar fi perceperea culorii albastre. Ideea fundamentală este că orice ar fi conștiința, ea nu poate fi redusă la o procedură computațională.

De ce nu sunteți de acord cu punctul de vedere predominant în neuroștiința actuală, conform căruia conștiința este o proprietate emergentă a creierului, rezultată din activitatea neuronală și schimbul de informații între neuroni?

Roger Penrose: Nu cred că există un „alt lucru” misterios care să plutească în capul nostru, dar în același timp nu consider că avem încă răspunsul complet. Sunt convins că explicația conștiinței ține de procesele neurofiziologice care au loc în creier, dar cred că pur și simplu nu înțelegem încă în totalitate ce anume se întâmplă.

În ceea ce privește natura conștiinței, credeți că doar neuroștiința poate rezolva această problemă sau susțineți necesitatea a ceea ce numiți "o nouă

fizică" pentru a integra acest fenomen unic în viziunea științifică asupra lumii?

Roger Penrose: Cred că avem nevoie de o nouă fizică. Iar celelalte părți ale argumentului sunt: "Care sunt celelalte părți ale fizicii pe care le cunoaștem?" Puteți trece în revistă una dintre teoriile majore pe care le avem, fizica newtoniană, și ați putea să o puneți pe un computer, așa cum fac oamenii. Dar teoria generală a relativității a lui Einstein?

Deși am reușit să simulăm spirala găurilor negre și semnalele pe care acestea le emit, complexitatea acestor fenomene ridică întrebări despre ce anume poate fi calculat în mod realist. Aș spune că relativitatea generală nu este complet computațională.

Cât despre mecanica cuantică, avem ecuația lui Schrödinger, care poate fi simulată pe un computer. Însă, în timp ce eram student, am asistat la un curs susținut de Paul Dirac, unde a explicat principiul superpoziției cuantice. Ideea era că un atom poate fi „aici” și „acolo” simultan. Pentru a ilustra acest lucru, Dirac a rupt o bucată de cretă în două și a încercat să explice cum bucățile ar putea exista în același timp în două locuri diferite.

Problema este că, în timpul acelei prelegeri, mintea mea a rătăcit; priveam pe fereastră, pierzând un fragment esențial din explicația sa. Când am revenit cu atenția, el deja trecuse la alt subiect, iar eu am rămas cu o

întrebare fără răspuns. Probabil încerca să ne liniștească, spunând că energia implicată în astfel de fenomene este prea mare pentru a fi relevantă la nivelul obiectelor macroscopice. Dar mie mi-a rămas sentimentul că avem nevoie de o nouă fizică.

Aici intervine problema actuală a fizicii deoarece ecuația lui Schrödinger nu descrie complet lumea fizică, iar acest lucru este bine cunoscut de toți fizicienii, deși adesea nu o spun deschis. Schrödinger însuși era conștient de această limitare, iar celebrul său experiment mental cu pisica a fost conceput tocmai pentru a sublinia absurditatea acestei situații.

Mulți interpretează greșit ceea ce încerca Schrödinger să demonstreze. Unii cred că el sugera că, dacă am avea un experiment suficient de sofisticat, am putea pune o pisică într-o stare de superpoziție, unde ar fi vie și moartă în același timp, dar că nu am ajuns încă la acest nivel tehnologic. Nu acesta era mesajul său. De fapt, Schrödinger argumenta împotriva propriei ecuații, arătând că aceasta nu poate explica în mod adecvat realitatea. El demonstra că este ridicol să presupunem că o pisică poate fi vie și moartă simultan și că, prin urmare, mecanica cuantică așa cum o înțelegem astăzi este incompletă.

Așadar, dacă dorim să înțelegem conștiința, trebuie să căutăm în fizică ceva care transcende modelele actuale. De aceea, consider că explicația conștiinței nu poate

veni exclusiv din neuroștiință, ci necesită o extindere a fizicii fundamentale.

Ați putea să dezvoltați ideea că ar putea exista un hiatus sau o lacună în mecanica cuantică, sugerând că lipsește ceva pentru ca aceasta să fie considerată o teorie completă? În ce fel vă imaginați această incompletitudine și ce legătură ar putea avea cu înțelegerea noastră despre conștiință?

Roger Penrose: Da, există un decalaj uriaș. Însă acest decalaj este, în esență, unul de natură fizică, independent de problema conștiinței. Ceea ce vreau să spun este că, în acest spațiu necunoscut al mecanicii cuantice, s-ar putea afla și non-computabilitatea, iar aceasta ar putea avea implicații fundamentale asupra conștiinței. Când am scris *The Emperor's New Mind*, am încercat să argumentez această idee. Slăbiciunea cărții a fost că, la acel moment, nu aveam nicio idee despre cum ar putea neuronii să joace un rol în acest proces.

Am sperat că, până la finalul cărții, voi înțelege suficient despre neurofiziologie pentru a identifica un mecanism posibil, dar acest lucru nu s-a întâmplat. Am fost nevoit să finalizez cartea fără un răspuns clar, ceea ce m-a nemulțumit într-o anumită măsură.

Apoi, Stuart Hameroff, care mi-a citit cartea, mi-a scris pentru a-mi vorbi despre anumite structuri despre care

nu știam la acea vreme: *microtubulii*. El dezvoltase o teorie conform căreia acești microtubuli ar putea fi implicați în conștiință, iar în prezent această idee a câștigat susținere din mai multe perspective. Un argument în favoarea acestei ipoteze este legătura dintre *microtubuli și efectul anestezicelor generale asupra conștiinței*.

Hameroff, fiind anestezist, nu doar că induce și inversează starea de inconștiență la pacienți, dar și-a pus problema ce anume dezactivează conștiința într-un mod controlat și reversibil. Dacă putem aduce un pacient într-o stare de comă indusă și apoi îl putem trezi complet, înseamnă că există un mecanism fiziologic precis care controlează acest proces.

Interesant este că efectele anestezicelor generale nu par a fi pur chimice, așa cum s-a presupus inițial, ci mai degrabă legate de un proces fizic fundamental. Ce se întâmplă la nivel fizic în momentul în care conștiința este „oprită” și „reactivată”?

În anii 1990, această ipoteză era considerată marginală. Totuși, în prezent, ea a devenit una dintre cele câteva teorii luate în serios despre natura conștiinței. Este încurajator să văd că această idee primește acum atenția cuvenită și că există experimente tot mai detaliate care investighează efectul anestezicelor asupra structurilor cerebrale implicate.

Microtubulii ar putea fi un ingredient în această poveste a conștiinței. Direcția în care evoluează cercetarea ar putea, în cele din urmă, să confirme sau să infirme această ipoteză. Cert este că, în acest moment, suntem într-un punct în care aceste idei încep să fie testate riguros, ceea ce este un pas important în descifrarea naturii conștiinței.

În esență, criticați interpretarea de la Copenhaga a mecanicii cuantice, deoarece susțineți că ceea ce se întâmplă în realitate este contrar acestei interpretări: avem colapsul funcției de undă, dar acesta nu necesită prezența unui observator conștient. În acest context, cum poate apărea conștiința din procesul de colaps al funcției de undă, prin influența gravitației, așa cum susțineți?

Roger Penrose: Exact, așa cum ați spus, este invers față de ceea ce propune interpretarea de la Copenhaga. Sunt mai multe aspecte care trebuie luate în considerare. Unul dintre ele, pe care nu l-am menționat încă, este faptul că, în opinia mea, colapsul funcției de undă este rezultatul interacțiunii dintre gravitație și mecanica cuantică. Această ipoteză nu a fost încă confirmată sau infirmată experimental, dar există experimente în desfășurare care vizează testarea ei. Există, de asemenea, argumente teoretice care susțin această viziune.

Am propus în urmă cu câteva decenii un argument conform căruia există o incompatibilitate fundamentală între principiile mecanicii cuantice și cele ale relativității generale. Pe de o parte, mecanica cuantică se bazează pe principiul superpoziției, conform căruia un sistem cuantic poate exista simultan în mai multe stări. Pe de altă parte, relativitatea generală are la bază principiul echivalenței, formulat de Galileo, conform căruia efectele gravitației pot fi anulate prin cădere liberă.

Galileo și-a imaginat un experiment în care două obiecte de mase diferite, lăsate să cadă de pe Turnul din Pisa în absența atmosferei, ar atinge solul în același timp. Dacă două insecte aflate pe unul dintre aceste obiecte s-ar uita una la cealaltă în timpul căderii, ele nu ar percepe existența gravitației. Acest principiu al căderii libere este fundamentul relativității generale a lui Einstein, dar este incompatibil cu principiul superpoziției din mecanica cuantică.

Argumentul meu este că acest conflict necesită o nouă teorie, o teorie care încă nu a fost formulată. Tot ceea ce pot face este să estimez la ce scară s-ar putea observa efectiv colapsul funcției de undă ca un proces fizic măsurabil.

După cum ați menționat, această perspectivă este opusă celei susținute în primele zile ale mecanicii cuantice, când se credea că colapsul funcției de undă, care are loc în momentul unei măsurători, depinde de

prezența unui observator conștient. Această idee era promovată de oameni precum John von Neumann și Eugene Wigner, care considerau că procesul de măsurare și, implicit, colapsul funcției de undă, sunt legate de actul conștient al observației.

Am discutat personal pe acest subiect cu Wigner și am descoperit că nu era atât de dogmatic pe cât s-ar putea crede. Deși lua în serios ipoteza conform căreia conștiința joacă un rol în colapsul funcției de undă, era deschis și la alte explicații. De fapt, era conștient că această perspectivă nu era complet satisfăcătoare, întrucât ar însemna că universul însuși ar avea nevoie de un observator conștient pentru a funcționa, ceea ce este improbabil.

Eu cred că colapsul funcției de undă este un proces fizic obiectiv, independent de prezența unui observator conștient. Însă, în sens invers, creierul exploatează acest proces fizic pentru a genera conștiința.

Cu siguranță mai avem un drum lung de parcurs. Nu văd încă o legătură directă între aceste fenomene. Deși sunt impresionat de progresul realizat până acum, mai ales că nu mă așteptam să se avanseze atât de mult, suntem încă departe de a înțelege pe deplin cum se leagă aceste descoperiri de fenomenul de conștiință în sine.

Cât de aproape credeți că suntem de a stabili o legătură autentică între stările de conștiință fenomenală, cum ar fi emoțiile sau sentimentele, și reducerea fiziologică a stărilor de superpoziție cuantică în creier, în special în microtubuli?

Am participat recent la o conferință în Canada, unde au avut loc discuții serioase și s-au desfășurat experimente legate de aceste idei. Am fost surprins de cât de mult s-a progresat în înțelegerea microtubulilor, a proteinelor tubulină și a propagării semnalelor nervoase. Cercetările dezvăluie aspecte necunoscute anterior, ceea ce sugerează că avem încă multe lucruri interesante și neașteptate de descoperit.

Știu că nu face parte direct din întrebarea dumneavoastră, dar există un fenomen intrigant legat de viteza reacțiilor umane. De exemplu, în jocuri precum ping-pong sau tenis, unde trebuie să reacționezi extrem de rapid, pare că deciziile conștiente influențează mișcările mai repede decât ar trebui să fie posibil, având în vedere viteza transmiterii semnalelor nervoase și zonele cerebrale implicate. Din perspectiva fiziologiei clasice, aceste reacții nu ar trebui să fie posibile într-un timp atât de scurt, ceea ce ridică întrebări foarte interesante despre percepția timpului și procesarea informației în creier.

Reducerea stării cuantice este deja un fenomen enigmatic, mai ales când încercăm să îl corelăm cu

principiile relativității generale și ale mecanicii cuantice. Când încercăm să le combinăm, apar relații surprinzătoare privind modul în care timpul progresa. Aceste relații nu sunt inconsecvente, dar sunt extrem de bizare și sugerează că percepția noastră asupra timpului nu reflectă cu exactitate realitatea temporală a proceselor fizice.

Au existat câteva experimente timpurii pe această temă, pe care le-am descris în *The Emperor's New Mind*, inclusiv cele realizate de Benjamin Libet. Într-unul dintre ele, Libet a efectuat o operație pe creierul unui pacient (nu voi intra în detalii), iar rezultatele au arătat un conflict temporal greu de explicat: ordinea evenimentelor percepute nu corespundea ordinii reale în care acestea s-au întâmplat.

Acest lucru sugerează că există ceva bizar modului în care trăim timpul. Când credem că trăim un eveniment și când îl trăim de fapt pot fi două momente diferite, iar pentru a înțelege pe deplin acest fenomen va fi nevoie de o nouă abordare teoretică.

În concluzie, în ceea ce privește fenomenul spiritualității, numeroase cărți explorează teme precum sincronicitatea cuantică, chakrele, chi și alte concepte controversate. Dacă ar fi să oferiți un scurt comentariu, ce părere aveți despre aceste

abordări și despre popularitatea unor astfel de cărți? Credeți că există elemente care ar putea conferi o oarecare validitate acestor fenomene alternative?

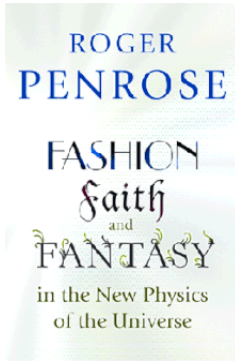
Roger Penrose: Nu sunt adeptul niciunuia dintre aceste concepte. Trebuie să recunosc că nu am citit prea mult în această direcție, dar, din prejudecata mea, nu cred că există prea multe lucruri valoroase în această zonă. Desigur, s-ar putea ca unele idei să ofere o perspectivă nouă asupra conștiinței, dar eu m-am concentrat doar pe un aspect foarte specific al acesteia: capacitatea de înțelegere, în special în ceea ce privește raționamentul matematic.

Conștiința este, fără îndoială, un fenomen mult mai vast și mai complex decât domeniul limitat pe care l-am explorat eu. Există numeroase alte aspecte ale experienței conștiente care merită studiate. Totuși, am ales să stau departe de aceste concepte ezoterice, în principal pentru că nu văd o modalitate clară de a le aborda într-un mod riguros și precis.

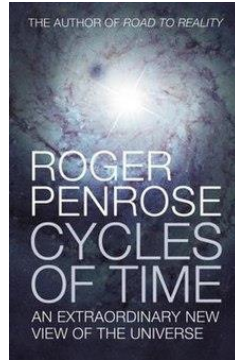
Cred că aceasta este principala problemă. Sunt matematician, iar ceea ce îmi place este să pot construi teorii bine definite, care să permită deducerea unor consecințe precise și verificabile. Fenomenele despre care vorbiți nu se prezintă în această formă, ceea ce face dificilă studierea lor într-un cadru științific riguros.

Așadar, nu le-am investigat și nu văd un beneficiu direct pentru mine în a le studia. Totuși, asta nu înseamnă că ele nu ar putea avea o anumită valoare pentru alții.

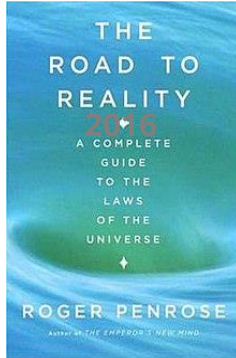
Cărți scrise de Sir Roger Penrose



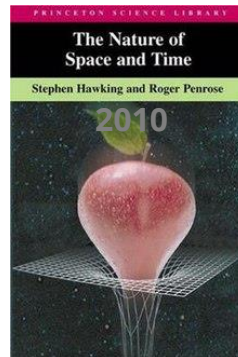
2016



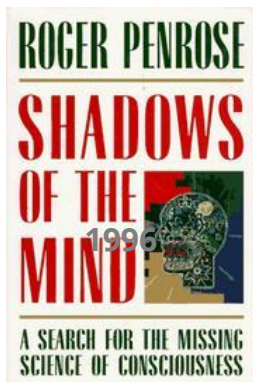
2010



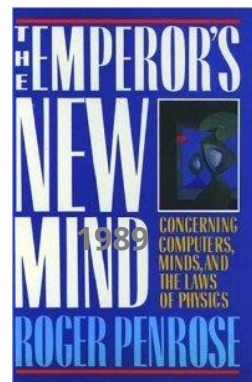
2004



1996



1994



1989

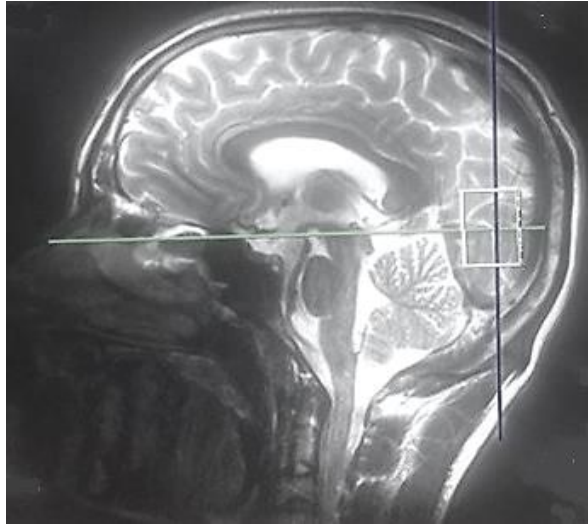
PARTEA II

Creierul

PARTEA II

Creierul

I. Creierul: o introducere



Imagine fMRI a creierului

Natura filosofică și științifică a creierului a fost un subiect de dezbatere timp de secole, fiind explorată în diverse discipline. Astăzi, avem mult mai multe cunoștințe despre creier decât aveau generațiile anterioare. De exemplu, știm că acesta cântărește, în medie, 1,5 kg și este alcătuit din numeroase tipuri de neuroni, care comunică prin impulsuri electrochimice. Acești neuroni formează rețele complexe, fiecare având un rol distinct în funcții mentale precum memoria pe termen lung, emoțiile și creativitatea.

Creierul uman conține aproximativ 86 de miliarde de neuroni, interconectați prin trilioane de sinapse. Este organizat în regiuni specializate. Cele mai primitive structuri, reminiscențe ale etapelor timpurii ale evoluției, controlează instinctele de supraviețuire. Deasupra acestora se află sistemul limbic, responsabil pentru emoții și procesarea memoriei. La exterior se află cortexul cerebral, stratul superior al creierului, împărțit în lobi specializați care coordonează procese cognitive complexe.

Deși cunoștințele despre creier au avansat semnificativ, încă există numeroase necunoscute, așa cum va fi explorat în această a doua parte a cărții. De-a lungul istoriei, creierul a fost subiect de controversă în încercarea de a înțelege natura umană. Doi gânditori proeminenți ai Greciei Antice au oferit perspective diferite asupra relației dintre minte și creier.

Într-o dezbatere aprinsă, Aristotel, în cartea a III-a a lucrării *Despre părțile animalelor*,⁵⁰ a susținut că inima este sediul senzațiilor, conștiinței și intelectului, iar creierul ar avea doar rolul de a răci sângele. Împotriva acestei idei, Hipocrate, în *Despre boala sacră*⁵¹, a argumentat că creierul, nu inima, este sursa plăcerii, bucuriei, tristeții și durerii.

⁵⁰ Versiune online: penelope.uchicago.edu/aristotle/parts3.html

⁵¹ Versiune online: classics.mit.edu/Hippocrates/sacred.html

Deși Grecia Antică este adesea considerată o eră deosebită privind istoria înțelegerii creierului, unele dovezi sugerează că primele reflecții asupra acestuia sunt mult mai vechi. La sfârșitul epocii de piatră, o perioadă care a început acum aproximativ 50.000 de ani în Europa și a evoluat în neolitic (între 3000 și 2000 î.Hr.), oamenii au dezvoltat agricultura, au domesticit animale și au construit primele așezări.

Tot în această perioadă au fost descoperite dovezi ale unei practici chirurgicale remarcabile. Fosile umane din acea vreme prezintă urme de perforații craniene, realizate cu unelte din piatră. Aceste descoperiri sugerează că oamenii neolitici efectuau o intervenție chirurgicală cunoscută sub numele de trepanație (din grecescul *trypanon*, care înseamnă „a străpunge”). Aceasta este cea mai veche procedură chirurgicală cunoscută, realizată probabil în scopuri medicale sau rituale.⁵²

Primul craniu recunoscut ca fiind perforat a fost descoperit accidental în 1864 din cauza... fecalelor de păsări. Care ar fi legătura dintre crani și fecale? În 1861, a izbucnit Războiul Civil American și a devenit necesară

⁵² Weber, J. e Wahl, J. (2006) "Neurosurgical aspects of trepanations from Neolithic times", *International Journal of Osteoarchaeology*, 16: 536-545.

garantarea fertilizării producției agricole, devenind una dintre prioritățile lui Abraham Lincoln.

Interesant este că cele mai bune îngrășăminte din lume, la acea vreme, veneau din America de Sud, odată cu extracția unei substanțe numite „guano” care provine din fecalele diferitelor animale, fiind bogată în fosfor și azot. Astfel, Ephraim G. Squier, arheolog și jurnalist, este trimis de Lincoln în America de Sud – mai exact, în Peru – pentru a garanta producția și expedierea de îngrășăminte în Statele Unite.

După ce și-a îndeplinit sarcina, Squier a decis să exploreze țara, ajungând în cimitirul incaș din Valea Yucay. Acolo, a descoperit un craniu perforat, având o gaură dreptunghiulară de 15x17 mm. Această descoperire l-a uimit, deoarece unghiurile drepte sunt extrem de rare în structurile naturale.

Conștient de importanța descoperirii, Squier s-a întors în Statele Unite și a prezentat craniul Academiei de Medicină din New York în 1865. Însă, din păcate, descoperirea sa nu a fost luată în serios. Mulți dintre oamenii de știință din acea vreme respingeau ideea că popoarele indigene ar fi putut poseda cunoștințe chirurgicale avansate, o atitudine influențată de prejudecăți rasiale.

Nedescurajat, Squier și-a continuat cercetarea și a decis să trimită craniul la Paris, unde avea să fie analizat de Paul Broca, fondatorul primei Societăți Antropologice

din Franța și una dintre cele mai mari autorități științifice în studiul creierului la acea vreme.

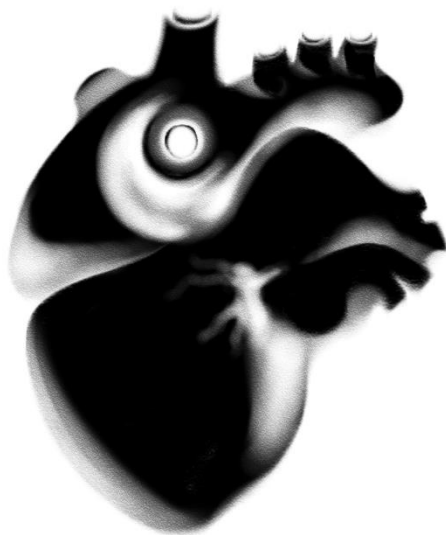
Broca a fost profund surprins când a examinat craniul în detaliu. Perforația prezenta semne clare că fusese realizată cu o unealtă de tăiere specializată, indicând existența unei forme de protochirurgie considerată până atunci imposibilă. Aceasta a fost prima dintr-o serie de descoperiri asemănătoare. În următoarele decenii, au fost găsite numeroase alte cranii perforate, multe dintre ele prezentând un detaliu și mai uimitor: semne de creștere osoasă în jurul perforației.

Ce înseamnă acest lucru? Motivul simplu a fost că indivizii supuși acestor intervenții chirurgicale au supraviețuit procedurii, uneori luni sau chiar ani după operație. Această constatare a revoluționat înțelegerea noastră asupra cunoștințelor medicale ale popoarelor preistorice.

O altă sursă istorică despre tratamentul craniului poate fi găsită în Egiptul Antic, unde ritualurile funerare și arhitectura piramidelor reflectă importanța pe care egiptenii o acordau vieții de apoi. Conform credințelor lor, pentru ca sufletul să ajungă în locul potrivit, era necesar ca trupul să fie îmbălsămat și conservat, o convingere care a dus la elaborarea procesului de mumificare.

Pentru egipteni, cea mai importantă parte a corpului era inima, considerată sediul identității, gândirii și sentimentelor. De aceea, în timpul mumificării, intestinele, plămâni, ficatul și stomacul erau îmbălsămate și păstrate în borcane canopice, alături de trupul mumificat.

În contrast, creierul era considerat lipsit de valoare. În timpul procesului de mumificare, acesta era scos prin nări cu un cârlig de fier și aruncat, ceea ce sugerează că egiptenii nu asociau mintea sau conștiința cu creierul.



Imaginea 14: "Pleasure" © Gaia Amoroso

Textele egiptene, în special inscripțiile funerare, confirmă această viziune. Inima nu era doar un depozitar al acțiunilor sufletului, ci era considerată și sediul rațiunii și al emoțiilor. Această concepție a

persistat chiar și în Biblie. Când Vechiul Testament a fost tradus în greacă, traducătorii au menținut ideea că intelectul și emoțiile umane rezidă în inimă, fără nicio referire la creier în întregul corp de texte scriptice.⁵³

Totuși, în ciuda acestei neglijări a creierului, egiptenii antici au lăsat cea mai veche referință scrisă despre acest organ. Un papirus medical, cunoscut sub numele de Papirusul Edwin Smith, descrie 48 de tipuri de leziuni ale capului și gâtului, oferind recomandări pentru tratament și intervenții chirurgicale. Acest text conține referințe datate până în anul 3000 î.Hr., demonstrând că egiptenii antici, chiar dacă nu îi atribuiau creierului un rol central în gândire, aveau deja o înțelegere medicală avansată asupra acestuia.

Autorii papirusului par să fi avut o anumită înțelegere a funcției creierului. Ei au descris în detaliu leziuni specifice ale capului și le-au corelat cu diverse simptome, precum paralizia sau pierderea vorbirii. Totuși, au comis și erori semnificative, cum ar fi afirmarea că o leziune în emisfera dreaptă ar afecta partea dreaptă a corpului, când, de fapt, partea stângă este cea care ar trebui să fie afectată.

⁵³ York, G.K. e Steinberg, D.A. (2010) "Neurology in Ancient Egypt" In Finger, S., Boller, F. e Tyler, L. (eds) *Handbook of Clinical Neurology*, vol. 95, Elsevier: Amsterdam.

O altă referință istorică fascinantă despre creier se regăsește în epopoile homerice, *Iliada* și *Odissea*, care au inspirat și titlul acestei cărți. În concepția grecilor antici, eul era alcătuit din mai multe forțe, una dintre ele fiind *psyché*, energia vitală care menținea ființa în viață. Homer descria existența mai multor „forme” ale sufletului, situate în piept, nu în creier.

Cea mai importantă dintre aceste forțe era *thymos*, menționat de peste 450 de ori în *Iliada*. Acesta era considerat sursa emoțiilor și a motivației și era situat în diafragmă. Capacitățile intelectuale, pe de altă parte, erau asociate cu *noos*, un alt aspect al sufletului, despre care se credea că se află tot în piept. La fel ca egiptenii, textele homerice nu atribuiau o mare importanță creierului.

Totuși, Homer a avut o influență lingvistică remarcabilă, introducând în vocabular trei concepte fundamentale: (1) *enkephalos*, care desemnează creierul propriu-zis; (2) *muelos*, referindu-se la coloana vertebrală (de unde derivă termenul „măduvă”); și (3) „*tendoane*”, un termen care a stat la baza conceptului de neuron.⁵⁴

Chiar dacă aceste tradiții plasau conștiința și emoțiile în inimă sau în piept, în Grecia Antică au existat gânditori

⁵⁴ Singer, C. (1957) *A Short History of Anatomy and Physiology from the Greeks to Harvey*, Dover: New York.

care au contestat această viziune și au susținut rolul central al creierului.

Unul dintre aceștia a fost Alcmaeon din Croton, care a susținut că creierul este organul senzațiilor. Deși toate textele sale originale s-au pierdut, avem referințe indirecte despre ideile sale. Theophrastus din Eressos, de exemplu, afirma că Alcmaeon a fost primul care a studiat anatomia animalelor prin disecție. Descoperirile sale includeau identificarea a două canale care conectau partea din spate a ochiului cu creierul, ceea ce astăzi numim nervii optici. Această constatare l-a condus la ipoteza că toate simțurile sunt conectate la creier.

O altă referință indirectă provine de la Aetius, contemporan al lui Alcmaeon, care menționa că acesta considera creierul centrul inteligenței și al proceselor cognitive. Datorită acestor descoperiri, impactul ideilor lui Alcmaeon poate fi comparat, într-o anumită măsură, cu cel al lui Copernic sau Darwin în domeniul lor.⁵⁵

Cealaltă mare referință greacă privind importanța creierului provine de la Hipocrate, părintele medicinei. În *Corpus Hippocraticum*, există numeroase referiri la

⁵⁵ Mithen, S. (1996) *The Prehistory of the Mind*, Phoenix Books: Guernsey.

creier, dar cea mai importantă lucrare este *Despre boala sacră*, un text axat pe epilepsie.

Hipocrate a fost primul care a încercat să demonstreze că epilepsia nu putea fi explicată prin teoria medicală predominantă a epocii, cunoscută sub numele de "Teoria posesiunii demonice". Conform acestei teorii, epilepsia era văzută ca un semn al unei posesiuni demonice, o pedeapsă divină pentru păcatele comise de individ.

Hipocrate a respins această explicație, susținând că era doar un mit perpetuat de șarlatani și vindecători care nu căutau cu adevărat un tratament. El a argumentat că epilepsia nu avea o cauză supranaturală, ci era o afecțiune a creierului, provocată de un exces de flegmă, care bloca fluxul de aer în vasele de sânge. Conform acestei ipoteze, o criză epileptică era mecanismul natural al corpului pentru a elimina această obstrucție.

Pe lângă această explicație naturalistă, Hipocrate mai susținut că întreaga activitate mentală are loc în creier, nu în inimă. Merită să reflectăm asupra cuvintelor sale, care marchează un punct de cotitură în istoria medicinei:

"Ar trebui să se știe în general că sursa plăcerii, a veseliei, a râsului și a amuzamentului nostru, ca și a durerii, a neliniștii și a lacrimilor, nu este alta decât creierul. Acesta este în mod special organul care ne permite să gândim, să vedem și să auzim

și să distingem urâtul și frumosul, răul și binele, plăcutul și neplăcutul. [...] Tot creierul este și sediul nebuniei și al delirului, al fricilor și spaimelor care ne asaltează, adesea noaptea, dar uneori chiar și ziua; acolo se află cauza insomniilor și a somnambulismului, a îndatoririlor uitate și a excentricităților...".⁵⁶

Această descriere remarcabilă, realizată de Hipocrate în jurul anului 400 î.Hr., ar putea fi extrasă din orice manual modern de neuroștiințe. Ea subliniază rolul creierului în percepția senzorială, raționament, emoții și îl asociază cu tulburările mentale.



Imaginea 15: „Infinitesimale” © Gaia Amoroso

⁵⁶ Chadwick, J. e Mann, N. (eds.) (1983) *Hippocratic Writings*, Penguin: London.

În aceeași lucrare, Hipocrate demonstrează abilități anatomice remarcabile, oferind descrieri detaliate ale creierului uman. El observă similarități între creierul uman și cel al altor animale și menționează existența unei structuri care îl împarte în două jumătăți — o referire clară la corpul calos și cele două emisfere cerebrale.

Platon a încercat, de asemenea, să înțeleagă rolul creierului, asociindu-l cu conceptul de suflet. În filosofia sa, sufletul nu avea un sens religios, ci era împărțit în trei componente distincte:

- *Epithymetikon* – situat în ficat și intestin, responsabil pentru nevoile vegetative și instinctele de bază.
- *Thymos* – situat în inimă, considerat sursa emoțiilor, inclusiv furia, frica, mândria și curajul.
- *Logistikon* – partea cea mai nobilă a sufletului, legată de rațiune și inteligență, considerată nemuritoare și capabilă de reîncarnare.

Cea mai interesantă parte a acestei clasificări este faptul că Platon a plasat *logistikon*-ul în creier. De exemplu, în *Phaedo*, el afirmă: "Creierul este sursa percepțiilor auditive, vizuale și olfactive, iar memoria și opinia își au originea în el". În *Timaeus*, el susține că

"capul este partea cea mai divină a corpului și domină restul acestuia".⁵⁷

Împotriva acestei perspective platonice, se ridică Aristotel, ale cărui idei despre creier au fost menționate anterior. Gândirea aristotelică nu a acordat o importanță majoră acestui organ. Interesant este că Aristotel a ajuns la această concluzie pe baza observațiilor sale empirice, care, deși rudimentare, au fost interpretate greșit.

Fascinat de funcționarea corpului uman, Aristotel a fost printre primii care au disecat un număr mare de animale, inclusiv pești, reptile, mamifere și chiar elefanți. A realizat numeroase descrieri anatomice, observând două membrane care acoperă creierul, despre care știm astăzi că sunt meningele (dura mater și pia mater). De asemenea, a identificat o structură situată în partea posterioară a creierului, pe care a numit-o *parencephalis*, o referință clară la cerebel, pe care l-a recunoscut ca având o textură și un aspect diferit de restul creierului.

Cu toate acestea, Aristotel nu a considerat creierul un organ important. Era nemulțumit de structura sa uniformă și rece, preferând să acorde prioritate inimii, care era caldă și dinamică. Această preferință se baza

⁵⁷ Crivellato, E. e Ribatti, D. (2007) "Soul, mind, brain: Greek philosophy and the birth of neuroscience", *Brain Research Bulletin*, 71: 327-336.

pe una dintre convingerile fundamentale ale epocii: căldura era necesară pentru viață, deoarece corpurile vii erau calde, iar cadavrele erau reci.

Aristotel a studiat dezvoltarea embrionară și a observat că primul organ care apare este inima. Acest lucru l-a condus la concluzia că rolul principal al creierului este acela de a răci sângele. O altă observație empirică care i-a consolidat această convingere era faptul că "căldura se ridică". Văzând numeroasele vase de sânge care acoperă creierul, Aristotel a presupus că acestea servesc la disiparea excesului de căldură din inimă.

Această explicație l-a condus și la o altă idee interesantă despre dimensiunea creierului uman. El credea că, deoarece oamenii sunt mai „fierbinți” decât alte mamifere, au nevoie de un creier mai mare pentru a regla această căldură.⁵⁸

La câteva decenii după aceste experimente rudimentare ale lui Aristotel, în orașul Alexandria a apărut un alt filozof pe nume Herophilus, care a adus contribuții admirabile la anatomia creierului. El a recunoscut, de exemplu, conexiuni foarte specifice între diferite părți ale corpului și creier prin intermediul măduvei spinării, pe care noi le numim acum nervi cranieni. El a descris șapte perechi și a formulat originile

⁵⁸ Gross, C.G. (1995) "Aristotle on the Brain", *The Neuroscientist*, 1 (4): 245–250.

acestora: facial, auditiv, optic, hipoglosal, trigemen și oculomotor.

Astfel, Herophilus a contribuit la teza conform căreia creierul era responsabil de funcțiile mintale în general. Pentru contribuțiile sale cu adevărat remarcabile, a devenit cunoscut ca primul mare anatomist, după ce a fondat Școala de Medicină din Alexandria, unde și-a efectuat o mare parte din cercetările sale pe cadavre.

După ce am explorat Egiptul Antic și Grecia Antică, ne îndreptăm acum către următoarea etapă a acestei călătorii istorice: Imperiul Roman și epoca Colosseumului, unde îl întâlnim pe Galenus din Pergam. Considerat fondatorul fiziologiei, Galenus nu a avut privilegiul de a diseca cadavre umane, așa cum o făcuse Herophilus, ci și-a bazat concluziile pe disecția animalelor.

Una dintre cele mai remarcabile realizări ale sale a fost diferențierea sistemului nervos în două mari componente. Pe de o parte, a identificat căi nervoase care intră și ies din creier prin baza craniului. Pe de altă parte, a descris un set de căi nervoase conectate prin măduva spinării. În lucrarea sa *Despre procedeele anatomice*, Galenus reușește să descrie 10 din cele 12 perechi de nervi cranieni, deși este posibil să fi confundat unele dintre ele.

Aceste descoperiri, alături de altele din perioada medievală, au deschis calea către modernitate. Odată

cu René Descartes, apare o nouă perspectivă asupra relației dintre minte și corp, împărțind realitatea în două substanțe fundamentale:

- materia, care alcătuiește majoritatea entităților existente;
- mintea sau sufletul, care este imaterială și distinctă de corp.

Acest cadru conceptual este cunoscut sub numele de dualismul substanței, deoarece sugerează o separare clară între aspectele fizice și aspectele non-fizice ale existenței umane.

Însă, această concepție ridică o problemă majoră: Cum poate o substanță imaterială, precum mintea, să influențeze o substanță materială, precum corpul?

Descartes a încercat să răspundă acestei dileme, realizând un studiu rudimentar al creierului uman. Observând că majoritatea structurilor cerebrale sunt dublate în perechi (emisfera dreaptă și emisfera stângă), a remarcat că o singură structură părea să fie unică: glanda pineală.

Bazându-se pe această observație, Descartes a susținut că interacțiunea dintre minte și corp se realizează în acest organ, care ar funcționa ca un punct de legătură între domeniul fizic și cel mental.

Pe lângă faptul că a fost filosof și matematician, Descartes a fost și tutorele prințesei Elisabeta de Boemia, căreia i-a dedicat una dintre lucrările sale. Elisabeta a formulat una dintre cele mai puternice obiecții la adresa dualismului cartezian:

Cum poate ceva imaterial, precum sufletul, să influențeze în mod causal ceva material, precum creierul și corpul?

Să luăm un exemplu: dacă o persoană își formează convingerea mentală „vreau să închid această carte”, cum anume această intenție abstractă determină un act fizic, ridicarea mâinii și închiderea efectivă a cărții?

Această obiecție a fost atât de puternică încât a determinat numeroși filosofi să caute alternative pentru a explica relația dintre minte și corp.

O modalitate de a răspunde obiecției Elisabetei a fost negația interacțiunii directe între minte și corp.

Nicholas Malebranche, un filosof francez, a susținut o teorie numită "ocazionalism". Conform acestei perspective, nu există o relație causală directă între minte și corp. În schimb, Dumnezeu este cel care orchestrează această legătură, acționând ca un intermediar. De exemplu, atunci când cineva decide să închidă o carte, nu mintea este cea care influențează corpul, ci Dumnezeu care se asigură că acțiunea fizică are loc.

Gottfried Wilhelm Leibniz a propus o variantă ușor diferită, cunoscută sub numele de "paralelism". Conform acestei teorii, mintea și corpul nu interacționează deloc, dar funcționează într-o armonie perfect sincronizată, ca două ceasuri setate să indice aceeași oră. Aparenta conexiune dintre gând și acțiune este doar rezultatul unei sincronizări prestabilite.

Problema fundamentală cu aceste două teorii este că ambele presupun existența unui Dumnezeu, care fie acționează constant pentru a menține interacțiunea (*ocasionalism*), fie a programat această sincronizare încă de la începutul universului (*paralelism*).

Conform ocazionalismului, de fiecare dată când un cititor își formează intenția de a închide cartea, Dumnezeu intervine pentru a se asigura că mâna sa execută această acțiune.

În schimb, paralelismul susține că Dumnezeu nu intervine în fiecare moment, ci a stabilit de la începutul lumii un mecanism perfect sincronizat, astfel încât gândurile și acțiunile să pară conectate, fără a fi cu adevărat cauzate una de cealaltă.

Imaginează-ți două linii paralele: una reprezintă procesele mentale, iar cealaltă acțiunile fizice. Deși ele nu se întâlnesc niciodată, sunt perfect sincronizate,

creând iluzia unei conexiuni cauzale între gând și mișcare.⁵⁹

Criticile aduse dualismului au stârnit suspiciuni și au determinat filosofi și oameni de știință să ia în considerare o altă perspectivă asupra relației dintre minte și corp: *monismul materialist*, conform căruia mintea și creierul sunt, de fapt, constituite din aceeași substanță, aceasta fiind materia.

Această nouă viziune a deschis calea către un studiu științific riguros al creierului și al funcțiilor sale mentale, bazat pe observații sistematice și modele empirice. Astfel, s-a conturat o nouă disciplină: *neuroștiința*.

Fundamentele acestei discipline au fost puse în 1810, odată cu proiectul frenologiei, propus de Franz Joseph Gall și J. G. Spurzheim. Frenologia a încercat să coreleze forme și iregularități ale craniului uman cu diferite funcții mentale, cum ar fi limbajul, percepția și conștiința.

Cu toate acestea, această abordare "localizaționistă" a fost rapid criticată. În Franța, fiziologul Jean-Pierre Flourens a respins ideea că anumite funcții cognitive sunt strict localizate într-o anumită regiune a creierului. În schimb, John Hughlings Jackson, un neurolog englez care a studiat pacienți epileptici, a observat că anumite

⁵⁹ Gouveia, S. (2018) *Philosophical Reflections: Art, Mind and Justice*, Braga: Editora Húmus.

funcții erau într-adevăr asociate cu regiuni specifice ale creierului.

Un moment important în confirmarea localizării funcțiilor cerebrale a fost cazul celebru raportat de Paul Broca în 1861. Broca a analizat un pacient care suferise o leziune în lobul frontal stâng și care prezenta dificultăți severe de vorbire. Această regiune, cunoscută astăzi sub numele de "zona lui Broca", s-a dovedit a fi esențială pentru producerea limbajului.

La scurt timp după, în Germania, Karl Wernicke și-a obținut doctoratul în 1876, studiind un caz similar, dar care implica o altă regiune cerebrală afectată de un accident vascular cerebral. Această zonă, cunoscută astăzi ca "zona lui Wernicke", s-a dovedit a fi unul din elementele lipsă pentru înțelegerea limbajului.

În paralel cu studiile asupra localizării funcțiilor cognitive, în Italia, Camillo Golgi a adus o altă contribuție fundamentală în domeniul neuroștiințelor. El a dezvoltat colorația Golgi, o tehnică revoluționară care a permis vizualizarea neuronilor individuali prin infuzarea țesutului cerebral cu cromat de argint.

Această metodă a oferit o nouă perspectivă asupra arhitecturii neuronale, permițând studierea în detaliu a dendritelor, axonilor și corpurilor celulare.

Ulterior, în Spania, Santiago Ramón y Cajal a perfecționat această tehnică și a propus o idee

revoluționară: *neuronii sunt unități independente care transmit informația doar într-o direcție, de la dendrite la axoni*. Această descoperire contrazicea teoria dominantă la acea vreme, conform căreia creierul funcționează ca o rețea continuă de celule interconectate.

Aceste progrese au consolidat legătura dintre funcțiile mentale și structura creierului, punând bazele teoriei neuronale moderne.

Un alt factor cheie în progresul neuroștiințelor a fost studiul tulburărilor neurologice. În perioada Primului Război Mondial, neurologi precum Erwin Strauss, Kurt Goldstein și Hartmut Kühlenbeck au analizat leziunile cerebrale ale soldaților răniți. Inițial, cercetările lor s-au axat pe localizarea exactă a leziunilor și impactul acestora asupra comportamentului.

Totuși, acești cercetători au adoptat și o abordare unică: au intervievat pacienții, încercând să înțeleagă modul în care leziunile cerebrale le-au afectat experiențele corporale și percepția asupra lumii.

Această metodă a reprezentat prima tentativă de a corela experiențele subiective ale conștiinței cu caracteristicile neuronale ale creierului, deschizând calea pentru neuroștiințele cognitive moderne.

Progresele în dezvoltarea de tehnologii specifice pentru explorarea mai directă a creierului au fost, de

asemenea, fundamentale pentru obținerea mai multor informații despre sistemul nervos. De exemplu, cercetările savantului german Hans Berger au făcut posibilă înregistrarea, pentru prima dată, a activității electrice din craniile ființelor umane.

Această tehnică va fi îmbunătățită pe parcursul mai multor ani, până când, în anii 1930, Berger a dezvoltat tehnica electroencefalogrammei (mai bine cunoscută sub numele de EEG), care a permis identificarea mai multor ritmuri cerebrale particulare.

De exemplu, mai rar, găsim unda delta, care se situează în intervalul 0,5-4 Hz și este asociată cu etapele de somn profund. Mai frecvent, undele gamma, care se situează în intervalul 30-100 Hz., sunt asociate cu procesele superioare, cum ar fi memoria, percepția sau învățarea.

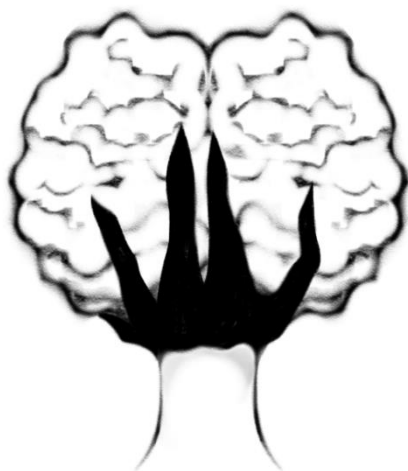
În cele din urmă, în ultimele decenii au fost dezvoltate și alte tehnologii relevante pentru studiul creierului, inclusiv tomografia cu emisie de pozitroni (PET) și imagistica prin rezonanță magnetică funcțională (fMRI). Aceste tehnologii nu măsoară direct activitatea electrică, ci evaluează activitatea biochimică sau metabolică a neuronilor.

De exemplu, folosind fMRI, putem cuantifica cu o anumită precizie cantitatea de molecule de oxigen consumate de neuroni și, pe baza acestor date, putem deduce nivelul activității electrice a acestora. Pe de altă parte, dacă folosim PET, utilizăm substanțe radioactive

pentru a urmări receptorii specifici sau diferite echilibre biochimice din neuroni.

Distincția fundamentală dintre tehnologiile EEG și fMRI/PET constă în tipurile de rezoluții care pot fi obținute: în cazul EEG, putem obține măsurători temporale de înaltă rezoluție, de ordinul milisecundelor, dar cu o rezoluție spațială scăzută.

Cu fMRI/PET, obținem o rezoluție spațială ridicată, dar cu o rezoluție temporală scăzută, de ordinul secundelor. Datorită acestei distincții, multe studii actuale din domeniul neuroștiințelor vizează combinarea atât a EEG, cât și a fMRI/PET pentru a obține o înțelegere mai cuprinzătoare și mai precisă a datelor.⁶⁰



Imaginea 16: Fisico e mentale” © Gaia Amoroso

⁶⁰ Gouveia, S. (2022) *Philosophy and Neuroscience: a Methodological Analysis*, New York: Palgrave Macmillan.

După această scurtă odisee istorică despre creier și reflecții filosofice asupra naturii, compoziției și relevanței sale, în această a doua parte vom prezenta câteva abordări stimulative ale naturii creierului. Vom aprofunda discuțiile cu patru cercetători de renume internațional - Anil Seth, Karl Friston, Christof Koch și Joseph LeDoux - cu care am avut privilegiul de a discuta aceste aspecte.

Prin urmare, să începem această a doua parte, dedicată creierului, prin introducerea unei viziuni mai contemporane asupra acestui organ, care susține că, spre deosebire de ceea ce s-a crezut până atunci, creierul poate fi văzut ca o "mașină de predicție" a cărei funcție este de a crea ipoteze despre lume.

II. Creierul predictiv

Primul dialog din această a doua parte, dedicată creierului, îl are în prim-plan pe neurologul care a popularizat ideea conform căreia creierul este o „mașină de predicție”. Acesta este Anil Seth, profesor de neuroștiințe cognitive și computaționale la Universitatea Sussex din Anglia și co-director al Centrului Sackler pentru Știința Conștiinței. Discursul său TEDx, *Your Brain Hallucinates Its Conscious Reality*,⁶¹ a acumulat peste 10 milioane de vizualizări, captând atenția publicului larg asupra unei perspective revoluționare asupra percepției.

Seth susține teoria procesării predictive (PP), o viziune care contestă ideea tradițională conform căreia creierul are un rol pur pasiv: acela de a recepționa stimuli din mediu și de a construi, pe baza acestora, o hartă mentală a realității. Din această perspectivă clasică, creierul nu face decât să reflecte fidel lumea exterioară, fără a interveni activ în modelarea experienței.

În contrast, teoria procesării predictive propune că, departe de a fi un simplu receptor, creierul funcționează ca un mecanism sofisticat de anticipare. Mai exact, el nu doar recepționează stimuli, ci îi

⁶¹ YouTube: <https://www.youtube.com/watch?v=lyu7v7nWzfo>.

combină și îi interpretează activ, formulând în permanență „ipoteze” despre cauzele acestora.

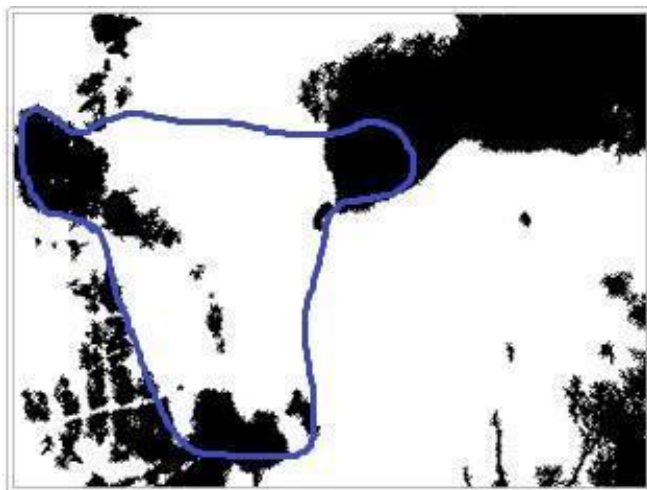
Conform viziunii tradiționale, cititorul acestei cărți primește stimuli vizuali din pagină, iar abia apoi creierul său construiește o reprezentare mentală a textului. Procesarea predictivă însă infirmă această explicație: percepția nu este un simplu proces reactiv, ci unul activ și constructiv. Creierul nu „așteaptă” să primească semnale din exterior, ci generează constant predicții despre ceea ce urmează să perceapă, ajustându-le ulterior în funcție de informațiile senzoriale primite. Astfel, conform acestei teorii, percepția poate fi văzută ca o „halucinație controlată” – un model anticipativ al semnalelor vizuale pe care le vom întâlni. Doar după ce creierul generează această predicție, el verifică dacă semnalele primite din exterior o confirmă sau dacă trebuie ajustată pe baza unor informații suplimentare. Dacă această explicație pare abstractă, cititorul este invitat să analizeze următoarea imagine:



Imaginea 17. Ilustrare a procesării predictive (percepție vizuală)

Dacă cititorul nu a văzut niciodată această imagine, este posibil să o perceapă ca fiind lipsită de sens – o compoziție abstractă, fără nicio semnificație recognoscibilă, care ar putea trece drept o lucrare haotică, asemănătoare cu ceea ce ar picta un copil în stilul lui Jackson Pollock. Nu există nimic familiar în acele forme și umbre care să sugereze un obiect sau o scenă recognoscibilă.

Cu toate acestea, deși creierul tău poate întâmpina dificultăți în a identifica ceva anume, teoria propusă de Anil Seth devine mai convingătoare atunci când privești a doua imagine, care conține doar o mică modificare. Prin simpla adăugare a unei linii ce conturează o parte din formele și umbrele inițial amorfe, creierul reușește brusc să descifreze ce reprezintă imaginea: o siluetă rudimentară a feței unei vaci.



Imagina 18. Ilustrare a procesării predictive (percepție vizuală) II

Cel mai fascinant aspect al acestei experiențe vizuale este că, odată ce ai văzut a doua imagine și ai identificat conturul vacii, vei putea recunoaște aceeași formă și în prima imagine, chiar și fără ajutorul liniei. Inițial, creierul tău nu putea descifra acest detaliu, dar acum, după actualizarea modelului perceptual, interpretarea devine automată.

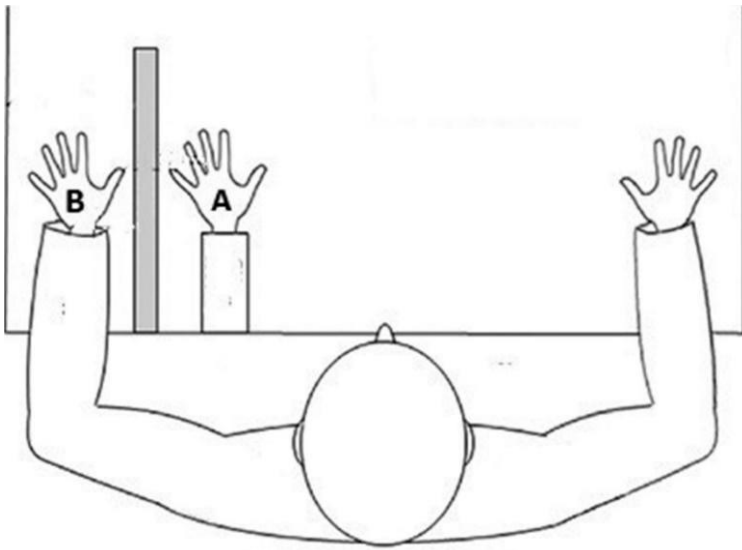
Cum explică procesarea predictivă acest fenomen? În esență, teoria sugerează că, după ce creierul a procesat a doua imagine, acesta a actualizat modelul său inițial, lipsit de sens, integrând noile informații. Astfel, ceea ce anterior părea haotic și neinteligibil capătă acum o semnificație clară.

Un detaliu esențial este că informația senzorială nu s-a schimbat deloc – prima imagine a rămas aceeași înainte

și după vizualizarea celei de-a doua. Ceea ce s-a modificat este „presupunerea” pe care creierul tău o formulează despre stimulul vizual, influențată de experiența anterioară. Aceasta demonstrează că percepția noastră nu este doar rezultatul stimulilor din lumea externă, ci și al modului în care creierul prezice și interpretează acești stimuli.

Teoria procesării predictive nu se aplică doar percepției, ci și altor domenii ale minții umane. Această perspectivă sugerează că multe fenomene mentale – de la conștiință și memorie până la afecțiuni psihiatrice – pot fi explicate prin același mecanism fundamental: creierul nu doar reacționează la realitate, ci o anticipează activ.

Un exemplu fascinant al aplicabilității acestei teorii la percepția sinelui este iluzia mâinii de cauciuc, un experiment bine cunoscut în psihologie. Acesta explorează modul în care creierul prezice limitele corpului nostru și ceea ce consideră că face parte din sine. Structura acestei experiențe este următoarea:



Imaginea 19. Ilustrație a procesării predictive de sine;
Mâna A = de cauciuc/falsificată; mâna B = adevărată.

Imaginează-ți că participi la acest experiment. Stai la o masă, iar mâna ta stângă se sprijină pe suprafața acesteia, însă nu o poți vedea deoarece este ascunsă de un obstacol (marcat în imagine cu punctul B). În fața ta, puțin mai la dreapta locului unde se află mâna reală, este plasată o mână de cauciuc, poziționată exact acolo unde ai percepe că ar fi mâna ta stângă dacă nu ar fi obstrucționată (punctul A în imagine).

Experimentul continuă cu o persoană care atinge simultan atât mâna reală (pe care nu o vezi), cât și mâna de cauciuc (pe care o poți vedea). După câteva minute de stimulare sincronizată, se produce un fenomen surprinzător: dacă, de exemplu, un personaj amenințător precum Michael Myers ar apărea brusc și

ar înfige un cuțit în mâna de cauciuc, ai reacționa imediat ridicându-ți mâna stângă reală, ca și cum ai fi fost rănit.

Ce s-a întâmplat aici? Explicația constă în modul în care creierul procesează informațiile senzoriale. Prin combinarea stimulilor tactili (atingerile simultane) cu cei vizuali (mâna de cauciuc aflată în câmpul vizual), creierul construiește un model intern al corpului tău, presupunând că mâna artificială face parte din tine. Această iluzie demonstrează că, din perspectiva procesării predictive, chiar și percepția propriului corp este o „halucinație controlată” sau o presupunere generată de creier pe baza indiciilor disponibile.

La fel cum putem fi induși în eroare de iluzii vizuale, modul în care ne experimentăm propriul corp poate duce la modele distorsionate. Această perspectivă are implicații profunde în înțelegerea unor tulburări psihiatrice care afectează percepția sinelui, cum ar fi schizofrenia sau tulburarea bipolară.

Dar cum funcționează, mai exact, această predicție continuă în creier, atât în raport cu realitatea externă, cât și cu semnalele interne ale corpului nostru? Un concept central pentru explicarea acestui proces este inferența bayesiană, un mecanism prin care creierul își

formează și își ajustează în permanență modelele predictive pentru a se adapta la noile informații.⁶²

În esență, inferența bayesiană presupune combinarea a două surse de cunoaștere:

1. predicțiile interne ale creierului, cunoscute sub numele de prior – ipotezele sau așteptările pe care le avem despre lume înainte de a primi noi date;
2. informațiile senzoriale din mediu, numite verosimilitudine – dovezile externe care pot confirma sau infirma predicțiile inițiale.

Aceste două surse sunt integrate pentru a forma o estimare actualizată, numită posterior, adică modelul îmbunătățit al realității. În mod constant, creierul ajustează acest echilibru între așteptările sale și dovezile externe, permițând o adaptare optimă la mediu.

Pentru a ilustra acest mecanism, să luăm un exemplu simplu. Presupunem că, înainte de a intra în casă, cititorul estimează că există o probabilitate foarte mică (1%) de a găsi o pisică înăuntru, deoarece știe că nu deține una. Acesta este prior-ul său – o presupunere inițială bazată pe experiența anterioară.

⁶² Gouveia, S. & Curado, M. (2020) (eds.) *The Philosophy and Science of Predictive Processing*, New York: Bloomsbury.

Însă, imediat după ce deschide ușa, aude un mieunat. Această nouă informație senzorială, care intră sub incidența verosimilității, trebuie integrată în modelul mental al realității. Având în vedere că pisicile miaună, creierul reevaluează situația și actualizează credința inițială, crescând probabilitatea ca o pisică să fie prezentă în casă la 90%. Prin combinarea predicției inițiale cu noile dovezi senzoriale, creierul ajunge la o nouă concluzie (posterior) – aceea că prezența unei pisici este acum mult mai probabilă.

Acest exemplu simplu demonstrează cum creierul își ajustează modelele predictive în mod constant, integrând noi date pentru a menține o reprezentare cât mai precisă a realității.

Pentru a înțelege mai bine implicațiile acestei abordări, trebuie să explorăm concepte fundamentale precum inferența activă, eroarea de predicție și principiul energiei libere, care vor fi analizate în secțiunea următoare.

Însă, înainte de a aprofunda aceste concepte, este util să examinăm modul în care această teorie poate contribui la înțelegerea diverselor tulburări psihiatrice.

Dacă, așa cum am argumentat anterior, creierul este o mașină de predicție care îmbină modelele interne cu stimulii din lume, atunci putem presupune că, în cazul schizofreniei, această capacitate de a face predicții

precise este alterată. Dar care sunt mecanismele implicate?

Un posibil răspuns este că indivizii cu schizofrenie pot avea dificultăți în a distinge între sursele interne și externe ale informației, ceea ce îi face să confunde propriile gânduri cu stimuli externi. De exemplu, unul dintre cele mai frecvente simptome pozitive ale schizofreniei sunt halucinațiile auditive: persoana percepe „voci” ca venind din exterior, deși acestea sunt generate intern de creier. Cu alte cuvinte, predicțiile eronate ale creierului determină individul să interpreteze un fenomen intern ca având o origine externă.



Imaginea 20: "Rifiuto della coscienza" © Gaia Amoroso

Un alt exemplu de aplicabilitate a acestei teorii poate fi observat în autism. Se poate argumenta că persoanele cu autism au dificultăți în integrarea informațiilor senzoriale externe, ceea ce le determină să acorde o importanță mai mare propriilor modele interne. Deoarece întâmpină dificultăți în procesarea noilor stimuli senzoriali, capacitatea lor de a-și actualiza modelele mentale și de a se adapta eficient la lume este compromisă. Acest lucru ar putea explica, cel puțin parțial, tendința lor de a prefera rutine rigide și dificultatea de a anticipa și interpreta comportamentele sociale ale celorlalți.

Mai departe, să ne imaginăm o persoană al cărei creier are o tendință negativă în anticiparea modelelor realității. Aceasta ar duce la așteptări pesimiste constante și la o capacitate redusă de a prevedea rezultate pozitive. Un astfel de individ ar dezvolta o „rigiditate” în modelul său relațional cu lumea, ceea ce înseamnă că percepțiile negative persistă și nu permit integrarea de noi informații pozitive. Această incapacitate de actualizare a modelului ar putea explica mecanismele depresiei, o tulburare care afectează milioane de oameni.

Se știe, de exemplu, că persoanele cu depresie tind să se perceapă pe sine într-o manieră negativă, filtrând selectiv informațiile care le confirmă stima de sine scăzută. Din perspectiva procesării predictive, acest lucru s-ar putea datora unei predicții eronate, greu de

modificat, care întărește o viziune distorsionată asupra sinelui și lumii.

Dar ce se întâmplă dacă un creier acordă o importanță exagerată stimulilor amenințători și generează în mod constant predicții despre evenimente viitoare periculoase? Această tendință ar duce la o anticipare excesivă a pericolului și la o stare de alertă permanentă. Rezultatul? O persoană predispusă la anxietate cronică, prinsă într-un ciclu de îngrijorări și stres anticipativ.

Prin urmare, înțelegerea creierului ca o mașină de predicție ne oferă o nouă perspectivă asupra tulburărilor psihiatrice. Aceasta sugerează că problemele mintale pot fi privite ca disfuncții ale proceselor predictive – fie că este vorba de predicții distorsionate (schizofrenie), inflexibile (depresie) sau hiperactive (anxietate).

Așa cum am încercat să vă arăt cu aceste scurte exemple din psihiatrie, "Teoria procesării predictive" prezintă un potențial explicativ semnificativ și ar putea contribui într-adevăr la elucidarea mecanismelor specifice în diverse boli mintale.

În plus, poate, de asemenea, să ne explice modul în care funcționează percepția, sinele, memoria, visele și chiar conștiința.

În continuare, vom prezenta câteva idei care constituie baza și fundamentul acestei perspective, așa cum au

fost dezvoltate de unul dintre cei mai influenți
nerocercetători la nivel mondial, profesorul Karl
Friston.

III. CREIERUL LIBER

Al doilea dialog îl are ca invitat pe profesorul Karl Friston, una dintre cele mai influente figuri în domeniul neuroștiințelor și un pionier al perspectivei discutate anterior. După introducerea în procesarea predictivă, urmează să explorăm în detaliu câteva principii fundamentale care stau la baza acestei teorii asupra creierului uman.

Unul dintre conceptele centrale ale acestei abordări este Principiul energiei libere. Acest principiu are rădăcini adânci în istoria științei, fiind dezvoltat de-a lungul timpului de gânditori din multiple discipline, dar a fost formalizat recent în neuroștiințe de Karl Friston, cel mai citat om de știință în viață.

În esență, principiul energiei libere sugerează că creierul uman are o tendință intrinsecă de a minimiza consumul de energie atunci când procesează informații și desfășoară activități cognitive. Pentru a conserva resursele, creierul caută să funcționeze într-un mod cât mai eficient, optimizând continuu predicțiile despre lume. Cu alte cuvinte, scopul său este de a construi modele interne precise ale realității cu un efort minim.

Acest principiu poate fi înțeles ca un echilibru între două procese esențiale:

1. menținerea unor modele interne stabile, pentru a evita consumul inutil de energie prin recalculări constante;
2. flexibilitatea necesară pentru a integra noi informații, astfel încât creierul să poată răspunde schimbărilor din mediu fără a rămâne rigid.

Prin această abordare, se explică modul în care organismele gestionează incertitudinea într-un mod eficient și economic, reducând surprizele neprevăzute și adaptându-se cât mai rapid la noi condiții.

Pe scurt, principiul energiei libere susține că economia de energie este unul dintre principiile fundamentale care guvernează organizarea și funcționarea creierului, având drept scop final supraviețuirea organismului.



Imaginea 21: "Libertà" © Gaia Amoroso

Matematic, principiul energiei libere are o formulare complexă, însă conceptul său de bază este relativ intuitiv. Imaginați-vă că cititorul se află la volan, conducând pe un traseu familiar, de exemplu, drumul de la serviciu către casă, pe care îl parcurge zilnic.

În timp, creierul a construit un model intern al acestui traseu, bazat pe experiențele anterioare. Acest model include detalii esențiale, precum poziția curbilor, a semafoarelor, a indicatoarelor rutiere și a altor repere. Scopul acestui proces este de a minimiza efortul necesar pentru a anticipa și reacționa la caracteristicile drumului de fiecare dată când îl parcurgeți.

Creierul utilizează datele acumulate pentru a face predicții rapide și eficiente. De exemplu, dacă știi că la capătul unei străzi se află un semafor, creierul tău îl va anticipa automat, reducând surpriza în cazul în care acesta devine roșu. În acest fel, creierul construiește un model stabil al drumului, pe care îl actualizează doar atunci când este necesar, un mecanism care permite economisirea semnificativă a resurselor cognitive.

Ce se întâmplă, însă, dacă pe traseu apare o modificare neașteptată, cum ar fi instalarea unui nou semafor? Creierul va trebui să se adapteze și să integreze această informație în modelul existent. În loc să construiască de la zero o nouă hartă mentală a drumului, acesta va

ajusta modelul anterior pentru a reflecta schimbarea. Aceasta este esența principiului energiei libere: creierul încearcă în permanență să reducă surpriza și să optimizeze predicțiile pentru a funcționa eficient, fără a consuma resurse inutile.

El nu doar explică modul în care creierul uman minimizează efortul cognitiv, ci și modul în care toate organismele vii, de la bacterii până la mamifere complexe, își optimizează strategiile de supraviețuire.

La un nivel fundamental, principiul energiei libere descrie o lege universală a organizării biologice: orice sistem viu, pentru a-și menține existența, trebuie să reducă incertitudinea și să optimizeze consumul de energie. Această perspectivă oferă o nouă înțelegere asupra minții umane, sugerând că procesele cognitive și comportamentele noastre sunt profund influențate de necesitatea eficienței și predictibilității.

Atât organismele simple, cât și cele sofisticate sunt sisteme dinamice, distincte de mediul în care există. Această separare presupune o interacțiune constantă și un schimb continuu de informații între organism și mediu. Scopul acestei interacțiuni este de a îndeplini un alt principiu esențial: minimizarea erorilor de predicție.

Până acum, am argumentat că modelele predictive ale creierului trebuie să fie cât mai precise pentru a reduce surpriza și a evita consumul excesiv de energie. Cu cât

un model este mai precis, cu atât se apropie mai mult de realitate, necesitând mai puține corecții și ajustări. Pe de altă parte, un model imprecis presupune un cost energetic ridicat, deoarece necesită actualizări frecvente.

Astfel, toate organismele tind să minimizeze erorile de predicție asupra mediului înconjurător prin aplicarea principiului energiei libere. Dar ce înseamnă mai exact energie liberă în acest context?

În acest cadru teoretic, energia liberă poate fi înțeleasă ca o măsură a surprizei sau incertitudinii predicțiilor creierului. Cu cât un organism poate minimiza mai eficient energia liberă, cu atât își poate asigura mai bine supraviețuirea.

Acest principiu nu se aplică doar în procesarea informațiilor despre lumea externă, ci și în reglarea stărilor interne ale organismului. Pe lângă anticiparea evenimentelor din mediu, creierul face predicții continue despre nevoile fiziologice și funcționarea organelor interne.

Acest mecanism poate fi înțeles astfel: dacă există erori în predicțiile despre mediul extern sau despre procesele interne ale corpului, se generează un nivel ridicat de energie liberă. Acest lucru corespunde unei stări de surpriză, imprevizibilitate sau stres fiziologic.

Pe de altă parte, atunci când predicțiile sunt precise, energia liberă rămâne scăzută, ceea ce înseamnă că organismul funcționează eficient, fără a consuma resurse inutile. Astfel, toate formele de viață, indiferent de complexitate, de la bacterii și plante la mamifere și oameni, sunt ghidate de acest mecanism fundamental de reducere a incertitudinii.

O altă caracteristică a acestui principiu este perspectiva diferită pe care o oferă asupra biologiei. În timp ce majoritatea teoriilor biologice încearcă să explice de ce un fenomen funcționează într-un anumit mod, abordarea energiei libere inversează întrebarea. În loc să descrie ce trebuie să facă un organism pentru a exista, această teorie încearcă să înțeleagă ce procese sunt necesare pentru menținerea existenței sale.



Imaginea 22: "Io esisto" © Gaia Amoroso

Această schimbare de perspectivă este importantă pentru a înțelege că principiul energiei libere nu este, în sine, o teorie falsificabilă despre comportamentul organismelor. În schimb, diferitele sale postulate pot fi testate empiric, în funcție de aplicabilitatea lor în diverse domenii ale științei cognitive și neuroștiinței.

Pentru a încheia această introducere, este necesar să discutăm un alt concept fundamental: inferența activă.

Acest mecanism explică modul în care creierul nu doar ajustează pasiv modelele predictive, ci și acționează activ pentru a obține informații care să reducă incertitudinea. Atunci când creierul detectează erori sau lacune în modelele sale predictive, el nu așteaptă pur și simplu noi date, ci explorează activ mediul pentru a obține informații relevante.

Astfel, inferența activă reprezintă capacitatea creierului de a influența informațiile pe care le primește din mediul înconjurător. Acest proces presupune:

- *selectarea stimulilor relevanți* deoarece creierul filtrează datele disponibile și se concentrează pe cele mai utile pentru ajustarea predicțiilor;
- *direcționarea atenției* atât spre factori externi (schimbări din mediu), cât și spre factori interni (stări fiziologice) care pot reduce ambiguitatea sau incertitudinea.

Pentru a înțelege mai bine modul în care inferența activă funcționează în reglarea stărilor interne ale corpului, să luăm un exemplu concret. Imaginează-ți că, dintr-o clipă în alta, ești transportat de pe plaja însorită a Copacabanei în impresionații Alpi elvețieni.

În această tranziție bruscă, corpul tău detectează instantaneu schimbarea de temperatură prin interocepție – procesul prin care sistemul nervos monitorizează stările interne ale organismului. Senzația de frig pe care o resimți este un semnal interoceptiv care indică faptul că temperatura corpului tău este pe cale să scadă dacă rămâi în acest mediu rece.

Aici intervine inferența activă: creierul nu doar recepționează acest semnal pasiv, ci începe să caute informații suplimentare din mediu și din propriul corp pentru a calibra răspunsul fiziologic optim. De exemplu, este posibil ca atenția ta să fie direcționată către tremuratul mușchilor, un mecanism fiziologic automat care generează căldură și care îți confirmă scăderea temperaturii interne.

Pe baza acestor informații, creierul își ajustează modelul predictiv al temperaturii corporale. Dacă, pe plaja Copacabana, corpul tău folosea transpirația pentru a regla căldura, în Alpi elvețieni, mecanismele fiziologice trebuie să se schimbe – tremuratul, constricția vaselor de sânge și căutarea unui adăpost devin prioritare pentru menținerea homeostaziei.

Astfel, acest exemplu demonstrează modul în care creierul optimizează în mod activ reglarea temperaturii, concentrându-se atât pe semnalele externe din mediu, cât și pe semnalele interne ale corpului. Dacă această ajustare nu ar avea loc, rezultatul ar putea fi dezastruos – în acest caz, hipotermia.

Deși principiul energiei libere a oferit o perspectivă unificatoare asupra funcționării creierului, acesta nu este lipsit de critici. Unii neurocercetători și filosofi⁶³ au susținut că nu toate procesele biologice pot fi explicate exclusiv prin încercarea de a minimiza consumul de energie.

De exemplu, funcții mentale precum creativitatea și imaginația par să implice un consum energetic ridicat, care nu pare să fie justificat doar prin optimizarea eficienței. În plus, variabilitatea neuronală între indivizi ridică întrebări cu privire la aplicabilitatea universală a acestui principiu – dacă toți oamenii ar funcționa exclusiv pe baza minimizării energiei libere, ar trebui să vedem o mai mare uniformitate în procesele cognitive și comportamentale, ceea ce nu este cazul.

O altă provocare majoră constă în dificultatea măsurării și cuantificării mecanismului predictiv la nivelul întregului organism. În timp ce modelele matematice

⁶³ Articol de Professor Karl Friston: Friston, K. et al. (2023) "The free energy principle made simpler but not too simple", *Physics Reports*, 1024: 1-29.

oferă o explicație teoretică solidă, demonstrarea empirică a acestui proces în toate aspectele biologiei umane rămâne o sarcină complexă.

În continuare, vom explora o altă abordare asupra creierului și a rolului său în generarea proceselor conștiente. Vom analiza o teorie care încearcă să ofere o metodă clară de cuantificare a stărilor conștiente, dezvoltată de profesorul Christof Koch, neurocercetător și președinte al Institutului Allen pentru Creier din Seattle.

IV. CREIERUL INTEGRAT

Acest al treilea dialog introduce una dintre cele mai influente teorii contemporane despre creier și rolul său în generarea minții conștiente, dezvoltată de profesorul Christof Koch, autorul cărții *The Feeling of Life Itself: Why Consciousness Is Widespread but Can't Be Computed*.

Cunoscută sub numele de Teoria Integrată a Informației (Integrated Information Theory, IIT),⁶⁴ această perspectivă a revoluționat modul în care conștiința este înțeleasă în neuroștiințe. Datorită naturii sale teoretice și conceptuale, IIT a avut un impact semnificativ nu doar în științele creierului, ci și în filosofie, fizică și chiar în cercetarea inteligenței artificiale. Pe lângă implicațiile sale științifice, teoria a generat și dezbateri etice importante privind natura conștiinței și criteriile prin care putem determina dacă un sistem, biologic sau artificial, este conștient.

IIT își propune să ofere un răspuns științific și obiectiv la problema conștiinței în lumea fizică. Mai exact, încearcă

⁶⁴ Primul autor care a dezvoltat această teorie a fost Giulio Tononi, în anul 2004, alături de care Christof Koch s-a alăturat ulterior pentru a contribui la dezvoltarea ei, după ce colaborase anterior cu laureatul Premiului Nobel, Francis Crick. Publicație originală: Tononi, G. (2004) "An Information Integration Theory of Consciousness", BMC Neuroscience, 5 (1): 42.

să identifice proprietățile esențiale ale conștiinței, denumite axiome, și să deducă condițiile necesare și suficiente pe care trebuie să le îndeplinească orice substrat fizic pentru a susține o experiență conștientă, numite postulate. Aceste principii sunt exprimate printr-o formulare matematică specifică, care nu va fi abordată în detaliu în această carte, din cauza complexității sale.

Ceea ce diferențiază IIT de alte teorii despre conștiință este abordarea sa inversă față de cercetarea neuroștiințifică tradițională. În loc să înceapă prin identificarea activităților neuronale asociate cu conștiința, Koch pornește de la însăși fenomenologia experienței conștiente și încearcă să determine ce tip de mecanisme fizice ar putea explica această experiență.

Această abordare reflectă o perspectivă similară cu cea a lui René Descartes, care considera conștiința o proprietate fundamentală a realității. Pornind de la această ipoteză, IIT formulează cinci axiome fundamentale:

- Axioma existenței (Aaxis): Această axiomă reformulează celebrul cogito cartezian, înlocuind „gândirea” cu „experiența”: „Experimentez, deci exist”. Conform acestui principiu, avem o certitudine absolută a propriei noastre perspective asupra lumii, iar conștiința este o

realitate fundamentală, independentă de orice altă justificare externă.

- Axioma informației (AInfo): Fiecare experiență conștientă este definită prin ceea ce o face distinctă de toate celelalte experiențe posibile. Cu alte cuvinte, o experiență conștientă nu este doar un simplu eveniment neutru, ci reprezintă o formă de informație unică, care ne diferențiază prezentul de orice altă posibilă stare mentală.
- Axioma integrării (AInteg): Conținutul unei experiențe conștiente nu poate fi redus la elemente separate, ci este integrat într-un întreg coerent. De exemplu, atunci când cititorul privește aceste cuvinte, percepe simultan atât forma, cât și culoarea literelor într-o singură experiență unificată, fără a le procesa separat.
- Axioma compoziției (AComp): Orice experiență conștientă are o structură internă formată din multiple aspecte și combinații între acestea. De exemplu, experiența de a citi un text nu este o percepție unidimensională, ci conține multiple caracteristici fenomenologice – cum ar fi forma, culoarea și contrastul literelor.
- Axioma excluderii (AExcl): Conștiința implică o selecție strictă: în fiecare moment, doar o anumită experiență conștientă este posibilă, iar altele sunt automat excluse. De asemenea, experiențele conștiente au o durată limitată în

timp, fiecare având un început și un sfârșit bine definit.

Pentru a înțelege mai bine cum funcționează Teoria Integrată a Informației (IIT), să aplicăm axiomele sale unui exemplu familiar: citirea unei cărți.



Imaginea 23: "No brain, never mind" © Gaia Amoroso

Imaginează-ți că ești acasă, relaxat pe canapea, răsfoind această carte. Această situație demonstrează prezența unei experiențe conștiente intrinseci (Axioma Existenței – AExis), deoarece ești pe deplin conștient de ceea ce citești.

Pe măsură ce parcurgi textul, informațiile din pagini sunt prelucrate în conștiința ta într-un mod unic:

cuvintele, ideile și cunoștințele transmise de autor formează o experiență distinctă, ceea ce reflectă Axioma Informației (AInfo), fiecare experiență este specifică și diferită de altele.

Toate aceste informații sunt integrate într-o singură experiență coerentă, legătura dintre cuvinte, imagini și emoțiile pe care le simți în timp ce citești creează o percepție unificată (Axioma Integrării – AInteg). Nu percepi fiecare element separat, ci ca parte a unui întreg.

Desigur, această experiență conștientă de lectură este formată din elemente distincte, precum structura textului, ideile autorului sau stilul narativ. Aceste componente formează o compoziție unică a experienței (Axioma Compoziției – AComp), în care fiecare detaliu contribuie la întreaga percepție a lecturii.

În plus, în timp ce citești, alte evenimente au loc în jurul tău, poate un meci de fotbal este difuzat la televizor. Totuși, dacă această carte îți captează atenția, conștiința ta poate exclude aceste alte experiențe, demonstrând caracterul finit și selectiv al conștiinței (Axioma Excluderii – AExcl).

Pe baza acestor axiome, "Teoria Integrată a Informației" definește un set de postulate care descriu proprietățile fundamentale pe care trebuie să le aibă un sistem fizic pentru a fi conștient. Să le examinăm pe rând:

- *Postulatul existenței:* conștiința nu poate exista fără un suport fizic care să aibă putere de cauză și efect. Pentru ca un substrat fizic să susțină conștiința, acesta trebuie să influențeze și să fie influențat de stările sale interne și de mediul său.
- *Postulatul informației:* dacă conștiința este într-adevăr informativă, atunci un sistem conștient trebuie să fie capabil să specifice și să diferențieze anumite stări de altele. Aceasta implică faptul că orice mecanism fizic care susține conștiința trebuie să aibă o structură causală care să îi permită să distingă între diferite stări posibile.
- *Postulatul integrării:* pentru ca un sistem să susțină o experiență conștientă unificată, acesta trebuie să fie ireductibil – adică părțile sale trebuie să fie interdependente. Fiecare element al sistemului trebuie să poată influența întregul și, la rândul său, să fie afectat de acesta. Dacă un sistem poate fi descompus în părți care funcționează independent, atunci acesta nu poate susține conștiința.
- *Postulatul compoziției:* conștiința nu este o experiență singulară, ci o combinație de aspecte distincte care interacționează. Orice sistem conștient trebuie să aibă o arhitectură fizică în care componentele sale se combină într-un mod structurat, formând un ansamblu coerent.

- *Postulatul excluderii*: conștiința este limitată și bine definită. La un moment dat, un sistem conștient poate avea doar o singură stare conștientă, iar aceasta trebuie să fie cea mai ireductibilă dintre toate structurile sale cauzale. Cu alte cuvinte, doar cel mai înalt nivel de integrare a informației dintr-un sistem va reprezenta conștiința sa efectivă.

Să luăm în considerare un alt exemplu pentru a înțelege relevanța a ceea ce tocmai a fost descris. Imaginează-ți că te trezești brusc din somnul tău odihnitor și vezi, în fața ta, peretele albastru al camerei tale. Mai mult decât atât, observi unde este patul tău și realizezi că, în acel moment, ești conștient de tine însuși și de ceea ce te înconjoară.

Acum, ceea ce afirmă Teoria Informației Integrate (IIT) este că această experiență există pentru tine și nu pentru altceva; este o experiență specifică (și nu una generică); este o experiență unitară: ochiul tău drept nu percepe nimic diferit față de ochiul tău stâng (ambele experimentează aceeași experiență); în plus, această experiență este definită în mod specific de ceea ce observi în fața ta și nu de ceea ce este în spatele tău sau în lateral; în cele din urmă, această experiență este structurată de mai mulți factori care se relaționează între ei (de exemplu, corpul tău, în patul tău, situat astfel încât să poți vedea peretele albastru din fața ta), ceea ce îți permite să ai exact acea experiență și nu alta.

Ai observat ce tocmai am făcut? Exact: prin acest exemplu, am aplicat toate postulatele IIT.

Este important de subliniat că, pentru teoreticienii IIT, aceste axiome (și, implicit, postulatele) fac parte dintr-o listă completă și finită: nu există nicio altă proprietate a experienței conștiente care ar trebui considerată esențială pentru existența sa. Dacă un sistem fizic – fie el biologic, artificial sau chiar extraterestru – vrea să fie conștient, va trebui să reproducă, prin postulate, toate axiomele prezentate.

Un alt aspect interesant al acestei teorii este că oferă o încercare de a cuantifica, într-o anumită măsură, capacitatea unui sistem de a integra informații. Această cuantificare este reprezentată prin măsura Φ – litera grecească Phi – care exprimă capacitatea (între zero și un maxim) a unui sistem de a integra informațiile procesate într-un anumit sistem.

Conform IIT, conștiința „apare” atunci când un sistem atinge un nivel ridicat de informație integrată: cu cât această informație este mai intrinsecă și ireductibilă, cu atât valoarea Φ este mai mare și, prin urmare, gradul de conștiință este mai ridicat. Calculul se bazează pe cantitatea de informație care este integrată global în sistem, în comparație cu suma informațiilor prezente în fiecare dintre părțile sale.

Astfel, de fiecare dată când un sistem atinge un Φ maxim, acel sistem va fi considerat conștient; sub această valoare maximă, se susține că sistemul nu dispune de suficientă informație integrată pentru a ajunge la un nivel conștient.

Această abordare cuantificabilă, împreună cu axiomele și postulatele IIT, oferă o teorie incredibil de elegantă și, mai ales, capabilă să-și testeze ipotezele. Să explorăm câteva predicții interesante, bazate pe date empirice reale, care ilustrează puterea explicativă a acestei teorii.

În continuare, vom explora câteva dintre cele mai interesante predicții ale Teoriei Integrate a Informației (IIT), bazate pe observații și experimente reale, care demonstrează puterea explicativă a acestei teorii.

Cerebelul este o regiune importantă a sistemului nervos, responsabilă de controlul mișcărilor și a echilibrului, situată în partea posterioară a capului. Deși reprezintă doar 10% din volumul total al creierului și cântărește aproximativ 150 de grame, cerebelul conține de patru ori mai mulți neuroni decât restul creierului la un loc și este puternic conectat cu alte regiuni cerebrale.⁶⁵

⁶⁵ Lemon, R. & Edgley, S. (2010) "Life without a cerebellum", *Brain*, 133: 652-654.

În ciuda acestor caracteristici impresionante, dacă cerebelul ar fi îndepărtat chirurgical, conștiința ar rămâne complet intactă. Cum este posibil acest lucru?

Conform IIT, răspunsul stă în structura cerebelului deoarece acesta este alcătuit din module neuronale relativ independente, ceea ce înseamnă că informația prelucrată acolo nu este suficient de integrată. Deoarece conștiința depinde de un grad ridicat de integrare a informației (Φ), cerebelul, având un Φ scăzut, nu joacă un rol esențial în generarea experiențelor conștiente. Acest fenomen explică de ce pacienții cu leziuni severe ale cerebelului pot avea dificultăți motorii, dar își păstrează complet conștiința.



Imaginea 24: "Il tassello mancante" © Gaia Amoroso

IIT face o altă predicție interesantă despre ceea ce se întâmplă cu creierul și conștiința în timpul somnului. Dacă presupunem că nivelul de conștiință este determinat de Φ , atunci, în timpul somnului profund, când nu avem experiențe conștiente (cum ar fi în fazele non-REM fără vise), ar trebui să observăm o scădere accentuată a integrării neuronale.

Studiile empirice asupra creierului uman în timpul somnului au confirmat această ipoteză: în fazele de somn profund, conectivitatea corticală se reduce semnificativ.⁶⁶ Aceasta sugerează că, pe măsură ce integrarea informației scade, conștiința se diminuează sau dispare temporar, exact așa cum prezice IIT.

O altă predicție remarcabilă a IIT se leagă de o procedură neurochirurgicală numită callosotomie, utilizată în tratamentul epilepsiei severe.⁶⁷ Aceasta presupune secționarea corpului calos, structura care conectează cele două emisfere cerebrale, pentru a preveni propagarea crizelor epileptice între cele două părți ale creierului. După această intervenție, IIT prezice că pacientul nu va mai avea o singură experiență

⁶⁶ Ediția originală: Massimini, M., ..., Tononi, G. (2005) "Breakdown of cortical effective connectivity during sleep", *Science*, 309: 2228–2232.

⁶⁷ Ediția originală: Gazzaniga, M. (1967) "The split brain in man", *Scientific American*, 217 (2): 24-29.

conștientă unificată, ci două fluxuri conștiente separate, fiecare în câte o emisferă cerebrală.

Dacă unui astfel de pacient i se prezintă o imagine doar ochiului drept, informația este procesată de emisfera stângă, care controlează vorbirea. Dacă îl întrebăm ce a văzut, ne va spune că nu își amintește nimic. Însă, dacă îl rugăm să deseneze imaginea folosind mâna stângă (controlată de emisfera dreaptă, care a procesat de fapt imaginea), acesta va putea desena cu exactitate ceea ce a văzut, chiar dacă verbal nu poate raporta experiența!

Acest fenomen nu are nimic supranatural, ci este o dovadă clară a principiului IIT. Odată ce corpul calos a fost deconectat, creierul a încetat să mai funcționeze ca un sistem unitar, formând în schimb două sisteme independente, câte unul în fiecare emisferă. Fiecare dintre acestea și-a maximizat propriul Φ , generând astfel două fluxuri de conștiință distincte și independente unul de celălalt.

O altă predicție interesantă legată de funcționarea defectuoasă a creierului se concentrează pe leziunile cerebrale deoarece această teorie prezice că orice tip de leziune care afectează capacitatea creierului de a integra informații va împiedica formarea stărilor conștiente.

Studiile recente confirmă că această predicție este într-adevăr corectă. Prin stimularea magnetică transcraniană, mai mulți pacienți cu leziuni cerebrale au

prezentat modificări substanțiale în creierul lor în comparație cu persoanele sănătoase.⁶⁸

Aceste exemple demonstrează puterea explicativă a IIT: teoria nu doar descrie conștiința, ci și face predicții testabile despre modul în care aceasta se manifestă în diferite condiții neurofiziologice. În secțiunea următoare, vom analiza implicațiile acestei teorii asupra înțelegerii conștiinței umane și a posibilității existenței conștiinței în sisteme artificiale.

Ar putea această teorie, pe lângă faptul că oferă o explicație a modului în care creierul biologic este conștient, să facă lumină și în ceea ce privește dezvoltarea sistemelor artificiale conștiente? Răspunsul este mixt: pe de o parte, da, iar pe de altă parte, nu.

În ceea ce privește partea pozitivă, în teorie, IIT nu specifică faptul că un sistem trebuie să fie biologic pentru a fi conștient: indică doar că trebuie să fie un sistem care poate integra informații conform axiomelor prezentate anterior.

Cu toate acestea, din punct de vedere negativ, este clar că inteligența artificială actuală, bazată pe structuri feed-forward în care informația este procesată într-o singură direcție (de la intrare la ieșire), nu va fi niciodată

⁶⁸ Casali, A. et al. (2013) "A theoretically based index of consciousness independent of sensory processing and behavior", *Science Translational Medicine*, 5: 198ra105–198ra105.

conștientă. Acest lucru se datorează faptului că un sistem de acest tip nu va avea niciun Φ , chiar dacă reproduce perfect comportamentul conștient; ar fi doar o simulare. Pentru conștiință, sistemul ar trebui să aibă o structură în care ieșirile pot fi utilizate ca intrări și viceversa.

Pentru a încheia aceste note introductive privind IIT, rămâne să abordăm o critică, care poate fi văzută ca un avantaj, adusă acestei teorii. Pentru unii autori, IIT poate avea consecințe panpsihiste contraintuitive. Și ce este panpsihismul? Panpsihismul este o teorie dezvoltată de diverși intelectuali de-a lungul istoriei, de la Spinoza la Leibniz sau Bertrand Russell⁶⁹ și, mai recent, Philip Goff,⁷⁰ care susține că tot ceea ce există în lumea fizică are proprietăți mentale fundamentale.

Acum, acest lucru înseamnă să afirmăm că conștiința este o proprietate fundamentală, ceea ce contravine ideii avansate de IIT conform căreia unele sisteme complexe, atunci când ating niveluri ridicate de informații integrate, devin conștiente.

În plus, atât panpsihismul, cât și ITT sunt de acord că conștiința nu este un fenomen "totul sau nimic":

⁶⁹ Russell, B. (1927) *The Analysis of Matter*, London: Kegan Paul, Trench, Trubner & Co.

⁷⁰ Goff, P. (2019) *Galileo's Error: Foundations for a New Science of Consciousness*, London: Pantheon Book.

dimpotrivă, conștiința trebuie privită ca un spectru, între inconștientul absolut și conștientul complet.

În ciuda acestor asemănări, pare oarecum radical să identificăm IIT ca o teorie panpsihistă. Deși ambele au, de fapt, angajamente carteziane prin acceptarea conștiinței subiective ca fundamentală în natura sa, panpsihismul susține că conștiința este fundamentală pentru toate entitățile existente, în timp ce IIT sugerează doar condițiile necesare pentru apariția conștiinței în anumite sisteme complexe.⁷¹

Pentru a încheia aceste note introductive, vom aborda o altă perspectivă asupra creierului prin intermediul emoțiilor și rolul acestora în dezvoltarea minții noastre, prezentând unele dintre tezele centrale dezvoltate de profesorul Joseph LeDoux, neurocercetător și director al Institutului Creierului Emoțional de la Universitatea din New York.

⁷¹ Tononi, G. (2012) "Integrated information theory of consciousness: An updated account", *Archives italiennes de biologie*, 150: 56-90; Tononi, G. & Koch C. (2015) "Consciousness: Here, There and Everywhere?", *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 370 (1668): 20140167.

V. CREIERUL EMOȚIONAL

În acest al patrulea și ultimul dialog din această odisee în lumea creierului conștient, vom prezenta câteva dintre conceptele-cheie dezvoltate de profesorul Joseph LeDoux, neurocercetător și autor al cărții *Creierul emoțional*. Primul pas spre înțelegerea a ceea ce presupune acest creier emoțional este înțelegerea structurilor anatomice care alcătuiesc acest sistem vital pentru supraviețuirea fiecăruia dintre noi

Partea creierului care procesează diverse emoții este denumită "sistem limbic" în neuroștiințe. Compus din mai multe părți și situat pe suprafața mediană a creierului mamiferelor, acesta joacă un rol fundamental în funcționarea emoțiilor noastre și, în consecință, în modul în care ne raportăm la noi înșine și la ceilalți.

Prin intermediul sistemului nervos autonom, acesta influențează pozitiv sau negativ funcționarea viscerală și reglarea metabolismului în organism, controlând comportamente considerate esențiale pentru supraviețuirea tuturor mamiferelor.

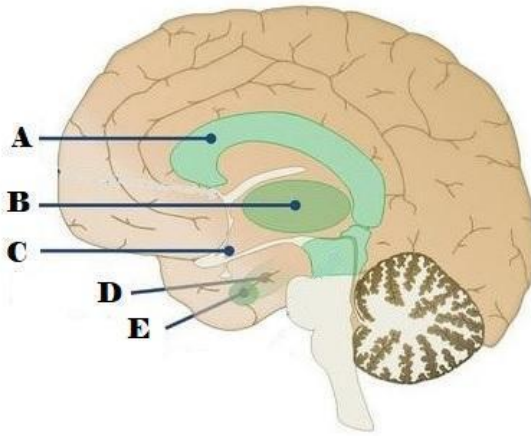
"Răspunsul emoțional adecvat este apoi generat prin integrarea informațiilor senzoriale într-o stare mentală specifică, atribuind conținut afectiv stimulilor înregistrați și asociindu-i cu amintiri preexistente."

Să analizăm câteva dintre principalele structuri anatomice care fac parte din sistemul limbic:

- *Amigdala* este situată în lobul temporal anterior, este conectată la hipotalamus și este considerată "centrul" care ne ghidează în situații periculoase, inducând stări de alertă sau senzații de frică; îndepărtarea sa la mamifere determină o postură fără discernământ sexual, indiferență față de situațiile riscante și pierderea simțului afectiv al stimulilor externi; stimularea sa electrică generează un comportament agresiv și violent;
- *Hipotalamusul* este situat în centrul creierului, se află chiar sub talamus și deasupra glandei pituitare, fiind considerat cea mai fundamentală parte a sistemului limbic, deoarece este responsabil de funcțiile vegetative ale creierului (menținerea echilibrului intern, temperatura corpului, foamea, setea, producția de hormoni etc.); în plus, este direct legat de comportament, jucând un rol esențial în emoții (partea centrală este legată de aversiune, iar părțile laterale de plăcere și furie);
- *Talamusul* este situat în diencefal, chiar sub corpul calos, în această parte a sistemului limbic sunt selectate informațiile senzoriale captate din mediul înconjurător, fiind responsabil și de partea motorie (activarea musculară); este legat și de stările de vigilență și este considerat

fundamental în percepția durerii, fiind important în reglarea stărilor de somn și de veghe;

- *Hipocampul* este situat în partea interioară a lobului temporal medial, este responsabil de memoria pe termen scurt și lung, permițând organismului să compare situațiile actuale cu experiențele trecute, sporindu-i astfel capacitatea de supraviețuire; este, de asemenea, responsabil de orientarea spațială și navigație.



Imaginea 25. Sistemul limbic: A = Corpus callosum; B = Talamus; C = Hipotalamus; D = Hipocampus; E = Amigdală.

Pornind de la această descriere neuroanatomică succintă, putem aprofunda acum ipotezele susținute de Joseph LeDoux. Acesta propune un nou cadru

conceptual care redefinește modul în care examinăm emoțiile, concentrându-se pe funcțiile și circuitele esențiale pentru supraviețuirea organismelor.

În această abordare, se renunță la metoda convențională care pornește de la experiențele conștiente ale oamenilor și încearcă să determine dacă aceleași emoții sunt prezente și la alte animale, fiind o strategie susceptibilă la antropomorfism. În schimb, întrebarea se inversează: în ce măsură funcțiile și circuitele fundamentale pentru supraviețuire la animale sunt prezente și la oameni?

Potrivit lui LeDoux, aceste circuite și funcții nu au o relație cauzală directă cu emoțiile, deși pot contribui indirect la acestea. Așadar, ce reprezintă exact circuitele de supraviețuire? În esență, acest concept mai larg cuprinde toate mecanismele relevante pentru menținerea vieții și succesul adaptativ al organismului în mediul său.

Această schimbare de perspectivă este justificată de un argument solid: LeDoux subliniază că termenul emoție este filosofic inconsistent și nu există un consens clar în comunitatea științifică asupra semnificației sale exacte. Prin urmare, pentru a evita capcana ambiguității conceptuale, el propune această inversare metodologică, mutând accentul de la emoții la circuitele fundamentale care guvernează comportamentul adaptativ.

Urmând această direcție, LeDoux își propune să studieze fenomenele relevante pentru înțelegerea emoțiilor, fără a apela însă la un limbaj bazat pe concepte emoționale tradiționale. În loc să discute despre „frica” unui organism, el analizează circuitele neuronale care îi permit să detecteze și să răspundă la amenințări. În loc să vorbească despre „pasiune” sau „entuziasm”, el examinează mecanismele care influențează motivația și comportamentul adaptativ. Emoțiile devin astfel etichetări secundare ale unor mecanisme fundamentale, nu categorii primare de analiză.

Aceste circuite reglează procese esențiale precum identificarea pericolelor, căutarea hranei și apei, apărarea și conservarea energiei, termoreglarea sau selecția partenerilor. Provocarea științifică constă în descrierea acestor mecanisme fără a recurge la un limbaj ambiguu, încărcat de semnificații subiective.

Această abordare permite evitarea dezbatelor interminabile legate de definirea exactă a conceptelor precum „tristețe”, „frică” sau „fericire” și oferă un cadru mai clar pentru investigarea emoțiilor. În loc să se concentreze pe interpretările subiective ale emoțiilor, această perspectivă le analizează prin prisma funcțiilor de supraviețuire care au fost conservate de-a lungul evoluției.

Ideea că emoțiile sunt legate de circuitele fundamentale de supraviețuire nu este nouă, ci își are originile în lucrările lui Charles Darwin (1872). Prin urmare, pentru neurocercetători nu este surprinzător faptul că aceste circuite sunt conectate la creier. Ceea ce diferențiază abordarea lui LeDoux este tocmai accentul pus pe aceste circuite ca bază pentru emoții. Aceasta reprezintă o schimbare semnificativă față de paradigma tradițională din neuroștiințe, care pornea de la definiții prestabilite ale emoțiilor și încerca ulterior să identifice circuitele cerebrale corespunzătoare. În schimb, LeDoux propune ca mai întâi să fie identificate circuitele, iar abia apoi să fie analizată relația lor cu emoțiile, evitând presupunerile apriorice despre natura acestora.

Un alt aspect esențial al cercetării emoțiilor, în care contribuțiile lui LeDoux sunt relevante, privește natura înnăscută sau dobândită a acestor circuite emoționale. Dezbaterea centrală este dacă așa-numitele „emoții de bază” sunt trăsături naturale, universale, sau dacă sunt produse ale interacțiunii dintre creier și influențele sociale și culturale.

Conceptul de emoții de bază se referă la emoțiile considerate universale, exprimate de toți indivizii și observabile la alte specii apropiate din punct de vedere evolutiv. Un exemplu clasic este lista propusă de Paul

Ekman (1972, 1992),⁷² care include fericirea, frica, tristețea, furia, surprinderea și dezgustul. Multe cercetări actuale încearcă să identifice substraturile neuronale ale acestor emoții.

Cu toate acestea, ideea de emoții de bază a generat numeroase controverse. Diferite teorii fie propun emoții suplimentare, fie elimină unele dintre cele incluse de Ekman. Acest lucru se datorează, în mare parte, dificultății de a corela un concept emoțional stabilit anterior cu diversitatea proceselor biologice care pot coincide sau nu cu acel concept.

În plus, această abordare ridică două provocări majore. În primul rând, unii cercetători susțin că emoțiile nu sunt fenomene strict biologice, ci construcții psihologice și sociale care rezultă din interacțiunea dintre individ și mediul său fizic sau social. Din această perspectivă, nu ar exista o determinare biologică rigidă a emoțiilor, ci acestea ar fi modelate cultural și contextual. În al doilea rând, majoritatea teoriilor despre emoțiile de bază sunt susținute de cercetări realizate pe creierul animalelor, însă acestea nu se aliniază întotdeauna cu emoțiile umane descrise de Ekman sau de alți teoreticieni. De

⁷² Ekman, P. (1972) "Universals and cultural differences in facial expression of emotion" In J. Cole (Ed.), *Nebraska Symposium on Motivation*, Lincoln, Nebraska: University of Nebraska Press: pp. 207-283. An interesting update can be found at: Ekman, Paul (1992). "An Argument for Basic Emotions", *Cognition and Emotion*, 6 (3/4): 169-200.

exemplu, lucrările lui Jaak Panksepp (1998, 2005) sugerează existența unor sisteme emoționale primare la mamifere, dar acestea nu se suprapun perfect peste clasificările clasice ale emoțiilor umane.⁷³



Imaginea 26: "Nosce te ipsum" © Gaia Amoroso

De exemplu, Lisa Feldman Barrett (2006)⁷⁴ consideră că este lipsit de sens să susținem că emoțiile de bază au

⁷³ Ediția originală: Panksepp, J. (1998) *Affective Neuroscience*, New York: Oxford U. Press. A relevant update can be found at: Panksepp J. (2005) "Affective consciousness: Core emotional feelings in animals and humans", *Consciousness and Cognition*, 14: 30–80.

⁷⁴ Ediția originală: Barrett, L. (2006) "Are Emotions Natural Kinds?", *Perspectives on Psychological Science*, 1:28–58.

fost conservate prin evoluție prin intermediul circuitelor neuronale din trei motive specifice:

- unele studii imagistice arată că zonele activate de același stimul au fost asociate cu emoții de bază diferite;
- o mare parte din dovezile în favoarea emoțiilor de bază au fost colectate în creierele animalelor prin tehnici retrograde (de exemplu, stimularea electrică a creierului), cu o lipsă evidentă de precizie;
- emoțiile de bază identificate în studiile pe oameni nu coincid cu emoțiile de bază identificate la animale.

Această controversă poate fi clarificată, încă o dată, prin abordarea lui LeDoux: în loc să discute despre "emoțiile de bază", oamenii de știință și filosofi ar trebui să favorizeze termenul de "circuite de supraviețuire". Această schimbare le-ar permite să abordeze și să rezolve confuziile conceptuale, care au implicații directe asupra cercetării empirice, cauzate de utilizarea în primă instanță a limbajului emoțiilor.

Aceste circuite de supraviețuire includ sisteme bazale pentru menținerea nutriției, echilibrul fluidelor, apărarea, reproducerea și termoreglarea. Se crede că aceste mecanisme fundamentale au fost prezente încă din cele mai timpurii forme de viață de pe Pământ. De exemplu, bacteriile demonstrează capacitatea de a

evita mediile toxice și de a căuta condiții mai favorabile dezvoltării lor, utilizând unele dintre aceste circuite.

Odată cu evoluția organismelor eucariote multicelulare, aceste circuite au devenit din ce în ce mai complexe, în special prin dezvoltarea receptorilor senzoriali specializați și a unui sistem nervos central capabil să coordoneze atât funcțiile corporale, cât și interacțiunile cu mediul.⁷⁵ Cu toate acestea, scopul lui LeDoux, atunci când descrie aceste circuite de supraviețuire, nu este de a le corela direct cu ceea ce numim „emoții de bază”. Intenția sa este, mai degrabă, de a separa limbajul neuroștiințific de conceptele bazate pe sentimentele subiective umane și de a se concentra, într-un mod neutru, asupra acestor circuite conservate de-a lungul evoluției.

Pentru a ilustra relevanța acestei abordări, putem analiza un exemplu legat de agresivitate. Știm că agresivitatea nu are o singură „reprezentare” fixă în creier, ci diferite tipuri de agresiune pot apărea în funcție de context. Aceasta poate fi declanșată de circuitul de apărare, în protejarea organismului, de circuitul alimentar, în competiția pentru hrană, sau de circuitul reproductiv, în competiția pentru parteneri.

Astfel, contrar multor studii din neuroștiințe, un circuit de supraviețuire nu este legat exclusiv de o anumită

⁷⁵ Shepherd, G. (1988) *Neurobiology II*, New York: Oxford.

emoție și nici de o anumită zonă neuronală specifică din creier. Dimpotrivă, mai multe circuite pot contribui la aceeași emoție, iar o emoție poate activa diferite circuite în funcție de circumstanțe. Aceasta sugerează că aceste circuite nu au o relație directă sau un rol causal asupra emoțiilor, ci mai degrabă funcționează ca mecanisme adaptative menite să asigure supraviețuirea organismului în interacțiunea sa cu mediul.

Această perspectivă propusă de LeDoux este esențială pentru înțelegerea emoțiilor umane. Prin studierea circuitelor de supraviețuire comune animalelor și oamenilor, și abia apoi analizând ceea ce numim „emoții”, ne apropiem mai mult de o explicație coerentă asupra naturii și funcției acestora.⁷⁶

După aceste două note introductive, cititorul este invitat să exploreze următoarele dialoguri cu patru intelectuali de renume internațional, fiecare oferind o perspectivă diferită asupra naturii creierului. Această incursiune va îmbogăți cu siguranță înțelegerea despre fascinanta călătorie care are loc în propria noastră minte.

⁷⁶ Capitol bazat pe: LeDoux, J. (1996) *The Emotional Brain*, New York: Simon and Schuster; LeDoux, J. (2003) "The Emotional Brain, Fear and the Amygdala", *Cellular and Molecular Neurobiology*, 23 (4/5): 727-738; LeDoux, J. (2012) "Rethinking the emotional brain", *Neuron*, 73 (4): 653-676.

DIALOGURI II

CREIERUL

DIALOGURI II

CREIERUL

VI. Dialogul cu Anil Seth



Anil Seth este profesor de neuroștiință cognitivă și computațională la Universitatea din Sussex, unde conduce Centrul Sackler pentru Știința Conștiinței. De asemenea, este co-director al Programului Institutului Canadian pentru Cercetări Avansate (CIFAR) privind creierul, mintea și conștiința, precum și al Programului de burse doctorale Leverhulme: „De la senzație și percepție la conștiință”.

În calitate de redactor-șef al revistei *Neuroscience of Consciousness* (Oxford University Press) și fost președinte al celei de-a 16-a conferințe a Asociației pentru Studiul Științific al Conștiinței (ASSC16, 2012), Seth este una dintre figurile de referință în cercetarea conștiinței. Activitatea sa științifică a fost finanțată de instituții de prestigiu, precum EPSRC (Leadership Fellowship), Consiliul European pentru Cercetare (ERC, Advanced Investigator Grant), Wellcome Trust și Canadian Institute for Advanced Research (CIFAR).

A publicat mai multe cărți, cea mai recentă fiind *Being You: A New Science of Consciousness* (Faber & Faber), considerată de *The Guardian* una dintre cele mai bune șapte cărți științifice din 2021. Mai multe informații despre activitatea sa pot fi găsite pe anilseth.com.

Sunteți unul dintre cei mai importanți neurocercetători care susțin cadrul teoretic al Procesării Predictive, o teorie cu o istorie îndelungată și implicații semnificative asupra înțelegerii creierului și minții umane. Când ați întâlnit pentru prima dată această idee conform căreia creierul este o mașină de predicție?

Anil Seth: Nu a fost nici pe departe ideea mea, dar mi s-a părut o perspectivă extrem de atrăgătoare asupra modului în care funcționează creierul. Această abordare nu doar că explică mecanismele generale ale creierului, dar mi s-a părut și foarte relevantă pentru subiectele care mă interesau în mod special la acea vreme și care continuă să mă fascineze și astăzi. Printre acestea se numără *corelațiile neuronale* ale conștiinței, relația dintre activitatea cerebrală și experiențele noastre conștiente, precum și modul în care putem înțelege emoțiile și sinele.

Am auzit prima dată despre conceptul de *codare predictivă* acum aproximativ douăzeci de ani, dar am început să lucrez în mod serios la acest subiect în urmă cu doisprezece sau treisprezece ani. Mi s-a părut un limbaj potrivit pentru a descrie și înțelege funcțiile creierului.

Un alt aspect care m-a atras la această teorie este versatilitatea sa: poate fi analizată la diferite niveluri. Se poate studia dintr-o perspectivă matematică riguroasă,

se poate explora la nivel computațional (prin scrierea unor algoritmi care să o implementeze), dar se poate aborda și conceptual sau filosofic. Această capacitate de a conecta discipline diverse o face extrem de valoroasă în peisajul actual al neuroștiințelor.

Aprofundând originile cadrului procesării predictive, puteți oferi o prezentare mai amplă a modului în care această perspectivă s-a dezvoltat? Cine sunt figurile-cheie care au propus inițial și au susținut ideea că creierul funcționează ca un agent activ în interacțiunea sa cu lumea?

Anil Seth: Platon! (râde) Depinde cât de departe dorim să ne întoarcem. Știm că, încă din filosofia greacă, exista ideea că ceea ce percepem drept realitate nu este neapărat ceea ce există cu adevărat. Un exemplu clasic este *Alegoria peșterii* a lui Platon, unde prizonierii, incapabili să vadă dincolo de umbrele de pe pereții peșterii, iau aceste proiecții drept realitatea însăși. Ei nu pot face diferența între iluzie și adevăr.

Un concept similar apare în tradiția filozofică hindusă sub numele de *Maya*, care subliniază ideea că percepția este doar o reprezentare a realității și nu realitatea însăși. Aceste concepte filozofice antice sunt strâns legate de principiile moderne ale procesării predictive.

Un alt moment important apare în secolul al XI-lea, când savantul arab Ibn al-Haytham a fost probabil primul care a conceput percepția ca un proces de inferență, susținând că ceea ce vedem este, de fapt, o interpretare a creierului, o estimare a ceea ce există în realitate.

În secolul al XIX-lea, Hermann von Helmholtz a dezvoltat această idee într-o formă mai detaliată, propunând că percepția este un proces inconștient de inferență realizat de creier. Această perspectivă s-a aliniat perfect cu principiile matematicii inferenței bayesiene, unde creierul face predicții despre lume și le ajustează pe baza noilor date senzoriale.

Abia în secolul XX, însă, ideea procesării predictive ca algoritm specific a fost dezvoltată inițial în inginerie și procesarea semnalelor. În anii 1950, cercetătorii căutau metode de compresie a semnalelor fără a pierde informații esențiale. Principiul era simplu: dacă poți prezice ceva, nu trebuie să îl transmiți explicit. Această idee a fost baza codificării predictive, utilizată ulterior în neuroștiință.

În jurul anului 1990, teoria a fost aplicată pentru prima dată sistemului vizual, într-o lucrare de referință publicată de Rajesh Rao și Dana Ballard. Aceștia au demonstrat că sistemul vizual uman ar putea funcționa conform principiilor procesării predictive. De atunci, această idee a fost extinsă și a devenit o teorie generală a funcționării creierului.

Sunteți cunoscut pentru critica dvs. asupra dihotomiei propuse de David Chalmers, care face distincția între „problema ușoară” și „problema dificilă” a conștiinței. Respingeți această diviziune și propuneți o abordare alternativă, denumită „problema reală” a conștiinței. Puteți detalia această schimbare conceptuală și modul în care facilitează o explorare științifică mai riguroasă a conștiinței?

Anil Seth: Am numit-o „problema reală a conștiinței” parțial ca o glumă prietenoasă la adresa lui David Chalmers, deoarece el a contribuit enorm la filosofia minții și cred că a formulat foarte clar distincția dintre problema dificilă și problema ușoară a conștiinței.

Totuși, cred că această distincție poate fi și un obstacol, deoarece întărește intuiția că nicio explicație a creierului sau a unui sistem fizic nu ar putea vreodată să clarifice de ce anumite părți ale universului sunt conștiente, iar altele nu. Este adevărat că conștiința nu pare genul de fenomen care ar putea fi explicat complet în termeni fizici, aceasta este o intuiție profundă. Iar, dintr-o anumită perspectivă, problema dificilă a conștiinței nu face decât să formalizeze această intuiție.

Dar efectul secundar al acestei abordări este că se ajunge la tratarea conștiinței ca pe un mister unic și absolut, ceea ce duce la căutarea unei soluții magice, poate printr-un mecanism cuantic enigmatic, care ar

trebui să transforme, ca prin miracol, un mecanism pur fizic în experiență subiectivă. Poate că într-adevăr avem nevoie de o nouă fizică sau de un moment "evrica", dar eu cred că există o altă abordare, ceea ce eu numesc abordarea problemei reale, care ar putea fi mai utilă.

Nu este o idee nouă, ci mai degrabă o recontextualizare în limbajul pe care Chalmers l-a popularizat. Problema reală a conștiinței pornește de la următoarea premisă: chiar dacă nu știm cum sistemele fizice generează conștiință sau dacă sunt identice cu aceasta, știm totuși că fenomenul conștiinței există și că este strâns legat, într-un mod sistematic, de creier.

Conștiința este un fapt al universului nostru și, ca orice alt fapt, are nevoie de o explicație. Problema reală susține că nu trebuie să abordăm direct întrebarea fundamentală (de ce există conștiința?), așa cum o face problema dificilă, ci că ar fi mai productiv să o descompunem în aspecte mai specifice: care sunt proprietățile conștiinței și cum le putem explica?

În cartea mea *Being You* (2022), folosesc o analogie pentru a ilustra această idee: cu ceva timp în urmă, oamenii credeau că viața nu poate fi explicată prin mecanisme fizico-chimice. Au petrecut mult timp dezbătând dacă există un elan vital, o scânteie a vieții, o forță mistică care face ca materia inertă să devină vie.

Însă, în final, nu s-a întâmplat nici ca oamenii să descopere existența scânteii vieții, dar nici să concluzioneze că viața nu există. În schimb, au realizat că viața nu este un fenomen singular, ci un ansamblu de multiple proprietăți, iar fiecare dintre acestea a putut fi explicată prin chimie și fizică.

Urmând aceeași strategie, problema dificilă a vieții nu a fost rezolvată, a fost dizolvată. Aceasta este abordarea pe care o consider mai utilă și pentru conștiință, chiar dacă, în final, nu va reuși să dizolve complet problema dificilă.

Dar singura modalitate de a afla dacă funcționează sau nu este să încercăm, iar abordarea reală exact asta face.

Această schimbare de perspectivă poate duce, în cele din urmă, la formularea unor probleme noi, care nu ar fi fost luate în considerare anterior. Prin dizolvarea distincției dintre problemele „ușoare” și „dificile” ale conștiinței, deschidem o cale de investigare care poate genera progrese semnificative. Înainte de această schimbare, dacă acceptăm premisa lui Chalmers că problema conștiinței este fundamental „dificilă”, posibilitățile pentru investigații empirice sunt sever limitate.

Anil Seth: Exact. Cred că aceasta este o observație foarte pertinentă și un aspect subapreciat al interacțiunii dintre știință și filosofie. Nu avem un „meniu” fix de întrebări pe care le rezolvăm, una câte una, până când ajungem la o înțelegere completă. Deși aceasta este adesea narațiunea prezentată publicului, progresul științific nu funcționează așa.

Este suficient să ne uităm la titlurile unor articole din ziare sau chiar din reviste științifice, unde vedem expresii de tipul „*Cele mai mari 10 mistere nerezolvate*”. Acest mod de a vedea lucrurile ignoră complet natura dialectică a progresului științific, unde întrebările, teoriile și datele evoluează în interacțiune constantă. Nu doar răspunsurile noastre se dezvoltă, ci și întrebările pe care le adresăm.

Și este posibil chiar să descoperim că întrebarea pe care încercam inițial să o rezolvăm nu avea, de fapt, sens, nu-i așa?

Anil Seth: Corect. Să revenim la analogia cu *élan vital* – dacă astăzi cineva ar încerca să obțină o finanțare pentru a căuta această „forță a vieții”, nu ar primi acel grant. Întrebarea în sine a devenit irelevantă, deoarece cadrul nostru conceptual s-a schimbat suficient de mult încât să înțelegem viața printr-un set diferit de principii biologice și chimice.

Vedem același tip de evoluție și în fizică, cu teoria eterului, un concept considerat esențial în secolele trecute pentru explicarea propagării luminii, dar care ulterior a fost complet abandonat odată cu dezvoltarea relativității.

Ce implică acest lucru? În esență, arată că aparatul conceptual pe care îl moștenim atunci când încercăm să investigăm o problemă ar trebui să fie întotdeauna pus sub semnul întrebării. Nicio întrebare științifică nu ar trebui preluată fără o analiză critică a presupunerilor care stau la baza sa. Ceea ce considerăm astăzi o „mare întrebare” s-ar putea dovedi, în timp, a fi doar o iluzie conceptuală care ne împiedică să vedem problema reală.

Credeți că abordarea procesării predictive a creierului și minții poate contribui la dezvoltarea inteligenței artificiale conștiente? Se vorbește mult despre *conștiința artificială*, dar pare dificil să ne așteptăm la progrese semnificative în această direcție fără a înțelege mai întâi conștiința umană.

Anil Seth: Pe scurt, cred că procesarea predictivă are deja un impact important asupra inteligenței artificiale, așa cum este înțeleasă în științele cognitive. După cum am menționat anterior, acest concept și-a avut originile în inginerie, ca metodă de compresie a semnalului.

În domeniul inteligenței artificiale, există deja modele care implementează aspecte ale procesării predictive. De exemplu, *mașinile Helmholtz*, numite după Hermann von Helmholtz, care a formulat ideea percepției ca inferență inconștientă, precum și *autocodificatoarele* și algoritmi de învățare profundă, care sunt structuri care aplică principiile acestei teorii.

Ingredientele fundamentale pentru procesarea predictivă și pentru IA sunt modelele generative, care pot anticipa intrările și le pot compara cu predicțiile anterioare. Pe baza acestor comparații, sistemul își ajustează predicțiile și își optimizează modelul de înțelegere a lumii. Cu un astfel de mecanism, percepția și învățarea pot fi realizate de același algoritm.

Un exemplu interesant este cercetarea lui Beren Millidge, un fost doctorand extrem de talentat, care a demonstrat că implementarea procesării predictive poate duce la *back-propagation*, algoritmul central utilizat în învățarea profundă a rețelelor neuronale artificiale.

Back-propagation a fost adesea criticat pentru că nu ar fi plauzibil din punct de vedere neurobiologic, deoarece necesită propagarea globală a semnalelor de eroare prin întreaga rețea neuronală, ceea ce nu pare a fi modul în care creierul uman funcționează. Însă codificarea predictivă pare să ofere un mecanism biologic realist pentru această propagare a erorilor.

Acesta este un indiciu puternic că procesarea predictivă nu doar explică învățarea artificială, ci poate reflecta și ceea ce face creierul uman.

În prezent, există o explozie de cercetări în IA care urmează această direcție. Nu este vorba doar de crearea unor rețele neuronale mai complexe, ci și de dezvoltarea unor modele generative capabile să anticipeze structurile și tiparele lumii. Vedem acest lucru în *rețelele generative adversative* (GANs), care sunt utilizate pe scară largă în diverse domenii, inclusiv în neuroștiință. Unele dintre aceste rețele sunt utile pentru înțelegerea creierului, în timp ce altele au scopuri mai aplicate, dar, în orice caz, există un teren de cercetare fertil în această intersecție dintre neuroștiință, procesarea predictivă și IA.

Ați publicat recent cartea *Being You: A New Science of Consciousness*, care a fost primită cu entuziasm și desemnată Cartea Anului la categoria știință de către *Financial Times* și *The Guardian*. Care sunt ideile principale pe care doriți să le evidențiați?

Anil Seth: Aș sublinia trei idei esențiale pe care le-am susținut în cartea mea.

În primul rând, conștiința poate fi abordată științific și filosofic fără a încerca să rezolvăm direct problema dificilă a conștiinței. Aceasta este abordarea problemei

reale, pe care am discutat-o anterior. De asemenea, această abordare se aliniază cu neurofenomenologia, propusă inițial de Francisco Varela, care a argumentat că putem studia conștiința corelând experiențele subiective cu mecanismele cerebrale.

A doua idee fundamentală este că procesarea predictivă poate explica nu doar cum organismele primesc și folosesc datele senzoriale pentru a ghida comportamentul, ci și modul în care predicțiile perceptuale se reflectă în experiențele noastre conștiente. Diferite tipuri de experiențe vizuale, de exemplu, pot fi înțelese ca fiind rezultatul unor tipare specifice de predicții perceptivă.

Al treilea și, probabil, cel mai important aspect al cărții este faptul că întreaga mașinărie predictivă se aplică și la sine. Aici se află punctul central al cărții: sinele nu este un agent care percepe realitatea și ia decizii, nu este un „mini-eu” în creier care analizează informațiile. În schimb, sinele este un construct format prin inferență perceptivă.

În cazul percepției obișnuite, semnalele senzoriale vin din mediul exterior. Însă în ceea ce privește sinele, semnalele vin în principal din interiorul corpului – de la ritmul cardiac, activitatea stomacului și alte procese fiziologice. Acest flux de informații interne este ceea ce numim interocepție, iar interpretarea acestor semnale de către creier creează experiența subiectivă a sinelui.

Marea teză a cărții este că toate fenomenele conștiinței sunt un produs al predicțiilor perceptuale. Experiențele noastre, inclusiv conștiința de sine, sunt fundamentate în procesele biologice ale corpului nostru. Conștiința nu este doar o proprietate a creierului, ci o expresie a corpului viu și a relației sale cu lumea.

Credeți că această perspectivă inovatoare ar putea contribui la o mai bună înțelegere a tulburărilor psihiatrice? Putem presupune că multe dintre aceste tulburări sunt rezultatul unui dezechilibru între predicțiile creierului și semnalele reale primite din mediu?

Anil Seth: Sper cu adevărat că da. Este un domeniu pe care l-am explorat intens la Centrul de Cercetare din Sussex în ultimii zece ani. Totuși, sunt precaut în fața promisiunilor exagerate. Deși această abordare pare promițătoare, drumul până la soluții aplicabile este încă lung.

În medicina psihiatrică, tratamentele actuale vizează în mare parte simptomele, fără a aborda cauzele subiacente ale tulburărilor. În multe cazuri, psihiatria acționează într-un mod similar cu administrarea de paracetamol pentru febră: simptomul este tratat, dar cauza profundă rămâne neabordată.

În alte domenii medicale, avem tratamente care vizează direct cauzele bolilor, de exemplu, antibioticele, care elimină infecțiile bacteriene. Însă, în psihiatrie, nu avem încă echivalentul unor „antibiotice psihiatrice” care să acționeze asupra mecanismelor fundamentale ale tulburărilor.

Privind tulburările psihiatrice prin prisma procesării predictive, putem începe să identificăm aceste mecanisme și să dezvoltăm intervenții mai specifice. Dar chiar și așa, suntem departe de a avea un tratament universal pentru bolile psihice.

Există semnale promițătoare în cercetare, dar nici măcar psihedelicele, care sunt uneori prezentate ca soluții revoluționare, nu pot oferi o rezolvare completă.

Dezbaterea despre liberul arbitru a devenit extrem de polarizată. Oameni de știință precum Robert Sapolsky susțin determinismul radical, negând existența liberului arbitru, în timp ce filosofi precum Daniel Dennett argumentează că avem un anumit grad de libertate. Care este poziția dumneavoastră în această dezbatere?

Anil Seth: Am scris un capitol despre acest subiect în cartea mea, iar redactarea lui a fost o adevărată provocare, deoarece era un aspect pe care nu îl mai abordasem anterior.

Conceptul de liber arbitru a început să aibă sens pentru mine atunci când l-am privit în același mod în care am analizat orice alt tip de experiență: ca o formă de predicție perceptuală.

Dar ce înseamnă acest lucru în contextul dezbaterii obișnuite? Există persoane precum Robert Sapolsky, care susțin cu fermitate că liberul arbitru nu există, și, de obicei, argumentează astfel din cauza unei suspiciuni că universul ar fi în întregime determinist, că traseul său este complet prestabilit. Acesta este, cu siguranță, punctul de vedere al lui Sapolsky. Însă astfel de argumente sunt adesea vinovate de construirea unui om de paie, adică combat o versiune a liberului arbitru care oricum nu merită apărută.

Această versiune problematică a liberului arbitru este ceea ce filosofii numesc liber arbitru libertarian. Eu sunt și mai puțin îngăduitor și îl numesc liber arbitru fantomatic, adică ideea conform căreia conștiința poate interveni brusc și modifica desfășurarea evenimentelor fizice, ca un fel de cauză necauzată cu o strategie proprie. Acest tip de liber arbitru nu are sens, dar nici nu este cel pe care ar trebui să ni-l dorim sau de care am avea nevoie.

Principala problemă în discuțiile despre liberul arbitru, din punctul meu de vedere, este că oamenii fac multe presupuneri fără să își dea seama că le fac. Întrebarea corectă pe care ar trebui să ne-o punem este: Este

creierul un sistem suficient de complex încât să poată reacționa la situații într-un mod flexibil, care nu este imediat determinat de mediul înconjurător? Iar răspunsul, în cazul organismelor ca noi, este: Da!

Noi putem realiza acțiuni ale căror cauze provin mai degrabă din interiorul corpului nostru și care își au originea adânc în istoria noastră personală, nu doar din lumea exterioară și din contextul imediat. Avem multiple grade de libertate, într-un sens strict ingineresc. Pot face multe lucruri, dar în final aleg să fac unul singur. Creierul nostru extrem de complex gestionează aceste grade de libertate, integrând numeroase cauze anterioare într-un rezultat comportamental unic, ceea ce unii numesc calea comună finală.

În plus, există momente în care, atunci când acționez, simt o experiență de voință și "agenție". Uneori, percep o acțiune ca fiind realizată în mod liber, alteori nu. Așadar, abordez această problemă în același mod în care studiez orice altă experiență perceptuală și mă întreb: „De ce această experiență se simte așa cum se simte?” și „Care este scopul acestei experiențe?”

Urmând această linie de gândire, putem înțelege experiența liberului arbitru asemănător experienței culorilor. Culoarea roșie, de exemplu, nu există în mod obiectiv ca proprietate a lumii externe, dar perceperea ei este extrem de utilă pentru noi, deoarece ne ajută să

urmărim anumite proprietăți importante ale obiectelor din jur.

Pentru mine, experiențele de liber arbitru funcționează în același mod: ele marchează un aspect esențial al modului în care organismele complexe se comportă. Aceste experiențe etichetează acțiunile care prezintă un anumit grad de libertate față de "imediatețe", folosind un termen propus de Mike Shadlen.

Aceasta este o caracteristică utilă nu pentru că experiențele de liber arbitru cauzează acțiunile, ci pentru că ele permit organismului să învețe și, data viitoare, dacă un rezultat nu a fost satisfăcător, să aleagă o altă acțiune.

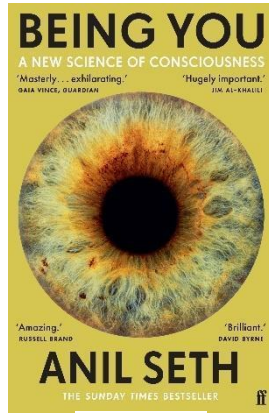
Greșeala este să credem că experiența liberului arbitru determină o acțiune, dar nu face acest lucru. În schimb, este utilă deoarece ne semnalează acele acțiuni care sunt mai degrabă cauzate de factori interni decât externi. Privit astfel, întregul conflict dintre determinism și stochasticitate devine în mare parte irelevant.

Putem considera liberul arbitru un fenomen biologic natural, dar el este real în același sens în care experiența culorilor este reală și este util în același fel în care percepția culorilor este utilă.

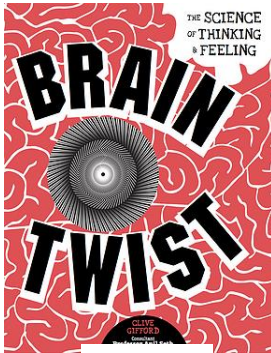
Privind liberul arbitru într-un mod mai nuanțat, ca ceva ce există în grade, putem analiza și modul în care variază acest fenomen. Desigur, acest lucru complică

discuțiile despre responsabilitatea morală și alte aspecte etice, dar această complexitate este inevitabilă.

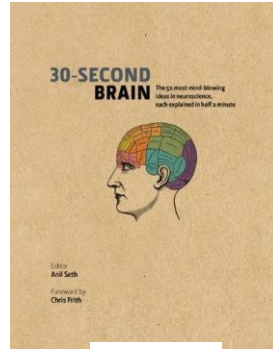
Cărți scrise de Anil Seth



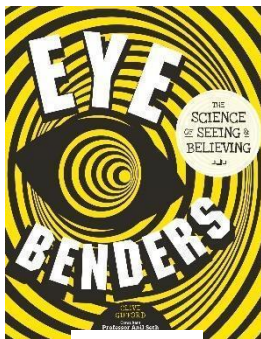
2021



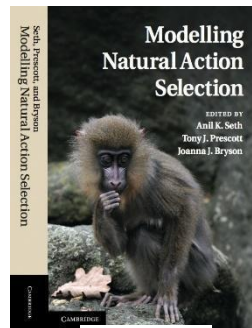
2015



2014



2013



2011

VII. Dialogul cu Karl Friston



Karl J. Friston este profesor la Institutul de Neurologie de la University College London, neurolog și o autoritate de renume mondial în imagistica creierului. Este inventatorul unor metode esențiale în analiza datelor cerebrale, precum cartografierea parametrică statistică, morfometria bazată pe voxel și modelarea cauzală dinamică (DCM).

De-a lungul carierei sale, a fost distins cu numeroase premii prestigioase, inclusiv Minerva Golden Brain Award, a fost ales Fellow of the Royal Society, iar în 2008 a primit Medalia College de France. În 2013, a fost laureat al premiului Weldon Memorial pentru contribuțiile sale la biologia matematică.

În 2016, a fost recompensat cu Premiul Charles Branch pentru contribuțiile sale excepționale în cercetarea creierului, precum și cu Premiul Glass Brain. Deține titluri onorifice de la Universitatea din Zurich și Universitatea Radboud.

Printre lucrările sale importante se numără Statistical Parametric Mapping: Analiza imaginilor funcționale ale creierului (2006), Principles of Brain Dynamics: Global State Interactions (2012) și, mai recent, Active Inference: The Free Energy Principle in Mind, Brain, and Behavior (2022). Mai multe informații despre activitatea sa pot fi găsite pe: <https://www.fil.ion.ucl.ac.uk/~karl/>

Sunteți considerat cea mai mare autoritate în ceea ce privește Principiul Energiei Libere (FEP) și implicațiile sale teoretice și empirice. Cum a apărut această idee în cercetările dumneavoastră?

Karl Friston: Putem susține că această idee există încă din vremea lui Platon, probabil fiind prezentă și în gândirea kantiană, iar mai târziu a fost formulată clar de Helmholtz prin conceptul de inferențe inconștiente.

O regăsim și în teoriile lui Richard Gregory, precum percepția ca testare a ipotezelor și formalizările aferente, dar și în lucrările lui Peter Dayan și Geoffrey Hinton, care au dezvoltat Helmholtz Machine, preluând noțiunea matematică de energie liberă de la Richard Feynman.

Feynman a folosit acest concept pentru a transforma o problemă de integrare inferențială imposibil de rezolvat într-o problemă de optimizare tratabilă, un aspect tehnic extrem de important, deoarece permite interpretarea creierului ca o mașină de optimizare, al cărei scop este optimizarea propriilor convingeri. Astfel, creierul poate fi privit ca o mașină inferențială, un organ statistic.

Aceste idei s-au dezvoltat de-a lungul timpului, dar au fost clar formalizate de oameni precum Richard Feynman și, în domeniul inteligenței artificiale și al învățării automate, de cercetători precum Geoffrey

Hinton și Peter Dayan. Totuși, generalizarea acestor idei către toate procesele cognitive necesită câțiva pași.

Primul pas ar fi extinderea acestui cadru conceptual nu doar la procesele de interpretare a informației, inferență, măsurare și observație, ci și la comportament și acțiune, pe care le putem înțelege tot ca o încercare de a minimiza energia liberă.

Dacă utilizăm machine learning ca un proces de optimizare a energiei libere (care, din perspectivă fizică, poate fi echivalentă cu energia liberă negativă), atunci, din perspectiva fizicienilor, vorbim de minimizarea energiei libere, reducerea erorii de predicție sau a surprizei.

În această logică, putem interpreta toate acțiunile noastre ca fiind ghidate de un mecanism de minimizare a surprizei. Anticipăm mișcarea, realizăm predicția că vom acționa, apoi transformăm acea predicție în acțiune efectivă, conform teoriei ideomotorii. Astfel, întregul proces poate fi înțeles ca reducerea aceleiași energii libere, erori de predicție sau surprize, pe care le folosim pentru a construi sensul realității.

Așadar, răspunsul meu, deși lung, este că nu am fost cel care a avut această idee inițial. Mai degrabă, am preluat-o și am contribuit treptat la formalizarea și cristalizarea ei de-a lungul timpului.

Conform Principiului Energiei Libere, organismele vii încearcă să minimizeze entropia energiei lor libere. Totuși, stările modificate de conștiință, cum ar fi cele induse de droguri sau de anumite afecțiuni psihiatrice, par să implice o entropie anormal de ridicată sau de scăzută. Credeți că aceste cazuri ridică probleme pentru teoria dumneavoastră?

Karl Friston: Nu consider că acest lucru este deloc problematic. Dimpotrivă, cred că astfel de fenomene reprezintă perspective extrem de valoroase și informative asupra modului în care creierul încearcă să înțeleagă lumea, făcând acest lucru printr-un proces activ de percepție. Atât psihopatologia, cât și stările de conștiință alterate de substanțe psihedelice, de exemplu, dezvăluie aspecte mecanistice esențiale ale acestei perspective.

Ați menționat ideea de entropie. Este important de reținut că entropia nu este altceva decât o descriere a unei distribuții de probabilitate, motiv pentru care trebuie să precizăm clar despre ce tip de distribuție vorbim. În aplicarea principiului energiei libere la auto-organizare – indiferent dacă ne referim la particule elementare sau la ființe umane –, entropia pe care o încercăm să minimizăm este entropia rezultatelor, adică a experiențelor pe care le trăim efectiv. Aceasta se aliniază cu perspectiva homeostatică, conform căreia încercăm să menținem interacțiunile noastre cu mediul într-un interval viabil.

Pe de altă parte, pentru a minimiza această entropie a rezultatelor, trebuie să maximizăm entropia convingerilor noastre, conform principiului lui Occam: căutăm explicații cât mai simple, fără a ne angaja prea rigid într-o singură interpretare. Aici se conturează un joc dialectic între realizarea unei inferențe cât mai eficiente – ceea ce presupune maximizarea entropiei credințelor – și consecința acestui proces, adică minimizarea entropiei fizice a rezultatelor noastre.

Aceași logică se aplică și entropiei activității neuronale, care codifică convingerile despre cauzele senzațiilor noastre. Pe măsură ce maximizăm entropia convingerilor, simultan minimizăm entropia dinamicii neuronale, făcând-o cât mai simplă și mai eficientă posibil.

Aș spune că trebuie să fim precauți atunci când discutăm despre entropie, deoarece există mai multe tipuri: entropia activității generale, entropia credințelor, entropia observațiilor. Cu toate acestea, fiecare dintre aceste tipuri de entropie este un instrument extrem de util pentru a înțelege diferite stări ale conștiinței, așa cum ați subliniat.

Dacă ne uităm la istoria psihopatologiei, la efectele drogurilor psihoactive și chiar la modificările fiziologice ale stării de conștiință – cum ar fi somnul sau visul –, observăm un aspect extrem de interesant al procesului de optimizare a energiei libere. Uneori, în teoria

codificării predictive, dacă facem anumite simplificări, putem interpreta energia liberă în sens fizic ca o eroare de predicție, mai precis, o eroare de predicție ponderată cu precizia sa.

În acest caz, nu încercăm să minimizăm toate erorile de predicție, ci doar pe cele care sunt considerate foarte precise, informative și de încredere. Dacă ne gândim la creier ca fiind ghidat de imperativul de a minimiza erorile de predicție ponderate cu precizie, realizăm că acesta nu doar că trebuie să facă predicții despre conținutul observațiilor, ci și despre precizia acestora.

Mai mult decât atât, creierul trebuie să cuantifice incertitudinea. Din punctul de vedere al unui fiziolog, aceasta înseamnă că trebuie să estimeze excitabilitatea diverselor structuri neuronale care raportează erorile de predicție către nivelurile superioare, responsabile pentru acumularea dovezilor necesare actualizării convingerilor. Pe scurt, acest lucru înseamnă că modularea excitabilității acestor structuri neuronale, care raportează erorile de predicție și energia liberă, trebuie ea însăși să fie precisă. Dacă acest proces dă greș, rezultatul va fi formarea unor convingeri și inferențe distorsionate.

Acest aspect este crucial din două motive. În primul rând, face legătura cu dimensiunea psihologică a ponderării deciziilor, adică atenția. Acordarea atenției unui stimul înseamnă, fundamental, că îl tratăm ca fiind

relevant pentru actualizarea convingerilor noastre. Procesul ar putea fi descris astfel: „Aceasta nu este o informație nouă, are o precizie scăzută, o ignor – este o știre falsă. În schimb, acest alt lucru este interesant, pare important, deci îi voi crește câștigul atențional și îl voi folosi pentru actualizarea credințelor mele.”

Din perspectiva farmacologiei, substanțele care afectează acest mecanism sunt exact aceleași care influențează excitabilitatea neuronală și ponderarea preciziei. Astfel, putem înțelege mai bine efectele psihedelice precum psilocibina sau LSD asupra percepției vizuale și mecanismele care duc la inferențe perceptuale anormale sub influența acestor substanțe.

Dar nu este nevoie să luăm psihedelize pentru a observa aceste efecte. Aceleași disfuncții de neuromodulare pot fi identificate ca fenotipuri ale unor condiții psihiatrice, conducând la inferențe eronate. Aceste inferențe se manifestă sub forma a două tipuri clasice de erori: eroarea de tip I, inferarea existenței unui stimul care, de fapt, nu este prezent (de exemplu, o halucinație) sau eroarea de tip II, eșecul de a detecta un stimul real (de exemplu, un sindrom disociativ sau un sindrom de neglijare). Ambele tipuri de erori sunt caracteristice multor tulburări neurologice și psihiatrice.

Teoria dumneavoastră pare să ofere o explicație atât pentru creierul „normal”, cât și pentru creierul „anormal”, incluzând stările modificate de conștiință și chiar visele lucide. Considerați această versatilitate a modelului dumneavoastră ca un semn că sunteți pe calea cea bună?

Karl Friston: Da, visul lucid este un exemplu fascinant.

Ați publicat o lucrare în 2018 despre visele lucide, din perspectiva inferenței active și a energiei libere, corect?

Karl Friston: Da, am lucrat cu Alan Hobson, un coleg și prieten apropiat, care a fost profund interesat de visele lucide. Hobson era convins că, dacă vrem dovezi că creierul este un organ constructiv, care generează ipoteze despre cum funcționează lumea, nu trebuie să căutăm mai departe de fenomenul visării.

Când visăm, avem o întrerupere neurochimică și neuromodulatorie a preciziei tuturor senzațiilor noastre, cu excepția celor provenite de la ochi, deoarece ochii fac parte din sistemul nervos central. Cu toate acestea, continuăm să percepem, să vedem și să experimentăm o realitate internă, chiar și în absența stimulilor externi.

Acest lucru demonstrează că creierul este un sistem activ, creativ și constructiv, care generează modele predictive chiar și în absența intrărilor senzoriale convenționale. De fapt, visul lucid ne arată cum creierul echilibrează ponderarea preciziei predicțiilor sale și controlul câștigului atențional, adaptând constant modelele interne la realitatea subiectivă a experienței.

Am o provocare filosofică pentru dumneavoastră. În unele lucrări, pare că susțineți o perspectivă întrupată sau chiar anti-reprezentationalistă. Însă, în alte lucrări, păreți să argumentați în favoarea unei viziuni reprezentationaliste. Care este, de fapt, poziția dumneavoastră în ceea ce este cunoscut drept „războiul reprezentărilor” în Știința Cognitivă și Filosofia Minții? Sunteți, în general, în favoarea sau împotriva reprezentărilor?

Karl Friston: Depinde cu cine vorbesc. Am auzit recent o replică amuzantă într-un sitcom australian despre a fi „un om care vrea să mulțumească pe toată lumea”.

Așadar, dacă vorbesc cu un sceptic, cineva care este anti-realist, atunci principiul energiei libere poate fi folosit pentru a spune: „Da, exact! Ai dreptate!”. De ce? Pentru că nu vei fi niciodată expus direct la realitate, dacă presupunem că există o realitate exterioară. Tot ceea ce ai este propriul tău sensorium. Practic, ești un

creier într-o cutie, iar tot ceea ce crezi despre lume este doar o fantezie construită pe baza percepțiilor tale. Această interpretare este perfect compatibilă cu abordarea sceptică.

Pe de altă parte, dacă vorbesc cu un realist sau un externalist, atunci inferența activă și aplicarea principiului energiei libere susțin acest punct de vedere, deoarece se bazează pe ideea de senzorialitate activă întrupată: este vorba despre modul în care interacționez fizic cu lumea mea pentru a obține exact tipul de senzații care fac ca sistemul meu intern de credințe să fie cât mai coerent cu acea lume. Astfel, self-evidencing poate fi citit literalmente ca fiind procesul de colectare a dovezilor din lumea reală pentru un model intern (I-model) al acesteia. Dacă eu sunt acel model, atunci îmi confirm existența prin colectarea dovezilor despre propria mea prezență. Acesta este, de altfel, un mod de a privi imperativele existențiale care sunt implicate în acest tip de auto-confirmare.

Ați menționat, însă, un concept interesant: reprezentationalismul. Dacă luăm în considerare principiul energiei libere, atunci, indiferent cu cine vorbesc, pot spune cu certitudine că acesta este, în esență, reprezentationalist. Absolut. Întreaga idee a separării creierului de lume este tocmai faptul că astfel lumea poate avea „credințe” despre lume, iar acestea nu sunt altceva decât reprezentări probabilistice.

Așadar, principiul energiei libere și inferența activă nu pot accepta enactivismul radical. Însă, dacă vorbim despre o formă mai soft de enactivism, atunci teoria poate, cu siguranță, să o integreze fără probleme.

Totuși, acest model nu acceptă enactivismul radical, adică ideea că percepția și cunoașterea nu implică reprezentări interne, ci sunt doar interacțiuni directe cu mediul. Dar dacă ne referim la un enactivism mai moderat, atunci da, “Principiul energiei libere” poate acomoda astfel de idei.

Cum poate “Principiul energiei libere” să ne ajute să înțelegem calitățile subiective ale experienței, așa-numitele qualia? Este abordarea dumneavoastră relevantă pentru această dezbatere filosofică?

Karl Friston: Aceasta este o întrebare extrem de interesantă și provocatoare, care a devenit un punct de interes recent pentru mai mulți dintre colegii mei care se ocupă de fenomenologia subiectivității.

Cea mai bună sinteză pe care o pot oferi în ceea ce privește gândirea actuală din domeniul filosofiei minții, în raport cu principiul energiei libere și inferența activă, este că pentru a avea o experiență calitativă, aceasta trebuie să fie o acțiune ascunsă (covert action). Cu alte cuvinte, trebuie să existe o acțiune internă.

Dacă aş privi din perspectiva unui psiholog, acest lucru ar fi echivalent cu ideea că trebuie să fiu capabil să îmi îndrept atenția asupra unui lucru. Dacă nu pot alege dacă să acord atenție sau nu, atunci nu pot avea o experiență calitativă a aceluia lucru. Acest aspect este adesea legat de conceptele lui Metzinger despre transparența și opacitatea fenomenală: pentru ca un fenomen să fie perceput ca fiind normal opac, iar aici interpretez opacitatea ca fiind echivalentă cu o experiență perceptuală, este necesar să avem capacitatea de a modifica precizia dovezilor senzoriale relevante, ceea ce, în esență, nu este altceva decât eroarea de predicție.

Dacă am privi această problemă din perspectiva unui neuroanatomist, a avea experiențe calitative (sau cel puțin acele tipuri de percepții care pot deveni opace) înseamnă că ar trebui să putem identifica proiecții neuronale anatomice în creier care au un efect neuromodulator asupra anumitor regiuni dintr-o ierarhie neuronală.

Din nou, din perspectiva unui psiholog, m-aș aștepta ca aceste proiecții să susțină capacitatea de a ne concentra endogen asupra diverselor aspecte ale sensoriumului. Astfel, ajungem la o explicație care leagă calitatea experienței de atenție, de selecția activă a surselor de informație relevante pentru actualizarea convingerilor și a reprezentărilor interne. Pentru mine, selecția activă nu este neapărat un comportament observabil.

Există, desigur, paralele cu teoria premotorie a atenției a lui Rizzolatti, în sensul că acțiunea asupra mediului pentru a obține informații precise implică uneori un comportament observabil, cum ar fi explorarea vizuală a lumii. Însă, dacă internalizăm acest proces și îl gândim în termeni de o „palpare” internă a sensoriumului, prin selectarea anumitor surse de erori de predicție sau dovezi senzoriale și ajustarea preciziei acestora, atunci vorbim de o formă de acțiune mentală, de o acțiune ascunsă. Aceasta ar putea fi considerată un test fundamental pentru calitatea experienței sau, altfel spus, pentru a interpreta qualia ca experiență calitativă.

Dacă acceptăm această perspectivă, există câteva condiții suplimentare necesare pentru a putea desfășura astfel de acțiuni mentale: este nevoie de un model generativ capabil să anticipeze precizia folosită ca bază pentru acțiunea mentală, ceea ce implică automat existența unui model generativ al viitorului. Acest lucru sugerează că doar entitățile suficient de sofisticate, care posedă modele generative ale viitorului, pot avea experiențe calitative, deoarece trebuie să poată anticipa consecințele acțiunilor lor, chiar și atunci când acestea sunt interne și ascunse.

Acest aspect este esențial. Există o dimensiune temporală inerentă acestei explicații mecaniciste a experienței calitative. Alte argumente, avansate recent de cercetători precum Maxwell Ramstead, Mark Solms și alții, explorează o formulare mai abstractă a senzației

conștiente, analizând experiența ca o proiecție pe un „ecran Markovian”, din perspectiva teoriei clasice a informației, sau ca un „ecran holografic”, din perspectiva teoriei informației cuantice.

Ideea de bază este că în creier ar putea exista un sistem distribuit fizic, delimitat de un fel de „pătură Markoviană”, sau conform teoriei informației cuantice, un „ecran holografic” care conține informația clasică, la care structurile interne ale acestui sistem pot scrie sau din care pot citi. Această delimitare ar separa structura internă de restul creierului. Există argumente care susțin că, pentru a avea conștiință și experiență conștientă, este necesară existența acestui ecran Markovian sau holografic, ceea ce, în mod interesant, ne aduce înapoi la ideea unui „Teatru Cartesian”.

Totuși, această abordare nu implică un dualism esențial, dar readuce în discuție, într-un mod fascinant, ideea unui ecran intern. Esențial este faptul că stările interne care „observă” acest ecran nu se pot vedea pe ele însele, deoarece nu pot exista fluxuri interne suplimentare care să reflecte direct aceste procese. În cele din urmă, întreaga explicație poate fi redusă la relația dintre atenție, acțiunea internă și experiența mentală, deoarece singura modalitate prin care interiorul poate modifica proiecțiile de pe acest „ecran intern” este prin acțiunea mentală. La finalul zilei, totul se reduce la această noțiune de acțiune mentală.

Neurocercetătoarea Grace Lindsay, în cartea sa *“The model of the mind”*, argumentează că “Principiul energiei libere” nu este falsificabil din punct de vedere științific. Susține că această teorie nu permite verificarea empirică și nu poate fi testată în mod concludent. Care este poziția dumneavoastră față de această critică?

Karl Friston: Aceasta este o observație corectă, dar trebuie înțeleasă în context. “Principiul energiei libere” este un principiu al acțiunii minime, similar cu “Principiul lui Hamilton” din fizică. Acest tip de principii nu sunt teorii sau ipoteze care pot fi falsificate direct, ci instrumente pentru formularea și testarea altor modele.

De exemplu, în matematică și informatică, avem concepte precum metodele de optimizare, care nu sunt teorii în sine, ci moduri prin care putem analiza și simula fenomene.

Astfel, “Principiul energiei libere” este o metodologie, un cadru formal care poate fi aplicat în diverse contexte pentru a testa alte ipoteze.

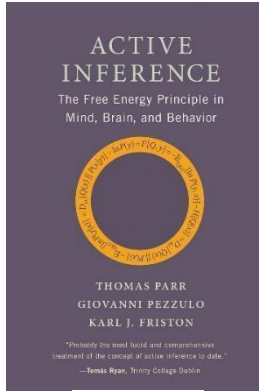
Dacă luăm un model generativ bazat pe acestuia și îl aplicăm unui anumit fenomen, de exemplu, percepția vizuală sau tulburările psihice, atunci acel model este testabil și falsificabil.

Dar "Principiul energiei libere" în sine este doar un principiu, este ca un T-Test, nu poate fi falsificat, îl folosiți pentru a face falsificarea. Așadar, nu principiul în sine trebuie să fie falsificabil, ci modelele specifice derivate din el. Aceasta este o distincție esențială pe care trebuie să o înțelegem.

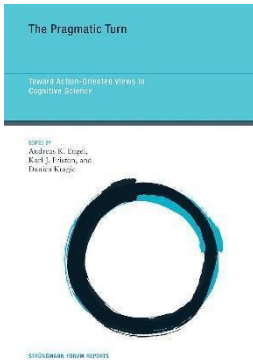
Într-un fel, putem spune că "Principiul energiei libere" oferă un cadru formal pentru înțelegerea falsificabilității deoarece el ne ajută să comparăm dovezile pentru o ipoteză nulă versus o ipoteză alternativă.

Astfel, critica este validă doar dacă încercăm să tratăm "Principiul energiei libere" ca pe o teorie direct testabilă, ceea ce nu este cazul. Este un instrument matematic, nu o teorie științifică în sine

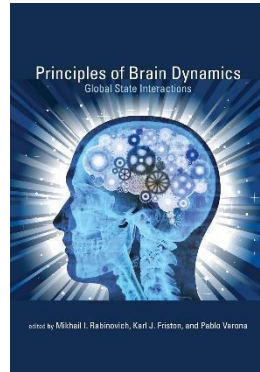
Cărți scrise de Karl Friston



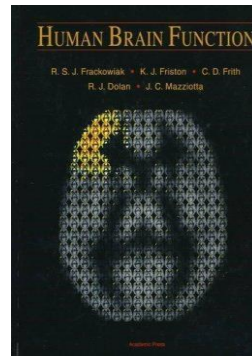
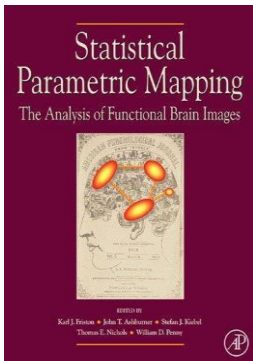
2022



2016



2012



VIII. Dialogul cu Christof Koch



Christof Koch este un neurofiziolog și specialist în neuroștiințe computaționale, cunoscut mai ales pentru lucrările sale privind bazele neuronale ale conștiinței.

Este președinte și cercetător-șef al Allen Institute for Brain Science din Seattle și cercetător-șef al Fundației Tiny Blue Dot.

Din 1986 până în 2013, a fost profesor la California Institute of Technology. Principalul colaborator al lui Koch în efortul de a localiza corelațiile neuronale ale conștiinței a fost Francis Crick, laureat al Premiului Nobel pentru Medicină.

Mai recent, Koch a colaborat îndeaproape cu psihiatrul și neurocercetătorul Giulio Tononi și a dezvoltat Teoria Informației Integrate a conștiinței.

Koch este autorul mai multor cărți, cum ar fi *The Quest for Consciousness: a Neurobiological Approach* (2004), *Consciousness: Confessions of a Romantic Reductionist* (2012) și *The Feeling of Life Itself - Why Consciousness is Widespread but Can't be Computed* (2019).

Mai multe informații: <https://christofkoch.com/>

Dată fiind complexitatea problemei conștiinței, credeți că paradigmele științifice actuale sunt suficiente pentru a o explica, sau este necesară o reformulare a fizicalismului pentru a face față acestui fenomen?

Christof Koch: Depinde ce înțelegeți prin fizicalism. Eu susțin o reformulare a *Teoriei Informației Integrate* (IIT), care afirmă că nu ne putem concentra doar pe puterea cauzală extrinsecă descrisă de fizica convențională, ci trebuie să studiem și puterea cauzală intrinsecă. Aceasta este conștiința.

Cu alte cuvinte, nu cred că avem nevoie de o nouă teorie a mecanicii cuantice integrate cu gravitația, așa cum sugerează Sir Roger Penrose. Chiar dacă vom avea, într-o zi, o teorie completă a fizicii fundamentale, nu cred că aceasta va schimba modul în care înțelegem conștiința. Creierul funcționează la 300 de grade Kelvin și nu operează la scările la care efectele cuantice ar putea deveni relevante pentru conștiință. Așadar, chiar dacă fizica se va schimba în viitor, nu cred că această schimbare va avea un impact semnificativ asupra modului în care explicăm conștiința. S-ar putea să mă înșel, dar aceasta este intuiția mea puternică.

Unul dintre aspectele controversate ale Teoriei Informației Integrate (IIT) este principiul excluderii.

Unii critici îl consideră un mecanism arbitrar, introdus ad-hoc. Credeți că acest principiu este esențial pentru IIT sau necesită o rafinare suplimentară?

Christof Koch: Nu sunt deloc de acord cu această critică. Principiul excluderii este fundamental pentru IIT. Dacă nu l-am avea, am ajunge la o multiplicitate de experiențe conștiente simultane și, implicit, la o multiplicitate de puteri cauzale, ceea ce este problematic.

Acest principiu nu a fost adăugat ulterior ca un artificiu teoretic, ci a fost mereu parte din fundamentul teoriei. IIT susține că, având în vedere toate combinațiile posibile de mecanisme dintr-un sistem, există un singur sistem conștient la un moment dat – acela care maximizează puterea cauzală intrinsecă.

Cu alte cuvinte, există o singură conștiință unitară și nu o infinitate de conștiințe fragmentate în creierul meu. Această conștiință este definită de granularitatea sa specifică adică analiza sa nu este nici la un nivel prea fin, nici la unul prea grosier. Dacă am avea mai mulți neuroni sau mai puțini neuroni, puterea cauzală a conștiinței ar fi mai mică.

Dacă ar fi să eliminăm acest principiu, întreaga teorie ar colapsa, pentru că ar deveni imposibil să stabilim o singură conștiință coerentă într-un sistem complex.

Desigur, teoria IIT ar putea fi greșită, dar nu putem elimina în mod arbitrar un principiu fundamental și să ne așteptăm ca teoria să supraviețuiască

Recent, a existat o controversă în studiul științific al conștiinței în care mai mulți oameni de știință și filosofi au semnat o scrisoare publică⁷⁷ în care susțineau că IIT este o teorie pseudoștiințifică, care nu poate fi falsificată. Ce fel de experimente credeți că ar putea falsifica teoria IIT?

Christof Koch: Multe dintre ele, de fapt. De exemplu, teoria face teste foarte specifice. De exemplu, vă puteți uita la o anumită rețea și vă puteți întreba dacă aceasta maximizează puterea cauzală intrinsecă. Este o predicție foarte specifică. Odată ce aveți matricea de probabilitate de tranziție a acelei rețele neuronale particulare, în principiu, o puteți calcula.

Și vă puteți întreba: maximizează puterea? Cum putem explica rezultatele din punct de vedere fenomenologic? Mă refer pe scurt la acest aspect în lucrarea despre fenomenologia spațiului;⁷⁸ sau cum rămâne cu

⁷⁷ Disponibilă aici: <https://osf.io/preprints/psyarxiv/zsr78>.

⁷⁸ Publicată aici:

<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0268577>.

fenomenologia timpului care curge? cum rămâne cu fenomenologia culorii?

În prezent, există o colaborare activă⁷⁹ pentru testarea predicțiilor între IIT și GNW. Un exemplu concret de testare implică compararea IIT cu Teoria Global Neural Workspace (GNW). GNW susține că substratul conștiinței se află în regiunile prefrontal-parietale. IIT afirmă că substratul conștiinței se află în partea posterioară a creierului. Această predicție este testată în prezent și poate fi verificată experimental.

Un alt exemplu este metoda de testare a complexității perturbării, utilizată în context clinic pentru a evalua dacă un pacient este conștient. Această metodă derivă direct din Teoria Informației Integrate (IIT) și este deja aplicată în studii clinice. Conform IIT, conștiința există atâta timp cât există experiență: dacă o persoană trăiește o experiență timp de 10 secunde, trebuie să existe un substrat fizic corespunzător acelei experiențe pe întreaga durată. În schimb, modelul *Global Neural Workspace* susține că experiența conștientă apare doar în momentul inițial al transmiterii informației, după care dispare. Aceste perspective divergente generează o gamă largă de teste experimentale, multe dintre ele fiind în desfășurare în prezent.

⁷⁹ Ediția originală: Melloni L, et al. (2023) "An adversarial collaboration protocol for testing contrasting predictions of global neuronal workspace and integrated information theory", *PLOS ONE*, 18 (2): e0268577.

Acesta este un semn bun, nu-i așa? Deoarece multe teorii ale conștiinței nu pot fi testate, atunci acesta mi se pare a fi un avantaj al IIT față de alte teorii ale conștiinței.

Christof Koch: Corect. Aș spune că este esențial.

În ceea ce privește inteligența artificială, susțineți că calculatoarele, cel puțin cele bazate pe arhitectura Von Neumann, nu pot fi niciodată conștiente. Cu toate acestea, dacă puterile cauzale intrinseci ale creierului ar fi reproduse pe un substrat artificial din siliciu, ar fi posibil să presupunem că aceste sisteme ar putea deveni conștiente?

Christof Koch: Pare plauzibil. De obicei, spun că nu există nimic magic în structura creierului – acesta nu are nimic supranatural. Un tranzistor tipic poate „comunica” cu alte două sau trei porți logice, în timp ce un neuron interacționează, în medie, cu aproximativ 50.000 de alți neuroni. Mai mult, între acești neuroni există o suprapunere semnificativă: dacă doi neuroni din proximitate proiectează fiecare conexiuni către 50.000 de alți neuroni, mulți dintre aceștia vor fi comuni.

Această conectivitate extinsă determină, în cele din urmă, puterea cauzală a sistemului nervos. Dacă un

astfel de model ar putea fi reprodus într-un hardware adecvat, indiferent de natura acestuia, atunci conștiința ar putea fi obținută. Totuși, trebuie să fim foarte atenți: inteligența și conștiința sunt două lucruri fundamentale diferite.

Nu există nicio îndoială că vom obține inteligență artificială generală – deoarece, în esență, inteligența presupune acțiune. Conștiința, pe de altă parte, nu este despre „a face”, ci despre „a fi”. Inteligența este orientată spre performanță și execuție, în timp ce conștiința reprezintă o stare de existență. Putem dota mașinile cu inteligență la nivel uman sau chiar superior, dar acest lucru nu implică în mod necesar că ele vor fi conștiente.

Ați menționat relația dintre inteligența artificială și conștiință, afirmând că dezvoltarea inteligenței artificiale este inevitabilă, dar nu știm dacă aceasta va fi sau nu conștientă. Totuși, este posibilă existența inteligenței fără conștiință? Nu vine inteligența după conștiință? Cum poate exista inteligența artificială fără a fi conștientă?

Christof Koch: O întrebare excelentă. În prezent, singurele două exemple cunoscute de inteligență flexibilă sunt cele umane și, într-adevăr, noi suntem conștienți. Totuși, mulți dintre noi credem că aceasta

este doar una dintre multiplele căi prin care poate apărea inteligența. Evoluția a ales acest traseu biologic, dar există și alte posibilități, poate prin siliciu și software, deși, până acum, nu avem nicio dovadă convingătoare în acest sens.

Pentru a pune problema altfel, îmi pot imagina fără dificultate computere extrem de puternice, mașini Turing universale, care ar putea fi la fel de inteligente sau chiar mai inteligente decât mine, cu siguranță mult mai rapide. Nu văd niciun motiv a priori pentru care conștiința ar fi necesară pentru asta.

Dacă ne uităm la progresele în învățarea profundă (*deep learning*), la rețelele generative și la arhitecturile bazate pe transformatoare, observăm că acestea se adaptează remarcabil de bine. Cu cât scala lor crește, cu atât performanța lor devine mai bună – iar conștiința nu pare să joace niciun rol în acest proces.

Desigur, este posibil ca aceste sisteme să atingă anumite limite în viitor. Însă, în acest moment, presupunerea mea este că putem atinge inteligența artificială generală, și chiar o formă de superinteligență, fără ca aceasta să implice apariția conștiinței.

Evoluția a urmat un drum diferit, dar noi, ca oameni, avem acum capacitatea de a ne abate de la acest model. Și, într-adevăr, exact asta facem. Dacă acest lucru este benefic pentru Homo sapiens-sapiens este o altă

întrebare. Sunt profund sceptic în această privință. Totuși, aceasta este direcția în care ne îndreptăm, fie că ne place, fie că nu.

Există mai multe metode pentru a studia creierul conștient. Ne putem concentra pe tulburările de conștiință, pe efectele psihedelicilor sau pe stimularea creierului, printre altele. Care dintre aceste abordări vi se pare cea mai promițătoare pentru înțelegerea naturii conștiinței?

Christof Koch: Aș spune că toate sunt valoroase. Nu există o metodă universal superioară, deoarece alegerea depinde de formarea și expertiza cercetătorului. Fiecare metodă are avantaje și limitări. Unele sunt mai obiective, altele permit o investigare mai directă a experienței subiective.

De exemplu, în studiile pe oameni, putem întreba participanții despre stările lor de conștiință, dar nu putem interveni direct în creierul lor așa cum o facem în studiile pe animale. Pe de altă parte, în cazul cercetărilor pe animale, putem manipula activitatea cerebrală, dar nu avem acces direct la experiențele lor subiective. Fiecare tehnică vine, așadar, cu propriile constrângeri.

Alegerea metodei depinde și de obiectivele cercetării. Psihedelicele, de exemplu, sunt utile deoarece pot

induce modificări profunde ale conștiinței. Totuși, există și dezavantaje deoarece multe dintre aceste substanțe sunt ilegale, iar cunoștințele noastre despre ele sunt încă limitate. Au fost realizate numeroase studii pe șoareci și șobolani, dar este dificil să determinăm cu certitudine ce experimentează aceștia sub influența psilocibinei sau a altor compuși psihedelici.

În concluzie, alegerea metodei depinde în mare măsură de formarea cercetătorului și de întrebările specifice pe care și le propune să le exploreze.

În contextul teoriilor procesării predictive, există o dezbateră continuă privind influențele „de sus în jos” asupra creierului. Aceste procese implică funcții cognitive de nivel superior care influențează percepțiile senzoriale de nivel inferior. Cum vedeți relevanța acestor influențe în înțelegerea conștiinței? Le considerați esențiale în formarea experiențelor conștiente? Și, dacă da, cum contribuie ele la înțelegerea noastră generală a conștiinței?

Christof Koch: În cazul majorității oamenilor, cu excepția celor care suferă de schizofrenie, percepțiile sunt relativ stabile. De exemplu, dacă priviți o iluzie vizuală timp de câteva minute, veți continua să percepeți același efect. Acest lucru sugerează că

influențele „de sus în jos” ale așteptărilor noastre nu sunt atât de puternice pe cât s-ar crede. Desigur, ele există, mai ales în condiții în care avem foarte puțin timp la dispoziție – dacă trebuie să raportăm ceea ce vedem într-un interval de 100 de milisecunde, influențele „de sus în jos” joacă un rol mai important.

Totuși, consider că rolul lor este adesea supraestimat. Dacă mă uit la un perete, văd pur și simplu un perete. Nu există nimic de anticipat, nimic care să se schimbe – doar un perete gol. Aceasta sugerează că, cel puțin în astfel de condiții, influența proceselor „de sus în jos” este mai redusă decât se crede. Și nu cred că, în acest caz, procesarea predictivă explică percepția.

Ce implicații etice anticipați că ar putea avea “Teoria informației integrate”? În special, în legătură cu panpsihismul, care sugerează că toate particulele ar putea poseda un anumit grad de conștiință. Ce consecințe are aceasta asupra eticii animalelor, având în vedere că teoria susține că multe specii, mai ales cele cu sisteme nervoase complexe, sunt capabile să experimenteze durere și plăcere? Credeți că IIT are implicații etice directe?

Christof Koch: Da, absolut. Acesta este unul dintre motivele pentru care, în urmă cu mulți ani, am devenit vegetarian. De fapt, încerc chiar să nu omor insecte, nici

măcar albine. Cred cu tărie că și ele simt ceva. Diferența este că nu au o voce. Ele nu se pot exprima așa cum o facem noi.

O albină, de exemplu, nu are un chip expresiv sau un limbaj articulat, dar creierul ei, deși mult mai simplu decât al nostru, îi permite să experimenteze plăcerea sau durerea. Îmi imaginez că simte bucurie atunci când bea nectar auriu și zboară prin lumina caldă a soarelui. La fel ca noi, și ea există între două eternități, nașterea și moartea, fiind aruncată în acest univers fără voia ei.

Prin urmare, avem o responsabilitate morală să minimizăm suferința tuturor ființelor, inclusiv a celor mai mici și aparent ne semnificative. Da, consider că IIT are implicații etice profunde pe care trebuie să le luăm în serios.

Cu toate cunoștințele acumulate despre creier și conștiință, v-au ajutat acestea să vă corecți anumite obiceiuri nesănătoase din viața de zi cu zi? Vă oferă această înțelegere un avantaj în schimbarea comportamentului, sau descoperirile științifice nu au un impact real asupra acestor aspecte?

Christof Koch: Nu, deloc. Trebuie să mă confrunt cu aceleași provocări ca toți ceilalți. Încerc să beau mai puțin, să fiu mai atent la ceilalți, iar aceasta nu este o

chestiune de neuroștiințe, ci o lecție din tradițiile înțelepciunii.

Sunt un admirator al lui Marcus Aurelius și al stoicismului, iar ceea ce pot spune este că schimbarea necesită disciplină și efort mental constant. IIT nu oferă nicio „superputere” care să îți permită să exclami: „Ah, acum înțeleg cum funcționează creierul, deci îmi pot corecta automat toate obiceiurile proaste!”. Nu, schimbarea necesită voință și un efort activ de evitare a tiparelor comportamentale dăunătoare.

De fapt, aceasta este adevărata libertate în sensul IIT. Teoria face și afirmații despre liberul arbitru deoarece susține că acesta există într-un sens fundamental. Îmi pot influența propria mea viață. Sunt „stăpânul convingerilor mele, căpitanul sufletului meu”.

Pentru a-mi exercita această libertate, trebuie să reflectez, să îmi analizez obiceiurile și să decid conștient ce vreau să întăresc și ce vreau să elimin. Acest proces necesită disciplină zilnică, iar în timp, schimbarea comportamentală modifică structura creierului și, implicit, cursul vieții mele.

Ce părere aveți despre tendința actuală de a încorpora aspecte „spirituale” în știință, prin concepte precum sincronicitatea cuantică și vindecarea cuantică? Considerați că aceste

abordări sunt problematice sau chiar imorale? Există dovezi științifice care să susțină relevanța acestor idei sau ele câștigă popularitate mai degrabă ca produse comerciale menite să atragă persoane care nu au capacitatea de a înțelege fundamentele științifice reale?

Christof Koch: În prezent, nu există nicio dovadă că mecanica cuantică microscopică, în special fenomenul de entanglement, esențial pentru calculul cuantic, joacă vreun rol în funcționarea creierului. Creierul este, din punct de vedere fizic, un sistem extrem de „zgomotos” și instabil pentru astfel de fenomene.

Dacă ne uităm la computerele cuantice de astăzi, acestea funcționează la temperaturi extrem de scăzute, în jur de 25-35 miliKelvin, adică de aproximativ 50.000 de ori mai rece decât temperatura creierului uman. Acest contrast face extrem de puțin probabil ca efectele cuantice să fie relevante pentru biofizica sau biochimia creierului.

Sigur, există multe persoane care vând idei precum „vindecarea cuantică” sau „terapia cu cristale”, dar acestea sunt complet separate de ceea ce face cu adevărat creierul. Nu există nicio bază științifică solidă pentru astfel de afirmații.

Ați avut o dezbatere interesantă cu Sanctitatea Sa, Dalai Lama, despre spiritualitate și creier. Ați putea împărtăși câteva idei care au apărut în timpul discuției? Cum a contribuit interacțiunea dintre cunoștințele științifice și perspectivele spirituale la o înțelegere mai profundă a conștiinței? Au existat aspecte ale conversației care v-au influențat gândirea în explorările ulterioare ale minții?

Christof Koch: Am fost de două ori în India, unde am avut ocazia să mă întâlnesc și să dezbăt cu Sanctitatea Sa, Dalai Lama. Am descoperit că împărtășim multe valori comune, în special ideea de a minimiza suferința tuturor ființelor conștiente. Acesta este un principiu fundamental în budismul tibetan, care subliniază că conștiința nu este exclusiv umană, ci este prezentă la toate creaturile capabile de experiență subiectivă.

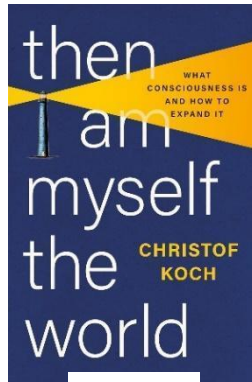
Totuși, am avut puncte de vedere divergente în ceea ce privește reîncarnarea. Eu susțin o perspectivă materialistă, pe care am rezumat-o într-o frază simplă: *No brain, never mind*. Cu alte cuvinte, odată ce creierul se degradează sau moare, substratul fizic al conștiinței dispăre. În această viziune, conștiința nu poate exista fără un suport material.

Chiar dacă ar exista un mecanism exotic al fizicii care să susțină o formă de continuitate, ar trebui să existe un suport clar identificabil în spațiu și timp, capabil să păstreze amintirile și identitatea individuală. În budism,

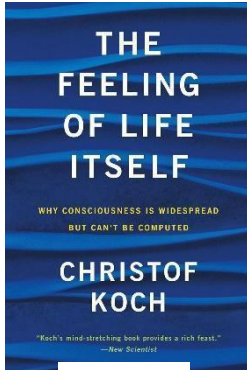
există conceptul de *bardo*, un spațiu de tranziție între moarte și renaștere. Însă, dacă ceva din viața mea ar supraviețui într-o existență viitoare, ar trebui să existe un substrat care să păstreze acea informație, iar până acum, nu avem dovezi pentru o astfel de posibilitate.

După ce i-am expus acest punct de vedere, Dalai Lama a râs cu poftă și a spus doar: "Ei bine, vom vedea". N-am mai avut nimic de adăugat.

Cărți scrise de Christof Koch

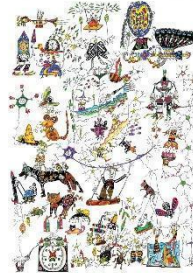


2024



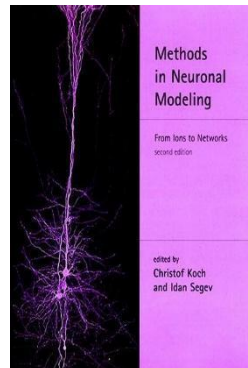
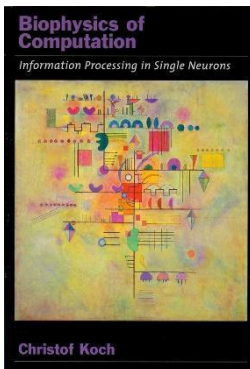
2020

CONSCIOUSNESS
CONFESSIONS OF A ROMANTIC REDUCTIONIST



CHRISTOF KOCH

2017



IX. Dialogul cu Joseph LeDoux



Joseph LeDoux este profesor de neuroștiințe, psihologie, psihiatrie și psihiatrie pediatrică și adolescentină la Universitatea din New York. Este directorul Emotional Brain Institute și al Nathan Kline Institute for Psychiatric Research.

De asemenea, este director adjunct al Centrului pentru Limbaj, Muzică și Emoții din cadrul Max Planck-NYU și membru al Academiei Naționale de Științe a Statelor Unite.

LeDoux este autorul unor lucrări fundamentale despre creier, printre care *Anxious* (2015), *Synaptic Self* (2002) și *The Emotional Brain* (1996). Pe lângă cariera sa academică, este cântăreț și compozitor în trupa folk-rock *The Amygdaloids* și în duo-ul acustic *So We Are*.

A primit numeroase premii prestigioase, printre care *Distinguished Scientific Contributions to Psychology* acordat de *American Psychological Association* (2010), *Karl Spencer Lashley Award* oferit de *American Philosophical Society* (2011) și *Gantt Medal* din partea *Pavlovian Society* (2012).

Pentru mai multe informații: www.joseph-ledoux.com

Sunteți unul dintre cei mai importanți experți în studiul emoțiilor și al mecanismelor prin care creierul, în interacțiune cu corpul și mediul, generează comportamente specifice. Cu toate acestea, vă exprimați un anumit scepticism cu privire la utilitatea conceptelor vernaculare sau populare legate de emoții. Ați putea detalia mai clar cum abordați această problemă?

Joseph LeDoux: Nu aș spune că sunt împotriva limbajului popular în sine. Cred că acesta joacă un rol esențial în modul în care ne raportăm la experiențele noastre mentale, deoarece trăim și comunicăm, în mare parte, prin astfel de concepte. Problema apare atunci când aplicăm aceste categorii la comportamente specifice, fără o analiză atentă.

De exemplu, creierul uman poate răspunde la un stimul periculos în mai multe moduri. Un reflex automat poate fi declanșat instantaneu, un model de acțiune fix, cum ar fi reacția de „înghețare”, poate fi activat, sau poate exista un răspuns bazat pe învățare și obiceiuri. La un nivel și mai complex, o persoană poate formula o strategie cognitivă orientată spre un obiectiv specific. Toate aceste răspunsuri pot avea loc fără conștientizare. Doar în ultima instanță, putem vorbi despre un răspuns conștient.

Prin urmare, nu putem reduce comportamente precum înghețarea sau fuga la simple expresii ale fricii. Frica

poate fi experimentată conștient în diverse moduri, iar multe dintre reacțiile defensive ale organismului nu implică neapărat o experiență subiectivă de frică.

În cercetările dumneavoastră, ați studiat îndeaproape amigdala, o regiune a creierului asociată cu detectarea și reacția la amenințări. Cu toate acestea, susțineți că asocierea acestei structuri cu emoția de „frică” este problematică. Ce v-a determinat să propuneți o schimbare conceptuală de la termenul convențional de „circuit al fricii” la „circuit de apărare și supraviețuire”?

Joseph LeDoux: Pentru că nu cred că frica își are originea în amigdala. Dacă ne uităm la evoluție, toate animalele bilaterale, adică toate vertebratele care au trăit vreodată, au trebuit să dezvolte un sistem capabil să detecteze și să răspundă la pericole. Acesta este un mecanism biologic primitiv, fundamental pentru supraviețuire.

Peste acest strat primar s-au adăugat, la speciile mai complexe, procese cognitive care permit conștientizarea acestor pericole și formarea experiențelor subiective. Dar, fără această auto-referențialitate conștientă, nu putem vorbi despre emoții în sensul strict al cuvântului.

Am creat un tricou cu sloganul „No self, no fear” – adică, fără un sine conștient, frica nu există. Pentru ca o emoție să fie experimentată, trebuie să existe o conștientizare a propriei implicări în situație. Răspunsurile reflexe sau instinctuale nu constituie, prin ele însele, o experiență emoțională.

Dacă vedem un câine lovit de o mașină, care se zbate de durere pe marginea drumului, avem tendința să-i atribuim trăiri similare cu ale noastre. Însă ceea ce observăm, în mod obiectiv, sunt reacțiile sale reflexe: zvârcolirile, mârâitul, spasmele musculare. Aceste manifestări nu sunt, în sine, dovezi ale suferinței conștiente.

Nu spun că animalele nu simt durere, ci că trebuie să facem o distincție clară între răspunsurile automate ale sistemului nervos și experiențele conștiente de frică, durere sau alte emoții.

Această perspectivă pare critică față de abordările convenționale în cercetarea științifică a emoțiilor. De exemplu, în studiile pe animale, cercetătorii observă adesea comportamente și le clasifică în funcție de emoții umane. Credeți că această metodă este problematică atunci când investigăm emoțiile în general?

Joseph LeDoux: Am abordat acest subiect în scrierile mele și, recent, am recitit un capitol dintr-o carte despre antropomorfism, scrisă de Elizabeth Knoll, care discuta perspectiva lui Darwin.

Darwin a trăit în Anglia victoriană, unde antropomorfismul era o normă culturală, un mod de a înțelege lumea. Întrucât teoria evoluției și selecției naturale era deja contestată din motive religioase, Darwin a făcut o alegere strategică: a preferat să descrie mințile animalelor în termeni umani, în loc să vorbească despre mintea umană în termeni animalii.

Această abordare a fost mai ușor de acceptat la acea vreme. Teoriile lui Darwin au fost punctul de plecare al cercetării moderne asupra emoțiilor, iar la sfârșitul secolului al XIX-lea, discipolii săi considerau comportamentul drept un „ambasador al minții”. Deși există un adevăr în această idee, comportamentele observate la animale sunt adesea răspunsuri reflexe, înnăscute, mai degrabă decât expresii ale unei experiențe conștiente de frică.

Atribuirea automată a unor stări conștiente fără dovezi clare a fost, de altfel, unul dintre motivele care au dus la revoluția behavioristă în psihologie. Nu putem presupune că „pare conștient, deci este conștient” fără o metodă clară de testare. Iar acest lucru este extrem de dificil de demonstrat la animale.

Unii mă acuză că neg emoțiile animalelor, dar nu este adevărat. Ceea ce susțin este că este extrem de dificil, din punct de vedere metodologic, să demonstrăm prezența emoțiilor conștiente la animale.

Susțineți că ceea ce a fost conservat prin evoluție nu este comportamentul în sine, așa cum consideră mulți oameni de știință și filosofi, ci mai degrabă circuitele neuronale care generează comportamente. Prin urmare, același comportament poate apărea din circuite diferite. Puteți oferi un exemplu concret care să ilustreze această distincție?

Joseph LeDoux: Cred că e important să clarific acest punct.

Nu spun că comportamentul nu este relevant. El este important, însă, la un alt nivel de analiză.

Sistemul nervos controlează comportamentul, iar interacțiunea dintre un animal și mediul său este reglată de acest sistem. Când speciile evoluează și se diversifică, ele dezvoltă moduri distincte de a reacționa la mediu. Dar ideea mea a fost că, dacă ne întoarcem la originile vieții, nu comportamentele specifice au fost transmise mai departe, ci cerințele fundamentale ale supraviețuirii.

Prima celulă capabilă să se reproducă, acum aproximativ 3,7-4 miliarde de ani, a trebuit să poată detecta și evita pericolele din mediul său. De asemenea, trebuia să identifice și să proceseze nutrienți, să mențină un echilibru al fluidelor și ionilor și să se reproducă pentru a asigura continuitatea speciei.

De exemplu, o bacterie primitivă expusă la un mediu acid își folosește flagelii pentru a se îndepărta de pericol. Dacă nu ar avea acest mecanism, nu ar supraviețui suficient de mult pentru a se reproduce. Aceste cerințe fundamentale ale vieții sunt comune tuturor formelor de viață, de la organisme unicelulare la specii complexe, precum noi.

Ceea ce am moștenit, așadar, nu sunt comportamentele specifice, ci necesitatea de a reacționa la pericole, de a detecta surse de hrană, de a menține homeostazia și de a ne reproduce. Acestea sunt cerințele fundamentale ale supraviețuirii tuturor organismelor, fie că vorbim despre o singură celulă sau despre un organism complex.

Nu spun că am moștenit amigdala de la bacterii, ci că am moștenit nevoia de a supraviețui. Capacitatea de a distinge între ceea ce este dăunător și ceea ce este benefic, precum și de a acționa în consecință, este un principiu fundamental care traversează întreaga istorie a evoluției.

Eu numesc aceste mecanisme „nevoi de supraviețuire” sau „strategii de supraviețuire”. Ele au fost implementate biologic în moduri specifice fiecărui grup de animale, fiecărei specii și, până la un anumit punct, fiecărui individ în parte.

Pe scurt, toate animalele au aceste cerințe de supraviețuire, dar doar organismele cu un sistem nervos au circuite specializate pentru supraviețuire.

Spre deosebire de mulți neurocercetători, nu considerați emoțiile drept procese biologice înrădăcinate în creier, ci susțineți că ele rezultă din procesarea cognitivă inconștientă. De ce credeți că există o astfel de divergență de opinii între dumneavoastră și alți neurocercetători în ceea ce privește natura emoțiilor?

Joseph LeDoux: Depinde de neurocercetătorul în cauză și de interesele sale. Teoria cognitivă a emoțiilor nu este o idee nouă. William James a avut o variantă proprie, dar forma modernă a acestei teorii își are rădăcinile în anii 1950, în teoria disonanței cognitive formulată de Leon Festinger, și a fost dezvoltată ulterior de Stanley Schachter și Jerome Singer în anii 1960, prin teoria cognitivă a emoțiilor. M-am bazat pe aceste idei în cercetările mele asupra pacienților cu creier scindat, în anii 1970.

Mentorul meu, Mike Gazzaniga, și cu mine am observat un fenomen fascinant: atunci când emisfera dreaptă a unui pacient cu creierul scindat genera un comportament, iar noi întrebam emisfera stângă „De ce ai făcut asta?”, aceasta fabrica o explicație, adică confabula o poveste pentru a da un sens acțiunii. Acest fenomen este în acord cu teoria lui Festinger, care afirmă că, atunci când ne confruntăm cu informații contradictorii, creierul nostru caută să le reconcilieze pentru a menține un echilibru cognitiv.

În acest caz, comportamentele generate inconștient de emisfera dreaptă ar putea produce stres sau anxietate, deoarece oamenii au o nevoie fundamentală de a crede în ideea de liber arbitru, fie că acesta există sau nu.

Dacă corpul nostru generează comportamente pe care nu le conștientizăm, acest lucru poate fi profund deranjant, iar creierul încearcă să găsească o explicație pentru a-și menține coerența narativă. Gazzaniga a propus că sistemele emoționale ar putea fi printre acele structuri cerebrale care produc astfel de comportamente inconștiente și care necesită ulterior o interpretare cognitivă pentru a fi integrate în autoevaluarea noastră mentală.

Pentru a investiga aceste idei, am apelat la studii pe șobolani, deoarece am vrut să înțeleg comportamentele inconștiente care, la oameni, ar putea declanșa interpretări cognitive. Teoria emoțiilor a

lui Schachter și Singer s-a dezvoltat pornind de la teoria disonanței cognitive a lui Festinger. Deși a evoluat mult de la formularea sa inițială din anii 1960, ea rămâne o abordare solidă asupra emoțiilor. Din perspectiva mea, oferă o viziune mult mai realistă asupra modului în care ne formăm emoțiile.

Imaginați-vă că vă aflați în vârful unui munte, iar rucsacul v-a căzut peste o stâncă. Soarele apune, nu aveți mâncare, apă sau haine călduroase. Începeți să simțiți frică și anxietate cu privire la soarta dumneavoastră. Dar dacă analizăm acest scenariu, ne dăm seama că sistemul amigdalei, așa-numitul „sistem al fricii”, sau, cum îl numesc eu, sistemul amenințării, a evoluat inițial ca un mecanism de apărare împotriva prădătorilor. Însă, în această situație, nu există niciun prădător care să vă sperie. Frica nu este determinată de un anumit tip de stimul extern, ci este o interpretare a contextului. Creierul monitorizează semnalele corpului – stomacul trimite indicii despre epuizarea rezervelor de energie, termoreceptorii raportează că temperatura corporală scade. Această combinație de semnale fiziologice și evaluare situațională declanșează îngrijorarea că ați putea muri de foame, îngheța sau suferi de deshidratare.

În realitate, în viața de zi cu zi ne putem teme dintr-o varietate de motive care nu au absolut nicio legătură cu prezența unui prădător. Acesta este motivul pentru care consider că emoțiile nu sunt reacții biologice fixe la

stimuli specifici, ci mai degrabă procese interpretative, modelate de cogniție și context.

Cum influențează perspectiva dumneavoastră alternativă înțelegerea clinică a tulburărilor de anxietate? Ce perspective teoretice noi oferă această abordare, comparativ cu perspectivele mai convenționale?

Joseph LeDoux: Tocmai am publicat un articol în *Molecular Psychiatry* intitulat „Putting the 'mental' back into 'mental' disorders”.⁸⁰ Ideea centrală este că abordarea tradițională a tratamentului tulburărilor mentale, care a început în anii 1950, a fost dominată de o viziune behavioristă care a marginalizat complet experiența conștientă subiectivă.

În acea perioadă, două abordări au devenit dominante: terapia comportamentală și tratamentul psihofarmacologic.

Terapia comportamentală, inspirată direct din lucrările lui Skinner și Watson, a deschis calea terapiei cognitive. Inițial, terapia cognitivă avea o dimensiune mentală mai pronunțată (fondatorul său, Aaron Beck, a fost psihanalist), însă, treptat, aceasta s-a transformat într-o

⁸⁰ Ediția originală: <https://www.nature.com/articles/s41380-021-01395-5>.

abordare din ce în ce mai behavioristă, bazată pe măsurători obiective.

În ceea ce privește psihofarmacologia, trebuie să ne întrebăm cine erau cercetătorii care lucrau în laboratoarele industriei farmaceutice în acea perioadă. Majoritatea fuseseră formați de behavioriști în anii '40, '50 și '60. Presupunerea lor era simplă: dacă un șobolan plasat într-o situație stresantă primea un medicament care îi reducea comportamentul de evitare, atunci acest medicament ar trebui să reducă și anxietatea umană.

Dar, în practică, această abordare nu a funcționat atât de bine. În prezent, industria farmaceutică se retrage din cercetarea anxietății, deoarece nu a reușit să producă tratamente noi. Majoritatea claselor de medicamente utilizate astăzi au fost descoperite accidental în anii 1960, precum inhibitorii recaptării norepinefrinei sau benzodiazepinele.

Problema este că un medicament care modifică comportamentul unui șobolan (de exemplu, reducând reacția de înghețare) nu schimbă în mod necesar angoasa psihologică umană, cum este anxietatea, îngrijorarea, frica trăită subiectiv. Trebuie, desigur, să controlăm reacțiile fiziologice ale pacientului, dar acest lucru nu este suficient. Experiența subiectivă trebuie luată în serios.

Având în vedere perspectiva dumneavoastră asupra emoțiilor, care atribuie originea acestora procesării cognitive inconștiente, mai degrabă decât legăturilor biologice directe cu creierul, vedeți vreo influență din teoriile lui Sigmund Freud? El a explorat, de asemenea, rolul inconștientului în modelarea comportamentului și emoțiilor umane. Identificați vreo legătură între ideile sale și dezvoltarea teoriilor dumneavoastră?

Joseph LeDoux: Aș spune că evoluțiile din anii 1950, care au dus la dezvoltarea unor noi forme de terapie, au fost de fapt o încercare de a se distanța de Freud, considerat prea subiectiv. Însă, în procesul de eliminare a subiectivității aduse de Freud în psihiatrie, a fost eliminat și un aspect esențial al psihologiei umane.

În loc să spunem: „Poate că Freud a greșit în privința inconștientului ca tărâm întunecat plin de conflicte refulate și dorințe sexuale reprimite, dar asta nu înseamnă că trebuie să eliminăm complet studiul experienței subiective”, s-a mers în direcția opusă, adică s-a adoptat o viziune strict obiectivă, centrată pe comportament.

Chiar și terapia cognitiv-comportamentală (CBT), care inițial încorporase elemente subiective (în special la Beck și Ellis, în anii 1950-60), s-a transformat treptat

într-o abordare mai rigidă, bazată pe măsurători și standardizări.

Această schimbare a fost în mare parte motivată economic deoarece asigurările de sănătate au preferat terapii cuantificabile. În loc să se concentreze pe experiența subiectivă a pacientului, terapia a devenit o procedură birocratică, în care simptomele erau măsurate și clasificate pe baza unor scale standardizate. De exemplu, modelul RDoC al National Institute of Mental Health include numeroase dimensiuni obiective de clasificare a tulburărilor, iar raportul verbal al pacientului, cel mai apropiat indicator al experienței sale subiective, este tratat doar ca un simptom printre altele.

Dar, în fond, de ce merg oamenii la terapie? Nu pentru că vor să îmbunătățească un scor pe o scală de măsurare, ci pentru că se simt rău și vor să se simtă mai bine. Din perspectiva mea, adevărata misiune a psihoterapiei ar trebui să fie îmbunătățirea stării interioare a pacientului, nu doar modificarea comportamentului său exterior.

Reluați experimentul de gândire în care aveți sarcina de a proiecta de la zero un nou tip de ființă umană, inclusiv circuitele sale de supraviețuire. Credeți că este posibil să programați aceste circuite

astfel încât individul să caute fericirea, binele sau plăcerea?

Joseph LeDoux: Este o idee interesantă și cred că ar putea fi util să ne gândim la o astfel de reprogramare umană, având în vedere că, nu ne aflăm într-o situație foarte bună ca specie. Totuși, înainte de a începe, trebuie să ne întrebăm ce anume ar trebui să păstrăm și ce ar trebui să schimbăm.

De exemplu, un aspect fundamental al supraviețuirii este întărirea comportamentului. Când unui șobolan îi este întărit un anumit comportament, îi este plăcere să facă asta? Este plăcerea ceea ce determină această întărire? Olds a continuat să scrie articolul în revista Science, intitulat „Centrul Plăcerii în Creier”, iar acum câțiva ani, cât timp era încă în viață, l-am întrebat pe Peter Milner: „La ce vă gândeați cu adevărat când studiați aceste centre ale plăcerii?”

Iar el mi-a răspuns: „În primul rând, nu studiam plăcerea, ceea ce încercam să descoperim erau mecanismele de întărire din creier.” Dacă citiți articolul publicat de Olds, veți observa că termenul „plăcere” nu apare nici măcar o dată în conținutul lucrării; este prezent doar în titlu. Probabil că editorii revistei au ales acest titlu, iar Olds s-a lăsat sedus de el și a început să promoveze ideea „centrelor plăcerii”.

Mai târziu, Rory Wise a lansat ideea că dopamina este „substanța chimică a plăcerii”, însă foarte curând atât Wise, cât și Olds au respins această interpretare și au afirmat că nu studiau, de fapt, plăcerea, iar dopamina nu era „molecula plăcerii”. Însă, odată ce astfel de idei scapă de sub control, este imposibil să le mai retragi.

Așa că astăzi toată lumea vorbește despre întărire comportamentală în termeni de recompensă și plăcere, dar, în esență, ceea ce studiem este modul în care dopamina influențează activitatea celulară, determinând o schimbare comportamentală.

Acest lucru nu înseamnă că un animal nu experimentează plăcere sau că o persoană nu simte plăcere, dar aceste senzații nu sunt echivalente cu mecanismul de întărire a unui comportament.

În concluzie, sunt curios să aud perspectiva dvs. asupra unei întrebări care preocupă filosofii de ceva timp. Credeți că facem progrese în înțelegerea modului în care qualia apar și sunt menținute în creierul uman, sau suntem încă departe de a oferi un răspuns clar la „problema dificilă” a conștiinței?

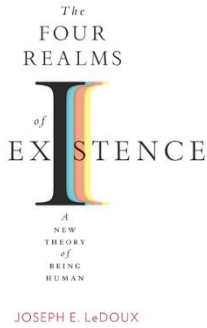
Joseph LeDoux: Un răspuns pe care l-aș oferi este că oamenii de știință trebuie să fie atenți la conceptele pe care le importă din filosofie. Filosofia este un sistem de reguli și raționamente, ceea ce este minunat, dar nu

înseamnă că acesta este modul în care funcționează de fapt creierul.

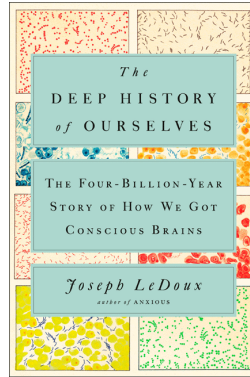
„Problema dificilă” a conștiinței a fost formulată într-un mod care face ca soluționarea ei să fie imposibilă, deoarece presupune un dualism minte-corp. Aceasta poate fi o idee validă în filosofie, dar din punct de vedere științific, este imposibil să avem o viziune dualistă asupra creierului.

Cred că cei mai mulți neurocercetători sunt materialişti și doresc să cunoască unele mecanisme care fac ca qualia să apară. Am scos o nouă carte intitulată “Cele patru tărâmurii ale existenței: O nouă teorie a ființei”, în care propun un mod posibil de a gândi despre qualia, așa că poate cititorul poate încerca acolo.

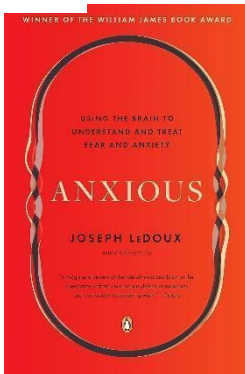
Cărți scrise de Joseph LeDoux



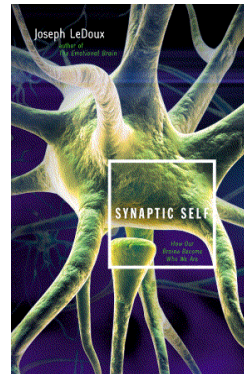
2023



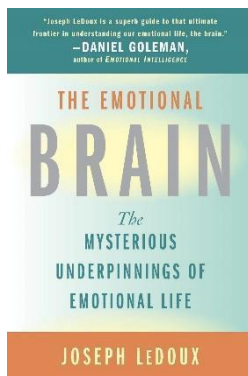
2019



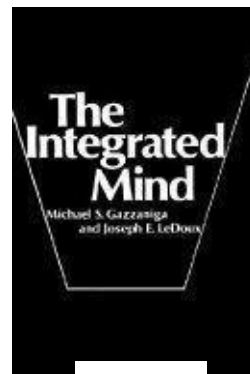
2015



2001



1996



1978

Concluzie

Această carte a avut ca scop introducerea cititorului în unele dintre cele mai actuale și fascinante dezbateri din filosofia și știința minții. Asemenea navigatorilor portughezi care au îndrăznit să traverseze Capul Bunei Speranțe, ne-am aventurat în complexitatea misterelor minții și în provocările pe care acestea le ridică.

Capul Bunei Speranțe, situat în apropierea orașului Cape Town, Africa de Sud, este unul dintre cele mai sudice puncte ale planetei. Datorită poziției sale, a fost considerat mult timp imposibil de navigat, din cauza furtunilor violente și a condițiilor extreme.

Până în 1488, nimeni nu reușise să străbată acest loc. Totul s-a schimbat odată cu Bartolomeu Dias, primul navigator portughez care a reușit să treacă dincolo de el. Ulterior, regele João al II-lea l-a redenumit *Capul Bunei Speranțe*, simbolizând speranța unei rute maritime directe între Oceanul Atlantic și Oceanul Indian. Această descoperire a deschis calea comerțului între Vest și Est, în special cu India, schimbând pentru totdeauna istoria economică, culturală și politică a lumii.

Asemenea lui Bartolomeu Dias, cred că multe dintre ideile prezentate în această carte vor continua să se dezvolte în anii următori. Poate că, într-o zi, vom reuși să transformăm “furtunile minții” într-un Cap al Bunei Speranțe.

Fiecare pagină a acestei cărți este precum o nouă zi pe mare, în căutarea unui tărâm care poate nu va fi niciodată găsit. Dar poate că tocmai curiozitatea noastră este vântul care va împinge viitoarele descoperiri, scoțând la lumină adevăruri încă ascunse în umbra necunoscutului.

A meritat această călătorie? Să ne încredem în cuvintele poetului portughez:

"A meritat? Totul merită
Dacă sufletul nu este mic.
Oricine vrea să treacă dincolo de Bojador
Trebuie să treacă dincolo de durere.
Dumnezeu a dat primejdia și abisul mării,
Iar în ea El a reflectat cerul."⁸¹

Fie ca această Odisee a minții să ofere cititorului un punct de sprijin sigur în porturile înțelegerii și reflecției.

⁸¹ Fernando Pessoa, partea a doua: X. Marea Portugheză din cartea *The Message*. Versiunea originală în portugheză: “Valeu a pena? Tudo vale a pena / Se a alma não é pequena./ Quem quer passar além do Bojador / Tem que passar além da dor. / Deus ao mar o perigo e o abismo deu, / Mas foi nele que espelhou o céu.

Cu orizontul deschis înaintea, sper ca fiecare dintre voi să găsească în această carte o busolă de încredere pentru propriile explorări și ca înțelepciunea dobândită pe parcursul acestei călătorii să lumineze noi drumuri spre viitoruri necunoscute, unde frontierele minții umane pot fi împinse și mai departe.

Pentru a încheia această concluzie, și cartea, într-o notă mai puțin poetică: această lucrare a fost inspirată de numeroasele cursuri online pe care le-am organizat, în calitate de profesor principal, în ultimii doi ani. În acest timp, am avut privilegiul de a ajunge la peste 600 de studenți din aproximativ 35 de țări.

Le sunt profund recunoscător tuturor, pentru că m-au învățat mai mult decât aș fi putut spera să transmit. Curiozitatea lor pătrunzătoare, întrebările provocatoare și intervențiile inspirate au dat naștere unor dezbateri fascinante, făcând fiecare curs nu doar o experiență educațională, ci și una intelectuală stimulantă, pentru toți, dar mai ales pentru mine.

Fie ca Odissea minții să continue!

Mulțumiri

Marcus Aurelius, filosof stoic și împărat roman, își începe celebra carte *Meditații* mulțumind tuturor oamenilor care s-au intersectat în viața sa și de la care a învățat ceva deosebit. Or, o lucrare de o asemenea amploare nu ar putea avea loc fără contribuția directă și indirectă a unui număr extrem de variat de persoane, motiv pentru care calc pe urmele împăratului mulțumindu-le tuturor.

Încep, desigur, prin a-mi exprima recunoștința față de cei 8 cercetători incredibili care mi-au oferit onoarea și încrederea de a împărți scena pentru această carte: David Chalmers, Susan Blackmore (și Alison Seldon), Sir Roger Penrose (și Helen McGregor), Nicholas Humphrey, Joseph LeDoux, Anil Seth, Christof Koch și Karl Friston, un imens mulțumesc.

În execuția acestei cărți, Tássia Vianna a fost persoana de bază în transcrierea dezbaterilor orale în limba engleză, transformându-le în dialoguri scrise care au fost apoi aprobate de fiecare dintre participanți. Ana Monteiro mi-a fost alături în producția creativă a copertei acestei cărți, ale cărei fețe au fost desenate manual datorită talentului și creativității sale pure. Pentru amândouă, o apreciere sinceră.

Colegilor de la Universitatea din Porto, în special profesoarei Sofia Miguens, coordonatoarea grupului de cercetare Mind, Language and Action Group, în cadrul căruia îmi desfășor cercetările, care a avut cea mai mare influență în gândirea filosofiei într-un mod deschis către lume, precum și profesorului José Meirinhos, directorul Institutului de Filosofie, și decanului Facultății din cadrul Universității din Porto, profesoara Paula Pinto Costa, pentru susținerea diferitelor proiecte fără a pune obstacole. Lui Isabel Marques, manager științific la Institut, îi mulțumesc pentru toată răbdarea.

În continuare, aș dori să le mulțumesc tuturor celor care au contribuit ca ideile mele să ajungă în lume, servind ca un stimulent constant în cercetarea mea.

În Brazilia, Gabriel Mograbi și Paulo Taddei, de la Universitatea Federală din Rio de Janeiro; Nythamar de Oliveira, de la Universitatea Catolică Pontificală din Porto Alegre; și Maria Luiza Ilenaco, împreună cu Osvaldo Pessoa Jr., de la Universitatea din São Paulo.

În România, Florin Piscociu și echipa sa de la Mindlifeline (Léa Chibany, Luana Aldea și restul membrilor) au fost un sprijin constant, a căror etică a muncii mă inspiră în fiecare zi.

În Malta, Ian Rizzo, Valdeli Pereira, Francois Zammit și restul prietenilor mei de la Philosophy Sharing Foundation mi-au arătat importanța unei filosofii publice, nepăstrată în turnul de fildeș și accesibilă

tuturor cetățenilor. În Cipru, entuziasmul și setea de cunoaștere a lui Panayiotis Stavrou sunt o motivație constantă.

În Chile, prietenia și admirația pentru munca depusă de Ricardo Ramirez, Camilo Garcia, Marcelo Martín, Nicole Nakousi, David Araya și alți colegi, care mi-au oferit privilegiul de o viață de a primi, la vârsta de 30 de ani, distincția de profesor onorific la Facultatea de Medicină Andrés Bello, o distincție pe care sper să o onorez pe tot parcursul vieții mele.

În cadrul proiectului meu privind etica inteligenței artificiale în medicină, ar trebui nominalizați mai mulți colegi: Pekka Mäkelä, Raul Hakli și Pii Telakivi (Uni. Helsinki); Radu Uzskai (Uni. București); Simona Tiribelli (Uni. Macerata); Heidi Mertes (Uni. Ghent); Inês Dutra (Uni. Porto), Sabina Leonelli (Uni. Exeter) și, mai recent, Antonio Chella și grupul de studii avansate de la Universitatea din Palermo.

Prietenilor mei de la Universitatea Catolică din Braga, cu care am avut o colaborare scurtă, dar intensă, al căror accent pe multidisciplinaritate a fost o încurajare: le mulțumesc lui Augusto Soares Da Silva, Carlos Morujão, Padre João Onofre, Paulo Dias, João Duque, António Melo, Bruno Nobre, Elton Marques, ngela Leite, Tânia Oliveira și directorului José Lopes - vă mulțumesc pentru că m-ați primit în casa voastră.

Filosofului Manuel Curado (Uni. Minho) și neurologului Georg Northoff (Uni. Ottawa), supervizorii doctoratului meu, le mulțumesc pentru toate îndrumările din această carieră academică încă scurtă, dar energică.

Doctoranzilor pe care am avut/au avut privilegiul de a-i supraveghea și care au avut încredere în îndrumarea mea, le sunt recunoscător lui Jaroslav Malík, Romeu Ivoela și Maria Luiza Ienacco pentru întreaga lor colaborare.

În mod aleatoriu, tuturor colegilor răspândiți în diferitele instituții care au apreciat împărtășirea cunoștințelor: Klaus Gärtner, Robert Clowes, Glorinda Andrada, Rui Vieira da Cunha, Sâmara Costa, Sara Fernandes, Bárbara Sousa e Brito, James Grayott, João Cordovil, João de Fernandes Teixeira, Leonel Moura, Ralph Bannell, Roberto Pereira, Diogo Gurgel, Diana Tavares, Dina Mendonça, Joana Rita Sousa, Tomás Magalhães Carneiro, ngelo Milhano, Björn Lundgren, Jorge Gonçalves, Luísa Neto, Peter Singer, Luiz Meirelles , Noam Chomsky, Yinchun Wang, printre mulți alții ai căror neuroni din hipocampus nu-mi permit să mi-i amintesc, dar care apar mereu în mintea mea.

Pentru a încheia această deja lungă notă de mulțumire - pe care datorez mult multor oameni -, Marcus Aurelius atribuie zeilor "faptul că constituția mea a supraviețuit acestui tip de viață atât de mult timp": nu pot decât să le mulțumesc marilor mei prieteni și familiei, care m-au ajutat în acest "fel" de viață dedicată cunoașterii: Dl.

Joaquim, Maria José și familia (Ana, Manuel, João și Maria Augusta); Rafael, Pedro, João Pedro și Carolina; frații Kevin și Dylan; Carlos și Zaida (adevărații responsabili pentru această "constituție"); bunica Arminda, verii Sandra, Nuno, António, Ana și unchii și mătușile; Monica și Kiara; Isabel și Isaura; Roberta și Bruno; Ana Maria, Gilberto și Ana Mafalda; Luna și pisicile Miró, Tareco, Mia, Didi, Manchinhas și Amarelinho (nume creative, știu).

Tuturor, un profund mulțumesc.

Biografia autorului

Steven S. Gouveia (Sion, 30/09/1992) este cercetător în cadrul Grupului Mind, Language and Action al Institutului de Filosofie al Universității din Porto (Portugalia), finanțat de Fundația pentru Știință și Tehnologie (CEECIND.02527.2022), unde conduce un proiect de 6 ani privind etica inteligenței artificiale în medicină, care are parteneriate cu Universitățile din Exeter, Yale și Helsinki (cf. <https://trustaimedicine.weebly.com/>).

A absolvit filosofia și la vârsta de 22 de ani și-a început doctoratul în filosofia minții la Universitatea din Minho (Portugalia), sub supravegherea filosofului Manuel Curado, după ce a fost cercetător invitat la Minds, Brain Imaging and Neuroethics la Royal Institute of Mental Health, Universitatea din Ottawa (Canada), sub supravegherea neurocercetătorului și psihiatrului Georg Northoff.

După finalizarea doctoratului (2021), a fost cercetător la Centrul de Studii Filosofice și Umaniste din cadrul Universității Catolice Portugheze, din Braga (Portugalia).

Din iunie 2023, este profesor onorific la Facultatea de Medicină a Universității Andrés Bello din Viña Del Mar, Chile, titlu acordat cu aceeași ocazie ca și Premiul Nobel Sir Roger Penrose. De asemenea, este consultant în

domeniul eticii pentru o companie de altruism eficient, "Altruistic Careers".

A publicat mai multe cărți academice. În 2016, a editat cartea *Philosophy and the Arts* în portugheză, având invitați unii dintre cei mai relevanți artiști din practica artistică națională, precum artista Joana Vasconcelos, rapperul Valete, muzicianul Fernando Ribeiro (Moonspell), Leonel Moura (IA), poeta Ana Luísa Amaral, printre alții.

În 2017, a co-editat cartea *Thinking Democracy* în portugheză cu prefața lui Noam Chomsky și, împreună cu Ana Figueiredo Sol, a editat *Bioethics in the 21st* în portugheză. În același an, a editat și *Philosophy of Mind: Contemporary Perspectives*, prima sa carte internațională, împreună cu Manuel Curado, la Cambridge Scholars Publishing.

În anul următor, a publicat prima sa carte de autor în portugheză la Editora Húmus, intitulată *Reflecții filosofice: Artă, minte și justiție*, cu o prefață a filosofului minții, João de Fernandes Teixeira.

În 2019, a publicat trei cărți editate: *Perception, Cognition and Aesthetics* și *Film and Philosophy: Bridging Divides*, la influenta editură Routledge și, pentru Vernon Press, a co-editat, împreună cu Manuel Curado, *Automata's Inner Movie: Science and Philosophy of Mind*.

El a publicat, de asemenea, cartea editată *The Age of Artificial Intelligence: an Exploration* la aceeași editură, la

care participă unii dintre cei mai influenți gânditori și transhumaniști, precum Daniel Dennett, Ben Goertzel (creatorul celebrului robot Sofia), David Pearce, Natasha Vita-More, Roman V. Yampolskiy și Vernor Vinge, printre alții.

În anul următor, a publicat o altă carte editată la Bloomsbury - aceeași editură responsabilă de publicarea lui Harry Potter - intitulată *The Philosophy and Science of Predictive Processing*.

În plus, a publicat cea mai căutată carte de autor de până acum, în portugheză: *Homo Ignarus: Rational Ethics for an Irrational World*, la Editura Minerva, care are prefața lui Peter Singer, o carte axată pe diverse probleme etice precum votul, umorul, eutanasia sau inteligența artificială.

În 2022, a publicat a treia sa carte de autor în limba engleză la editura Palgrave din New York, parte a grupului Springer Nature, cu titlul *Philosophy and Neuroscience: a Methodological Analysis*.

Tot în același an, el a publicat o carte editată intitulată *Artificial Intelligence: Conversations about the New World* cu participarea mai multor cercetători precum Peter Singer, Wulf Loh sau Sabina Leonelli. În 2023, co-editează din nou o carte cu Manuel Curado intitulată *Miști predictive: Old Problems and New Challenges* de Vernon Press.

În plus față de această producție academică intensă, este gazda și producătorul documentarului internațional *"The Age of Artificial Intelligence: a Documentary"*, la care participă cercetători internaționali și care este accesibil gratuit pe YouTube, cu peste 55 000 de vizualizări.

A fost vorbitor la zeci de conferințe științifice evaluate de experți din Portugalia și din străinătate. În plus, a fost invitat ca vorbitor la universități din întreaga lume, în Brazilia, Argentina, Chile, Malta, Italia, Republica Cehă, Coreea de Sud, Cipru, România, printre altele. Mai mult, a fost speaker la Conferința Science of Consciousness de două ori, cea mai relevantă conferință internațională privind studiile despre conștiință. De asemenea, a fost speaker la TEDx NOVA, în septembrie 2023.

A fost solicitat pentru diverse interviuri, participând la programe de televiziune și radio precum "A Minha Geração" (RTP3), "Sociedade Civil" (RTP2), "Linha da Frente" (RTP1), "Much Barulho para Nada" (RTP2), "Filosofia na Rua" (Antena 2), în plus față de mai multe podcast-uri, cum ar fi "45 Graus", "Despolariza", "Pergunta Simples", "Desassossego", "Smart Meat", printre multe altele.

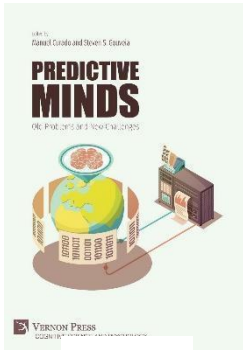
El este profesorul principal al mai multor cursuri online pe teme precum democrația, conștiința sau etica, cu participarea unor profesori invitați precum Sir Roger Penrose, Peter Singer, Noam Chomsky, Slavoj Zizek, Paul Bloom, printre alții.

A fost fondatorul și editorul principal al "Apeiron - Student Journal of Philosophy" (2012-2016), la care au participat cercetători precum Noam Chomsky, Peter Singer, Daniel Dennett sau Noël Carroll.

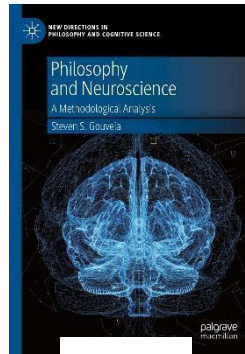
În sfârșit, a fost principalul organizator al mai multor conferințe internaționale, aducând în Portugalia academicieni precum Peter Singer (Princeton), Luciano Floridi (Yale), William Child (Oxford), Shaun Gallagher (Memphis), Dan Zahavi (Copenhaga), Karl Friston (Londra), David Papineau (New York), Tim Crane (Cambridge), printre alții. Mai multe informații pot fi găsite la www.stevensgouveia.weebly.com.

Proiect finanțat de FCT CEEC Individual Project 2022.02527.CEECIND de Fundação da Ciência e Tecnologia, la Mind, Language and Action Group, Institute of Philosophy, University of Porto, Address: Faculdade de Letras, Via Panorâmica s/n, P-4150-564 Porto, Portugal.

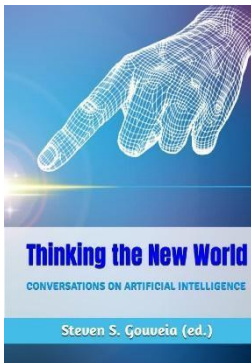
Cărți scrise de Autor



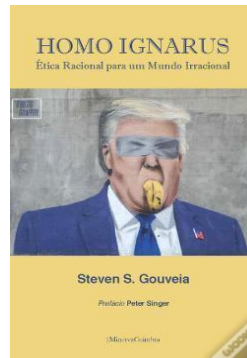
2023



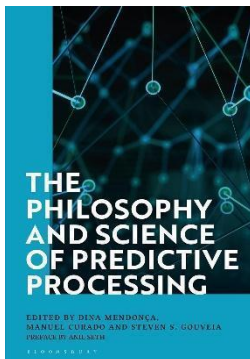
2022



2022



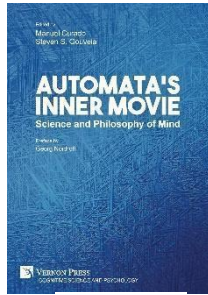
2020



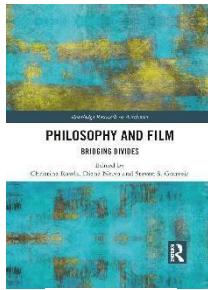
2020



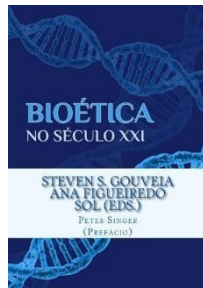
2020



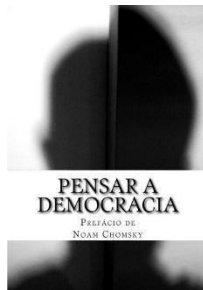
2019



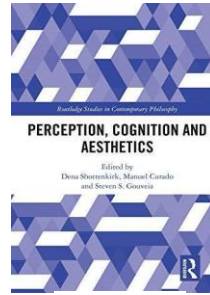
2019



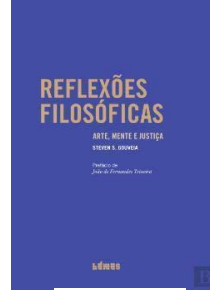
2018



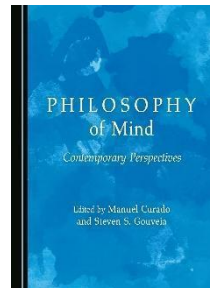
2017



2019



2018



2017



2016