

---

# DESCARTESOVA DEDIŠČINA IN VPRAŠANJE MIŠLJENJA IN ZAVESTI V KOGNITIVNI ZNANOSTI

O l g a M a r k i č

## 1. Uvod

Znanstvena revolucija v 17. stoletju je za raziskovanje narave vpeljala mehanicistični pristop in s tem pomembno spremenila razumevanje narave in mesta človeka v njej. Pred tem, v srednjem veku in renesansi, so o naravi razmišljali kot o vrsti organizma, o živi naravi. V skladu z Aristotelovim sistemom, ki je predstavljal temelje znanstvenemu pristopu tistega časa, so tako organske kot anorganske stvari počele tisto, kar je bilo v njihovi »naravi«. To pomeni, da so težile k svojim naravnim mestom. Na primer, kamni so padali na tla, ker so tla njihovo naravno mesto. V 17. stoletju pa je aristotelsko metodo razlage, ki je uporabljala pojma »končnega cilja« in »narave«, nadomestila mehanična metoda, ki je razlagala s pomočjo urejenega determinističnega obnašanja materije v gibanju (Crane 1995: 3). Galileo Galilei, Francis Bacon, René Descartes in Isaac Newton so za razliko od svojih predhodnikov, ki so študirali in interpretirali Aristotela, uvedli kot glavno metodo pri iskanju zakonitosti v svetu opazovanje, poskuse in natančno merjenje interakcij v naravi. Ta novi mehanični pogled na svet je Galilei opisal kot branje knjige, zapisane v matematičnem jeziku. Merjenja in oblikovanje zakonov narave v zapisu matematičnih enačb pa so pripomogli k razvoju znanosti (predvsem je tu mišljena fizika), kot jo poznamo danes.

Znanstveniki, ki so sprejeli mehanični pogled, so razlagali obnašanje stvari s sklicevanjem na vzročne zakone. Pojmovanje mehaničnega sistema je bilo omejeno s takratnim stanjem naravoslovne znanosti, predvsem fizike. Tako je bil mehanični sistem determinističen, interakcija pa je potekala zgolj z neposrednim dotikom. Kot primer zapletenega mehaničnega sistema je pogosto nastopala ura, ki je bila večkrat uporabljena

tudi kot metafora ali analogija v filozofskih argumentih. Kasnejši razvoj fizike je sicer zavrnil mehanično sliko sveta, razumljeno v tem specifičnem smislu, ni pa spodkopaval splošnega pogleda, da svet deluje v skladu z naravnimi zakoni ali zakonitostmi. Nekateri avtorji uporabljajo izraz mehaničen v tem širšem smislu, sinonimno z izrazom naturalističen, na primer Tim Crane (1995) in John Haugeland (1985).

Če so se znanstveniki pred znanstveno revolucijo v 17. stoletju pri razlagi neorganskih stvari zgledovali pri živi naravi, kjer je vse imelo svoje naravno mesto v skladu s harmoničnim delovanjem sveta, se je s sprejetjem mehanične slike sveta zadeva obrnila. Prvenstvo je prevzelo neorgansko; vzročna razlaga v neorganskem svetu je postala zgled tudi za razlago organskega. Mnogi znanstveniki so želeli s pomočjo enake metode, kot je v uporabi za neživa bitja, razložiti tudi vedenje živih bitij. Tako je René Descartes menil, da bi bilo mogoče tudi vedenje živali razložiti s povsem mehaničnim pristopom. Vendar je živalim s tem, ko jih je pojmoval kot stroje, odrekel zavest in mentalna stanja. Slednje po Descartesu ne sodi v mehansko sliko sveta, kajti mišljenja, zavesti in prostovoljnih dejanj ni mogoče razložiti z mehansko metodo naravoslovnih znanosti. Kot ugotavlja Cottingham (1995), pa je pogosto spregledano, da je Descartes v delu *Razprava o človeku (Traité de l'homme)* navedel tudi mnogo funkcij človeka, ki jih je mogoče razložiti mehanično, brez sklicevanja na posebno mislečo stvar (*res cogitans*). Mednje ne sodijo zgolj funkcije avtonomnega živčnega sistema, kot na primer dihanje in bitje srca, ampak tudi take, ki jih danes uvrščamo med psihološke, na primer čutno zaznavanje, spomin in notranji občutki lakote in strahu. Mehanično razlago je po Descartesu mogoče uporabiti tudi za razlago dejanj, na primer hoje ali petja, če se dogodijo brez mentalne pozornosti. Če pa je pozornost prisotna, potem terja postavitev posebne »racionalne duše«.

Descartes nas v petem delu *Razprave o metodi* povabi, da si zamislimo, da bi obstajali stroji, ki so povsem podobni ljudem in posnemajo vsa naša dejanja. Do spoznanja, da ti stroji niso resnični ljudje, naj bi po njegovem mnenju vodili dve poti. Prva pot oziroma prvi argument se opira na ugotovitev, da »ti stroji nikoli ne bi imeli daru govora ne drugih znakov, ki jih uporabljamo ljudje za izražanje svojih misli drugim« (Descartes 1957: 82). Čeprav bi bilo mogoče, da bi se tak stroj odzival

na zunanje dražljaje, pa po Descartesu ne bi mogel »iz lastnega nagiba nikoli postavljati besed, da bi smiselno odgovarjal na vse, kar se izreče vpricho njega, kot to zmorejo katerikoli ljudje« (*ibid.*). V drugem argumentu izpelje razliko med mehanskim strojem in človekom iz dveh trditev. Prvič, da je človekov razum univerzalen, »organi strojev pa morajo biti za vsako delovanje posebej razvrščeni«, in drugič, da »je praktično nemogoče, da bi bilo v enem stroju dovolj raznovrstnih delov, ki bi mogli sprožiti njegovo delovanje v vseh življenjskih prilikah, kakor dela to za našo dejavnost razum« (*ibid.*). S tema preizkusoma bi lahko spoznali tudi razliko med živalmi in človekom. Po Descartesu živali nimajo le manj razuma kot človek, ampak ga sploh nimajo. V pismu Morusu (5. februarja 1649) opozarja, da ljudje zmotno menijo, da živali mislijo. Po razmisleku je Descartes prišel do sklepa, da gibanje živali izvira iz enega samega principa, ki je telesen in mehaničen, in da ni mogoče dokazati, da bi bila v živalih misleča duša (Descartes 1957: 276).

Descartes je postavil ločnico med živalmi, katerih delovanje je podvrženo naravnim zakonom, in človekom, ki se s sposobnostjo mišljenja izmika mehanični razlagi. Stuart Shanker (2004) tako poudarja, da je Descartes zavračal doktrino »velike verige bitij«, kajti med živalmi in človekom naj bi bila prekinitev, ki je ne more zapolniti noben manjkajoči člen. Telo je lahko stroj, a človek s sposobnostjo mišljenja, jezikovnega sporazumevanja, usmerjanja dejanj in zavedanja je kategorično drugačen od živali. Descartes je poudaril razmejitev med mehaničnim, refleksnim in nehotenim gibanjem na eni strani ter namernim in hotenim gibanjem na drugi. Po njegovem prepričanju ne moremo izvajati racionalne duše iz možnosti materije, temveč mora biti posebej ustvarjena.

Razvoj znanosti, še posebej Darwinova teorija evolucije, je filozofe in znanstvenike spodbudil k iskanju poti, kako preseči prepad med živalmi, pojmovanimi kot mehaničnimi avtomati, in človekom kot racionalnim bitjem. Zagovorniki kontinuuma imajo po Shankarjevem mnenju dve možni poti: pokazati morajo, da je (i) vedenje živali inteligentno, ali (ii) da je vedenje človeka mehanično. Obe jemljeta Descartesa za duhovnega vodjo, čeprav se razlikujeta v tem, kar sprejemata za njegovo dediščino (Shankar 2004: 318).

V nadaljevanju bomo pokazali, kako sta ti dve poti zaznamovali dva pomembna pristopa iskanja naturalistične razlage duševnih procesov.

Zagovorniki klasične simbolne paradigme so začeli s pojmi poljudne psihologije in »od zgoraj navzdol« pokazali, kako so duševna stanja fizično realizirana oziroma da je vedenje človeka mehanično. S pristopom »od spodaj navzgor« pa skušajo znanstveniki z natančnim raziskovanjem živčnih mehanizmov tako pri živalih kot pri človeku ugotoviti, kako duševni pojavi in racionalno vedenje vzniknejo kot produkt evolucije. Ta dva pristopa kažeta tudi na razlike v reševanju dveh nalog, povezanih z vprašanjem naturalistične razlage: prvič, razložiti, kako fizični sistem lahko racionalno razmišlja o nečem (vprašanje intencionalnosti), in drugič, ugotoviti, kako je tak sistem lahko zavesten (vprašanje fenomenalne zavesti).

## 2. Kako je mogoča racionalnost v mehaničnem svetu?

### *Pot od zgoraj navzdol*

Descartes je kot enega temeljnih problemov materializma izpostavil vprašanje racionalnega mišljenja. Mnogi materialistični filozofi so zato skušali najti mehanično razlago za mišljenje. Na primer, francoski materialistični filozof La Mettrie (1709–1751) je izzval svoje sodobnike s knjigo *L'Homme machine* (*Človek stroj*, gl. La Mettrie 1996). Za razvoj kognitivne znanosti najbolj zanimivo idejo, da je mišljenje računanje, pa je zagovarjal angleški filozof Thomas Hobbes (1588–1679). Po njegovem mnenju naj bi bilo mišljenje notranji »mentalni pogovor«. Podobno kot pri glasnem pogovoru ali računanju s pomočjo svinčnika in papirja uporabljamo simbolne operacije, naj bi tudi pri mentalnem pogovoru uporabljali simbole, zapisane v ustrezni »živčni kodi«. In tako kot tisti, ki računa, sledi točno določenim pravilom (npr. pravilo prenosa čez desetico pri seštevanju večmestnih celih števil), tako mora tudi racionalno mišljenje slediti pravilom razuma. Kadar razmišljamo, naši možgani računajo.

Temeljni problem, kako razložiti racionalnost kot povsem fizikalni mehanizem, je ameriški filozof John Haugland imenoval *paradoks mehaničnega razuma*: »Manipulator je bodisi pozoren na to, kaj simboli pomenijo, ali pa ne. Če je pozoren na pomene, potem ne more biti povsem mehaničen – kajti pomeni (karkoli že so), ne uveljavljajo fizične sile. Po drugi strani, če manipulator ni pozoren na pomene, potem te manipu-

lacije ne morejo biti primeri mišljenja, kajti to, ali je nekaj razumno ali ne, je pomembno odvisno od pomena simbolov. Na kratko, če je proces ali sistem mehaničen, potem ne more misliti; če pa misli, potem ne more biti mehaničen« Haugland (1985: 39).

Klasična kognitivna znanost se poskuša izogniti omenjenemu paradoksu in vsaj načeloma pokazati, kako je mogoča fizična realizacija racionalnega mišljenja. Gre za pristop »od zgoraj navzdol«, ki za izhodišče jemlje poljudno psihologijo in skuša pokazati, kako s pomočjo funkcionalne razlage oziroma funkcionalne dekompozicije pridemo do mehanizma, pri katerem ni več duševnih razsežnosti. Ned Block opredeli funkcionalno razlago takole: »Funkcionalna razlaga temelji na razčlenbi sistema na njegove sestavne dele; delovanje sistema razloži v pojmih zmožnosti delov in z načinom, kako so deli med seboj povezani. Na primer to, kako tovarna izdeluje hladilnike, lahko razložimo s sklicevanjem na delovanje posameznih tekočih trakov, delavcev in strojev ter z organizacijo vsega tega« (Block 2007: 147, citirano po Markič 2011: 95). Postopek poteka tako, da kognitivno sposobnost, ki jo želimo razložiti, razstavimo na podspodobnosti, ki jih nato obravnavamo vsako posebej. Ta proces nadaljujemo, dokler ne pridemo do mehanizma v možganih.

Klasična kognitivna znanost temelji na dveh predpostavkah. Prva je prepričanje, da vsako inteligentno vedênje predpostavlja zmožnost ustrezne predstavitve sveta in da razlaga vedênja brez te predpostavke ni mogoča. Zato je eden od osrednjih pojmov klasične kognitivne znanosti pojem reprezentacije. Druga predpostavka odraža temeljno značilnost funkcionalističnega pristopa, da določanje duševnih stanj in procesov poteka neodvisno od nevrofizioloških stanj in procesov (materialne realizacije).

Filozofske temelje predstavlja računski reprezentacijski teoriji duševnosti (Fodor 1975; 1987), kombinacija funkcionalistične teorije in računski analogije (t.j. analogije med kognitivni procesi in računalniškimi programi). Funkcionalistična teorija duševnosti nam pove, kako prepoznavati in razvrščati duševna stanja. Vsako duševno stanje ali proces ima svojo vzročno vlogo znotraj sklopa vzročnih relacij. Te vključujejo vzročne dražljaje iz okolice organizma, medsebojne vzročne relacije z drugimi duševnimi stanji in procesi v organizmu (postavke poljudne psihologije, kot so prepričanja, želje, namere, itd.) in vzročne učinke, ki

se kažejo v vedênju. Na primer, kaj je tisto skupno, zaradi česar različna prepričanja uvrščamo med prepričanja, različne želje med želje in različne bolečine med bolečine. Ideja funkcionalizma je, da duševna stanja razvrščamo glede na njihovo funkcionalno vlogo, ki je enaka njihovi vzročni vlogi. Na primer, za funkcionalista je bolečina duševno stanje, katerega funkcionalna/vzročna vloga je izogibanje ali odstranitev izvora poškodbe organizma. Bolečina je prepoznana kot bolečina zaradi svoje funkcionalne vloge, ne pa zaradi določenega tipa živčnega stanja ali posebnega občutka. Za funkcionaliste posamezna duševna stanja, ki imajo isto vzročno vlogo, predstavljajo le primerke določenega tipa duševnega stanja. Gre za večvrstno realizacijo, saj je isti tip lahko realiziran v različnih živčnih stanjih. Na primer, isti tip duševnega stanja (bolečina) je lahko v različnih vrstah (npr. ljudje, psi, mačke) realiziran v različnih živčnih bazah.

Duševni procesi so v reprezentacijski teoriji obravnavani kot vzročna zaporedja uprimerjanj duševnih reprezentacij. Sosednje misli je vzročno zaporedje duševnih reprezentacij, duševne reprezentacije pa izražajo propozicije, ki so objekti misli. Avstrijski filozof Franz Brentano je stanja, ki se nanašajo na določen predmet ali situacijo in imajo vsebino, imenoval intencionalna stanja. Intencionalnost je po njegovem mnenju tista bistvena značilnost, ki razlikuje duševna stanja, stanja »o nečem«, od fizičnih stanj.

Računska reprezentacijska teorija in klasična simbolna kognitivna znanost skušata podati naturalistično teorijo intencionalnosti. Fodor in Pylyshyn sta izpostavila dve značilnosti simbolnega pristopa: (i) kombinatorna sintaksa in semantika za duševne reprezentacije, in (ii) strukturna občutljivost procesov (Fodor in Pylyshyn 1988: 12–13). Prva značilnost vzpostavlja notranji jezik misli (*language of thought*, Fodor 1975), katerega semantika je določena s pomenom osnovnih sintaktičnih gradnikov jezika in kompozicionalno semantiko. Ta za vsako možno kombinacijo določa, kako pomeni sintaktičnih delov sestavljenega izraza prispevajo k pomenu celote. Druga značilnost, strukturna občutljivost procesov, pa omogoča, da so načela, po katerih se transformirajo duševna stanja, definirana nad strukturnimi lastnostmi duševnih reprezentacij. Gre za formalne računske procese, ki so občutljivi zgolj na formo, sintakso ali

obliko simbolnih struktur, nad katerimi so definirani, ne pa na njihovo vsebino.

Po analogiji med programi in kognitivnimi procesi so duševne reprezentacije simboli, ki imajo vsebino in so hkrati vzročno odgovorni za vedénje. Računalnik je stroj, ki deluje v skladu z fizičnimi lastnostmi simbolov in nima neposrednega dostopa do njihovih semantičnih vrednosti. Kljub temu so operacije omejene s semantiko, kajti vsako pomensko razliko, ki je pomembna za program, programer zakodira v sintaksi simbolnega jezika. Če stroj sledi znotraj sistema definiranim sintaktičnim pravilom pri tvorjenju novih formul, potem bodo interpretirane formule ohranjale svojo semantično vrednost. Newell in Simon (1976) sta take sisteme (tako možgane kot računalnike) imenovala »fizični simbolni sistemi«. Njuna hipoteza in hipoteza klasične kognitivne znanosti je, da fizični simbolni sistem poseduje nujna in zadostna sredstva za izvajanje inteligentnih dejanj.

Klasična simbolna kognitivna znanost, ki temelji na analogiji med računalniškimi programi in kognitivnimi procesi, odgovarja na Descartesov izziv, da je skoraj nemogoče, da bi lahko zgradili fizični stroj, ki bi kazal razumno vedénje. Ponuja odgovor, kako je *v principu* mogoče, da je vrsta duševnih stanj (predvsem propozicionalna stanja kot so prepričanja, želje, namere...) fizično realizirana in dejansko povzroča vedénje.

### 3. Kako zavest in racionalno vedenje vznikneta kot produkta živčnega sistema in evolucije? *Pot od spodaj navzgor*

Descartesa bi lahko, kar se tiče razlage vedenja živali in razlage refleksnih procesov v človeku, označili za mehničnega redukcionista. Tisto, kar se izmika redukciji, so po njegovem sposobnost mišljenja, jezikovnega sporazumevanja, hotene in namerne dejavnosti ter zavedanje. Teh procesov naj ne bi bilo mogoče razložiti, če ne sprejmemo dualističnega stališča, po katerem sta duh in telo v medsebojni vzročni zvezi. Vendar zagovorniki evolucijske teorije tako velikega prepada med živalmi in človekom niso priznavali. Zato so iskali razlage, kako z naravnimi procesi razložiti omenjene sposobnosti človeka. Te naj bi se pojavile kot posledica večje kompleksnosti možganov in naj bi bile emergentne lastnosti živčnega sistema. Biolog Thomas Huxley (1825–1895), ki je bil velik za-

govornik teorije evolucije, je ob koncu 19. stoletja na podlagi takrat znanih dejstev fiziologije živali in človeka zagovarjal teorijo kontinuiranosti.

»Nauk o kontinuiranosti je preveč trdno uveljavljen, da bi si drznil predpostaviti, da neki kompleksen naravni pojav nastane iznenada in ne da bi mu predhodile preprostejše modifikacije; prav tako bi potrebovali zelo močne argumente, da bi dokazali, da se tako zapleten pojav, kot je zavest, prvič pri človeku. [...] Prav tako vemo, da nižje živali posedujejo, dasiravno v manj razviti obliki, tisti del možganov, za katerega upravičeno domnevamo, da je pri človeku organ zavesti; in da imajo živali, čeprav morda ne posedujejo naše stopnje zavesti in čeprav zaradi nezmožnosti rabe jezika ne morejo imeti miselnih, marveč le čustvene tokove, vendarle zavest, ki v večji ali manjši meri napoveduje našo zavest.« (Huxley 2007: 207)

Čeprav je zanikal Descartesovo pojmovanje živali kot nezavednih strojev, pa je sprejemal tezo, da so živali »bolj ali manj zavedni avtomati«, ki jih vodi instinkt in ne razum. Posedovali naj bi poseben aparat, ki bi lahko priklical občutke, čustva in ideje. Po Huxleyju spremembe v možganih povzročajo zavestna stanja, ne pa tudi obratno.

»Zdi se, da zavest živali nastane le kot stranski produkt delovanja mehanizma njihovega telesa in da ji manjka vsakršne moči, da bi to delovanje predrugačila, podobno kot zvok piščali na paro, ki spremlja delo parnega stroja lokomotive, ne more vplivati na ustroj sam. Njihova volja, če jo sploh imajo, je čustvo, ki priča o fizičnih spremembah in ne vzrok tovrstnih sprememb.« (Huxley 2007: 209)

Podobno kot pri živalih pa tudi pri človeku Huxley ne najde dokaza, da bi zavestna stanja povzročala spremembe v možganih. Po njegovem mnenju so zavestna stanja le epifenomeni, »simboli, ki jih v zavesti gojimo o spremembah, do katerih v organizmu prihaja spontano« [in] »občutek, ki ga imenujemo volja, ni vzrok hotenega dejanja, temveč simbol možganskega stanja, ki je neposredni vzrok tega dejanja« (Huxley 2007: 210).

Od Huxleyjevih razmišljanj v članku *O hipotezi, da so živali avtomati, in njeni zgodovini* je minilo že več kot sto let. V tem času, posebej še v zadnjih dvajsetih letih, smo bili priče velikemu napredku v raziskovanju možganov. Z novimi slikovnimi metodami kot so elektroencefalografija (EEG), računalniška tomografija (CT), slikanje z magnetno resonanco (MRI), pozitronska emisijska tomografija (PET) in slikanje z funkcio-

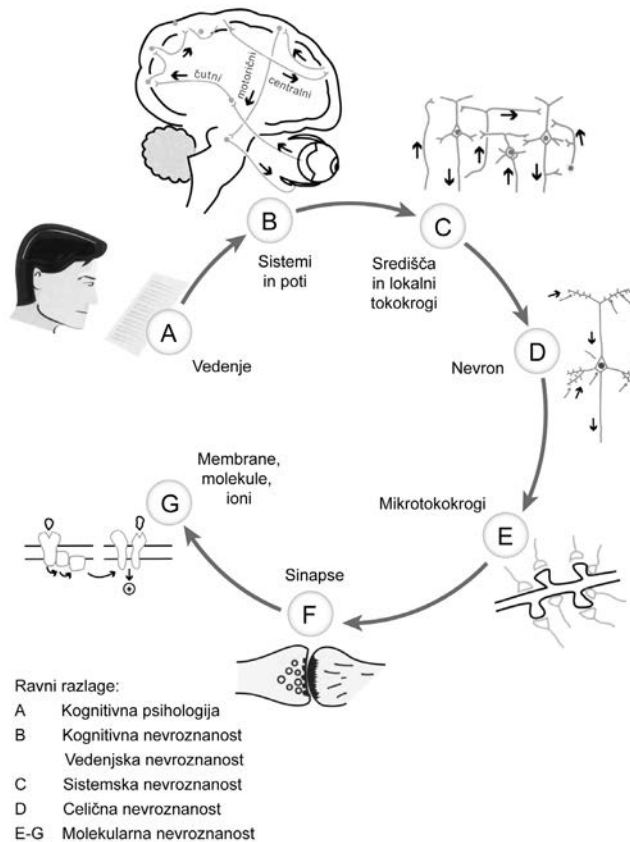


nalno magnetno resonanco (fMRI) so nevroznanstveniki dobili orodja, s katerimi lahko »pokukajo« v možgane, včasih tudi med samim delovanjem. Na ta način je vsaj do neke mere omogočeno spremljanje možganskih procesov v realnem času. Meritve lahko potekajo med različnimi nalogami, ki jih takrat opravlja človek, ali pa v časovnem obdobju, za katerega je človek poročal, da je imel določene zavestne izkušnje (npr. Damasio 2008).

Idejo, da bi v možganih našli predel, ki bi bil sedež racionalne duše, so sodobne raziskave v nevroznanosti zavrnil. Daniel Dennett (1991) je tak pogled, ki trdi, da obstaja področje v možganih, v katerem je shranjena vsebina zavestnega izkustva, poimenoval kartezijanski materializem. Po njegovem gre le za prenos Descartesovega dualizma v materializem. Dejansko pa procesiranje v možganih poteka vzporedno v različnih predelih možganov. Sodobna nevroznanost skuša integrirati različne ravni razlage, ki sledijo različnim ravnam organizacije v možganih in živčnem sistemu. Kratko predstavitev in sliko povzemam po Markič (2011: 89).

Najvišjo raven predstavlja organizacija v možganih, ki jo opisujemo z izrazi nevronskega sistema in nevronske poti (B). Z njo se ukvarjata kognitivna in vedenjska nevroznanost, ki preučujeta večje sisteme v možganih, ki se uporabljajo v dejavnostih na kognitivni ravni. Sistemski nevroznanost raziskuje delovanje nevronskega sistema. Nevronski sistemi so določeni glede na lokacijo v možganih (npr. primarni vizualni korteks), poti pa omogočajo širjenje informacij po možganih (npr. dorzalna pot, če se informacije širijo po zgornjem delu možganov in ventralna pot, če se širijo po spodnjem delu možganov). Aktivnost na tej ravni organizacije je rezultat aktivnosti na nižji ravni organizacije, na ravni središč, lokalnih krogotokov in mikrotokokrogov (C in E). Skupinska aktivnost populacije nevronov kodira informacije tako, da organizira in usklajuje informacije posameznih nevronov. Raven lokalnih tokokrogov je zelo kompleksna in najmanj razumljena raven v živčnem sistemu. Nevroznanstveniki imajo orodja in tehnike, na primer funkcionalno slikanje, za raziskovanje obnašanja nevronskega sistema. Prav tako so razvite tehnike snemanja posameznih nevronov, ni pa primerljivih načinov za neposredno raziskovanje populacij nevronov. Aktivnost populacije nevronov je funkcija obnašanja posameznih nevronov. Povezovanje med delovanjem nevronskega sistema in delovanjem posameznih

nevronov je ena osrednjih tem računske nevroznanosti. Vendar nevroni ne predstavljajo najbolj osnovne ravni. Za razumevanje njihovega delovanja moramo razumeti, kako med seboj komunicirajo preko sinaps (F). Nekatere sinapse so električne, večina pa je kemičnih. Te delujejo preko neurotransmiterjev, ki jih sproži električni signal (aktivacijski potencial). Prenašanje neurotransmiterjev je določeno z molekularnimi lastnostmi sinaptične membrane. To so lastnosti, ki so določene genetsko (G). Z biološkimi lastnostmi nevronov se ukvarjata celična in molekularna nevroznanost.



Slika 1: Ravni organizacije in ravni razlage v živčnem sistemu (Shepherd 1994, prirejeno po Bermúdez, 2010: 96, povzeto po Markič 2011: 89)

Na podlagi vedno bolj podrobnega znanja o ravneh od G do B so različni avtorji predlagali teorije, kako na mehaničen način, to je s sklicevanjem na procese v živčnem sistemu, pojasniti zavestno in razumno vedenje.

Ameriški filozof Ned Block (1995) je poudaril, da zavest ni enoten pojem, in izpostavil dva pomena: *zavest kot dostop in fenomenalno zavest*. Prva obravnava duševno stanje *D* kot dostopno zavesti, v kolikor je vsebina *D* lahko uporabljena kot premisa razmišljanja, ali za racionalni nadzor govora ali dejanja. Lahko gre za stanja, v katerih se človek trenutno nahaja, ali pa za spomine. Po drugi strani pa je duševno stanje *D* fenomenalno zavestno zaradi fenomenalno zavestnih lastnosti, kot so »načini, kako občutimo doživljanje«, »kako je biti ...«, »kvalitativni značaj«, *qualia*, »neposredne fenomenološke kvalitete«, itd. Gre za subjektivno zavedanje, občutenje, doživljanje, kot so npr. doživetja ob gledanju, poslušanju, dotikanju, vonjanju, čustvih itd.

Empirični pristopi se ukvarjajo predvsem z raziskovanjem funkcionalne in vzročne vloge zavestnih duševnih stanj. Primer takega pristopa je teorija globalnega delovnega prostora (*Global Workspace Theory*) Bernarda Baarsa (1988, 1997). Baars ponazori razliko med zavestnim stanjem in nezavednimi procesi z metaforo gledališča. V »gledališču zavesti« je zavestno stanje tista točka na odru delovnega spomina, na katero posije svetloba pozornosti, ostalo gledališče pa je v temi. Igralci se premikajo in komunicirajo v temi (nezavedno) in le včasih jih oplazi snop svetlobe na odru. Takrat nezavedna stanja postanejo zavestna. Baarsovo prispodobo razumem tako, kot da so zavestna stanja izenačena s posebnimi, »s svetlobo pozornosti« osvetljenimi možganskimi stanji, njihova vzročna vloga pa je vzročna vloga teh posebnih možganskih stanj.

Baars pravi, da imajo vsi elementi teorije globalnega delovnega prostora verodostojne interpretacije v možganih, kar omogoča postavljanje možganskih hipotez o zavesti in njenih mnogih vlogah v možganih, kot tudi izdelavo modelov z nevronskimi mrežami. Znanstveniki zdaj te hipoteze lahko preizkušajo in Baars pravi, da so doživele precejšnjo empirično podporo (Baars 2003). Teoretične hipoteze so, na primer: zavestni dogodki omogočajo skoraj vse vrste učenja: zavestna zaznavna povratna zanka omogoča hoten nadzor nad motoričnimi funkcijami in morda celo nad populacijami nevronov in posameznimi nevroni; zavestne vse-

bine lahko sprožijo izbrano pozornost itd. Zanimiv primer za slednje je pojav nenamerne slepote, ko zaradi usmerjene pozornosti ne opazimo nekaterih sicer očitnih sprememb. Primer zabavne ilustracije si lahko ogledate na strani *Quircology* psihologa Richarda Wisemana.

Po razdelitvi Davida Chalmersa (1995, 1996) bi Baarsovo teorijo verjetno uvrstili med pristope, ki rešujejo »lahke« probleme zavesti. Tako je Chalmers označil probleme, ki so takšne narave, da so dostopni znanstvenemu raziskovanju (ne pa, ker bi bili dejansko lahko rešljivi). V to skupino med drugim sodi iskanje odgovorov na vprašanja, kot so: Kako človek razlikuje čutne dražljaje in na njih reagira? Kako možgani usklajujejo informacije iz različnih virov in uporabijo te informacije za nadzor vedenja? Na kakšen način oseba ubesedi svoja notranja stanja? Raziskovanja povezana s temi in podobnimi vprašanji predstavljajo sicer velik izziv za znanstvenike, a po Chalmersovem mnenju ne morejo dati zadovoljivega odgovora na najbolj zanimivo vprašanje: »Kako lahko vznikne zavest iz nevronskega procesa v možganih?«. Huxley je že leta 1866 to zadrego ponazoril takole: »Kako lahko nekaj tako izrednega, kot je stanje zavesti, nastane kot posledica draženja živčnega tkiva, je prav tako nedojemljivo, kot je nedojemljiv pojav duha, potem ko Aladin podrgne po svoji svetilki.« Vprašanje, kako lahko fizični proces v možganih privede do subjektivnih doživetij, je tisto, kar po Chalmersu predstavlja »težak« problem zavesti.

Poskusa oblikovanja teorije, ki se ukvarja predvsem s fenomenalnimi vidiki zavesti in daje odgovor na težko vprašanje, se je lotil nevroznanstvenik Gerald Edelman (2006). Predstavlja teorijo, kako je zavest nastala skozi evolucijo in individualni razvoj določenih vrst. Po njegovem mnenju je zavest posledica (v smislu logičnega sledenja; ang. *is entailed by*) povratnih aktivnosti med kortikalnimi področji in talamusom ter interakcij znotraj korteksa in interakcij korteksa in s subkortikalnimi strukturami. Primarna zavest je proces, ki ga sestavlja veliko različnih takšnosti (*qualia*); razlikovanja, ki so posledica široko porazdeljenih in visoko dinamičnih aktivnosti talamokortikalnega jedra (Edelman 2006: 37). V primeru primarne zavesti sta zavedanje in zavestno načrtovanje še omejeni na zapomnjeno sedanost. V neki fazi evolucije višjih primatov so se nato razvile poti, ki so omogočile povratne povezave med pojmovnimi zemljevidi možganov in tistimi predeli, ki so zmožni simbolnega

ali semantičnega nanašanja. Zavedst višjega reda pa je morala počakati tisti čas v evoluciji človeka, ko se je pojavil jezik. Takrat zavest ni bila več omejena na zapomnjeno sedanost. Zavedst zavesti je postala mogoča (*ibid.* 38). Naj poudarim nekatere značilnosti Edelmanovega pojmovanja zavesti. Prvič, zavest je proces in ne stvar. Zato je po njegovem mnenju vprašanje, ali takšnosti obstajajo, napačno vprašanje. Drugič, vzročno moč imajo aktivnosti v talamokortikalnem jedru in ne fenomenalno izkustvo, ki je logična posledica teh aktivnosti. Tretjič, zavest je proces, ki ga moramo najprej subjektivno izkusiti, preden ga lahko razložimo. Zavedst je namreč nujno privatna, saj je posledica povratnih aktivnosti v talamokortikalnem jedru individuuma. Edelman se tako pridružuje Huxleyju in sodobnim skeptikom (npr. Wegner 2002), za katere je prepričan, da zavest lahko povzroča, da se zgodijo stvari, ena izmed koristnih iluzij. Vendar naj to ne bi pomenilo, da zavest ne počne ničesar. Kot pravi Edelman, nas informira o naših možganskih stanjih in zato zavzema osrednje mesto pri razumevanju samega sebe (Edelman 2006: 92).

V grobem bi lahko rekli, da teorije zavesti izhajajo iz dveh predteoretskih intuicij, ki ju je Güven Güzeldere povzel v naslednjih dveh trditvah: (i) Zavedst je to, kar zavest počne. (*Consciousness is as consciousness does.*) (ii) Zavedst je to, kar se zavesti dozdeva. (*Consciousness is as consciousness seems.*) (Güzeldere 1995: 36).

Pri prvi gre za vzročno intuicijo, ki kot temeljno za naše duševno življenje jemlje vzročno vlogo zavesti. Pri drugi, fenomenalni intuiciji, pa je poudarek na tem, kako občutimo in izkusimo naše duševno življenje, t.j., gre za pojavne kvalitete, ki označujejo naše zaznave, bolečine in druga mentalna stanja.

Ti dve intuiciji pogosto vlečeta vsaka v svojo smer. Zato se zgodi, da tisti, ki gradijo na osnovi ene intuicije, zanemarijo drugo in ne zajamejo njenih značilnosti. Güzeldere je prepričan, da diskusije pogosto zaidejo v slepo ulico zaradi še neke tretje, po njegovem mnenju napačne intuicije. To je intuicija bistva (esencializma). Ta intuicija pravi naslednje: če označujemo zavest kot bistveno vzročno, potem mora biti bistveno ne-fenomenalna in obratno, če je bistveno fenomenalna, potem je bistveno ne-vzročna. Güzeldere sam se zavzema za nasprotno intuicijo, po kateri to, kar zavest počne, ne moremo določiti, če je odsotna fenomenalna

zavest, oziroma še bolj pomembno, »kar se zavesti dozdeva« ne moremo pojmiti ob odsotnosti tega, »kar zavest počne« (*ibid.*).

Čeprav Baarsova in Edelmanova teorija poudarjata porazdeljenost procesov po različnih delih možganov in ne določata posebnega predela kot sedeža zavesti, pa na neki način še vedno izhajata iz Descartesove dediščine. Zavest poskušata pojasniti kot možganski proces.

#### 4. Težave in obet rešitve

Na prvi pogled se morda zdi, da so težave, s katerimi se srečujejo raziskovalci na eni ali drugi poti, zgolj praktične narave. Na primer, za težave so krive tehnološke omejitve računalnikov, ko pa bomo imeli močnejše računalnike, bomo lahko naredili verodostojen model mišljenja. Ali pa, raziskovalne tehnike nam še ne omogočajo, da bi dovolj natančno spoznali delovanje živčnega sistema na različnih ravneh, a ko bomo razvili nove metode in orodja, nam bo nevroznanost razložila odločanje.

Mnogi znanstveniki in filozofi pa so opozorili, da so težave globlje in da so posledica posameznih predpostavk ene ali druge poti. Naj jih nekaj naštejemo. Na primer, glede klasične simbolne kognitivne znanosti so se že kmalu pojavili skeptični ugovori. John Searle (1990) je z miselnim eksperimentom »Kitajska soba« opozoril na problematičnost omejevanja razumevanja mišljenja na zgolj manipulacijo s simboli. Postavil je tudi vprašanje, na kakšen način dobijo pomen posamezni simboli. Če jim ga pripiše programer, potem ne gre za izvorno intencionalnost. Podobno vprašanje o »ozemljitvi simbolov« je postavil tudi Harnad (1990). Eden prvih odmevnih kritikov Hubert Dreyfus je v knjigi *What Computers can't do* kritiziral domnevo, da je mišljenje simbolna manipulacija, ki poteka v skladu s pravili. Dokazoval je, da je tak pristop preveč tog in da nikoli ne bo mogel uspešno modelirati veččin, ki so del vsakdanjega znanja (Dreyfus & Dreyfus 1990). Izkazalo se je tudi, da zgolj zanašanje na propozicijsko znanje, ki ga je mogoče predstaviti s simbolnimi strukturami, ne more dati odgovora na vprašanje, katere informacije so relevantne v dani situaciji (problem relevance). Ljudje imamo neverjetno sposobnost, da hitro vidimo relevantne posledice določenih sprememb v dani situaciji. Dojamemo, kaj se dogaja in se hitro odzovemo s pravimi rešitvami. Eden od problemov klasičnega simbolnega pristopa je ravno

v tem, kako preprečiti nepomembna in časovno potratna sklepanja. Ali je rešitev znotraj danega pristopa sploh mogoča? Predlog za radikalno rešitev, ki zapušča zgolj razumski okvir klasične simbolne paradigme, sta dala Antonio Damasio (1994) in Ronald de Sousa (2007). Po njunem si pri težavah lahko uspešno pomagamo s hipotezo o vlogi čustev. Čustva naj bi v procesu racionalnega izbiranja omogočila, da postane relevanten samo majhen delež možnih alternativ in dejstev. Proces odločanja se odvija tako na zavestni kot na nezavedni ravni, zato za razlago ni dovolj poznavanje zgolj zavestne ravni (Markič 2009).

Na drugi strani pa kognitivni nevroznanstveniki niso dosledni, ko interpretirajo svoja raziskovanja. Včasih so zelo previdni in se skušajo izogniti metafizičnim stališčem glede odnosa med zavestnimi in možganskimi stanji. Zato se pogosto zatekajo k uporabi izraza »korelat« in govorijo o nevronskih oziroma živčnih korelatih zavesti (Koch 2004). Ker pa v tekstih pogosto govorijo kar o zavesti in zavestnih stanjih, dobimo občutek, kot da raziskujejo zavestne procese. Možnost ekvivokacije je zato precejšnja.

Dennett (1991) ter Bennett in Hacker (2003) ugotavljajo, da kognitivni nevroznanstveniki pogosto zgolj zamenjajo duševnost z možgani. Ob tem opozarjajo na pogosto neustrezno rabo vsakdanjih psiholoških pojmov pri opisovanju funkcij možganov ali živčnih podsistemov. To lahko privede do napačnega postavljanja vprašanj in do napačnih interpretacij eksperimentalnih rezultatov. Po njihovem mnenju lahko uporabljamo psihološke predikate samo, kadar se nanašamo na osebo (ali žival), ne pa, kadar se nanašamo na njene možgane ali predele možganov. Tako na primer rečemo: »človeka boli zob« ali »človek vidi in sliši«. Ne moremo pa reči: »možgane boli zob« ali »možgani vidijo in slišijo«. Ali celo: »leva hemisfera vidi in sliši« (Bennett in Hacker 2003: 73). Bennett in Hacker sta napako poimenovala »mereološka napaka« v nevroznosti (*ibid.*), Dennett pa je, da bi se izognil takim napakam, razlikoval osebno in podosebne ravni razlage.

Na tem mestu se ne bom ukvarjala s predlogi, kako izboljšati ali izbrati alternativni pristop, ki bi bil uspešen na prvi poti (o tem v Markič 2011), niti z empiričnimi teorijami kognitivne nevroznosti. Menim namreč, da se poleg težav, ki so specifične za vsako od poti, kažejo še težave, ki postavljajo pod vprašaj predpostavke obeh. Zame je zid, v ka-

terega sem vedno znova trčila, predstavljal problem mentalne vzročnosti in argument Jaegwona Kima (1993, 1999) o vzročni izključitvi, saj nobena rešitev ni bila zadovoljiva. Predpostavka, ki je morda razlog za težave, pa je naslednja: duševni procesi so realizirani v možganih; oziroma, če pogledamo iz druge smeri: duševni procesi vzniknejo iz možganov. V današnjem času hitrega razvoja nevroznanosti in fascinantnih raziskav, ki jih lahko spremljamo tudi v popularnem tisku, se zdi ideja na prvi pogled morda nora. Ker pa je reševanje težav še vedno znotraj okvirov naturalizma, možgani seveda ohranjajo pomembno vlogo, čeprav niso več edini igralec. Zavest in duševni procesi so utelešeni v možganih, telesu in v svetu – tako kot pravi Teed Rockwell v naslovu svoje knjige: *Niti možgani niti duh (Neither Brain nor Ghost, 2007)*.

## L i t e r a t u r a

1. Baars, B. J. (1988). *A Cognitive Theory of Consciousness*. Cambridge, MA: Cambridge University Press.
2. Baars, B. J. (1997). *In the Theater of Consciousness*. New York, NY: Oxford University Press.
3. Baars, B. J. (2003). »The Global Brainweb: An update on global workspace theory«. Guest editorial, *Science and Consciousness Review*, October 2003. (dostopno na [http://cogweb.ucla.edu/CogSci/Baars-update\\_03.html](http://cogweb.ucla.edu/CogSci/Baars-update_03.html))
4. Bennett, M.R., Hacker, P.M.S. (2003). *Philosophical Foundations of Neuroscience*. Malden, MA, Oxford: Blackwell Publishing.
5. Block, N. (1995). »On a Confusion about a Function of Consciousness«, *Behavioral and Brain Sciences* 18 (2), str. 227–287.
6. Bermúdez, J.L. (2010). *Cognitive Science: An Introduction to the Science of the Mind*. Cambridge: Cambridge University Press.
7. Chalmers, D. (1995). »Facing up to the Problem of Consciousness«, *Journal of Consciousness studies*, 3 (1), str. 200–219.
8. Chalmers, D. (1996). *The Conscious Mind*. Oxford: University of Oxford Press.
9. Cottingham, J. G. (1995). »Cartesian Dualism: Theology, Metaphysics, and Science«. V Cottingham, J. G. (ur.), *The Cambridge Companion to Descartes*. Cambridge: Cambridge University Press.
10. Crane, T. (1995). *The Mechanical Mind*. London: Penguin Books.
11. Damasio, A. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain*. New York: G.P. Putnam's Sons.



12. Damasio, A. (2008). *Iskanje Spinoze: Veselje, žalost in čuteči možgani*. Ljubljana: Krtina.
13. Dennett, D. (1991). *Consciousness Explained*. London in New York: Penguin Books.
14. Descartes, R. (1957). *Razprava o metodi*. Ljubljana: Slovenska matica.
15. De Sousa, R. (2007). *Why think? Evolution and the Rational Mind*. Oxford, New York: Oxford University Press.
16. Dreyfus, H.L. and Dreyfus, S.E. (1990). »Making a Mind Versus Modeling the Brain. Artificial Intelligence Back at the Branch Point«. V Boden, M. (ur.), *Philosophy of AI*. Oxford: Oxford University Press.
17. Edelman, G.M. (2006). *Second Nature*. New Haven and London: Yale University Press.
18. Fodor, J. (1975). *The Language of Thought*. New York: Crowell.
19. Fodor, J. (1987). *Psychosemantics*. Cambridge, MA, London: The MIT Press.
20. Fodor, J.A. and Pylyshyn, Z.W. (1988). »Connectionism and Cognitive Architecture: A Critical Analysis«. *Cognition* 28, str. 3–71.
21. Güzeldere, G. (1995). »Consciousness: What it is, how to study it, what to learn from it's history«, *Journal of Consciousness studies*, 2 (1), str. 30–51.
22. Huxley, T.H. (2007). »O hipotezi, da so živali avtomati in njeni zgodovini«. V O. Markič, J. Bregant (ur.) *Narava mentalnih pojavov*. Maribor: Aristej, str. 197–211.
23. Harnad, S. (1990). »The Symbol Grounding Problem«. *Physica D* 42, str. 335–346.
24. Haugeland, J. (1985). *Artificial Intelligence: The very idea*. Cambridge, MA., MIT Press.
25. Kim, J. (1993). *Supervenience and Mind: Selected Philosophical Essays*. Cambridge: Cambridge University Press.
26. Kim, J. (1996), *Philosophy of mind*. Boulder, Co. and Oxford: Westview Press.
27. Kim, J. (1999). »Making sense of emergence«. *Philosophical Studies* 95, str. 3–36.
28. Koch, C. (2004). *The Quest for Consciousness, Englewood*. Roberts and Company Publishers.
29. La Mettrie J. (1996). *Machine Man and Other Writings*. A. Thomson (ur.). Cambridge: Cambridge University Press.
30. Markič, O. (2009). »Odločanje: racionalnost in čustva«. V: Ule, A., Markič, O. Kordeš, U. (ur.), *Konteksti odločanja*. Maribor: Aristej, str. 17–33.
31. Markič, O. (2011). *Kognitivna znanost: Filozofska vprašanja*. Maribor: Aristej.
32. Rockwell, T. (2007). *Neither Brain nor Ghost*. Cambridge, MA, London: The MIT Press.
33. Searle, J. (1990). »Duhovi, možgani in programi«. Prevedeno v Hofstadter, D. in Dennett, D. (ur.), *Oko duha*. Ljubljana: Mladinska knjiga, str. 361–389.

34. Shankar, S. (2004). »Descartes' Legacy: The Mechanist/Vitalist Debates«. V Shankar, S. (ur), *Philosophy of Science, Logic and Mathematics in the Twentieth Century*. London, New York: Routledge.
35. Wegner, D. (2002). *The Illusion of Conscious Will*. Cambridge, MA, London: The MIT Press.
36. Wiseman, R. [http://www.quirkology.com/USA/Video\\_ColourChangingTrick.shtml](http://www.quirkology.com/USA/Video_ColourChangingTrick.shtml).