Teadvusega vaim

 Teel fundamentaalse teooria poole

David Chalmers

#### Ilmumisandmed

David Chalmers. ​*The Conscious Mind: In Search of a Fundamental Theory*​. Oxford University Press 1996. ISBN 0-19-510553-2.

PDF Internetis: [siin](https://personal.lse.ac.uk/ROBERT49/teaching/ph103/pdf/Chalmers_The_Conscious_Mind.pdf).

### ​**[lk.v]**​ Tänuavaldused

Teadvus ja keha–vaimu probleem hakkasid mind köitma matemaatikaõpingute ajal Adelaide’i ülikoolis. Jutuajamised terve hulga inimestega, eriti Paul Barteri, Jon Baxteri, Ben Hambly ja Paul McCanniga, aitasid mul oma ideedele kuju anda. Juba siis tundus see teema vaat et kõige põnevama probleemina üldse. Tundus lausa uskumatu, et inimesele võidakse anda täisajaga töökoht mõtisklusteks nii põneval teemal.

Hiljem, Oxfordi aspirandina, hakkasin märkama, et minu mõtted tegelevad matemaatika asemel kogu aeg teadvusega, ning ma otsustasin vahetada eriala ja lõpuks ka kontinenti. Paljud inimesed toetasid mind sel raskel ajal kannatlikult, eriti Michael Atiyah, Michael Dum­mett ja Robin Fletch­er. Olen tänulik ka kõigile, kes võtsid vaevaks kuulata minu heietusi oma järjekordsetest teadvuseteooriatest; ideed, mida ma siin raamatus arendama hakkan, on nonde mõttearenduste kauged järglased.

Üleminek Indiana ülikooli, et saada põhi alla filosoofias, kognitiivteaduses ja tehisintellektis, oli minu elu parimaid samme. Erilist tänu võlgnen Doug Hofstadterile: just tema kirjatööd olid esimesed, mis mulle juba noorukipõlves teadvuse saladusi tutvustasid, ning tema uurimislabor — Mõistete ja Kognitsiooni Keskus [​*Center for Research on Concepts and Cognition*​, CRCC] — oli nende ideede arenemiseks innustav ja õdus keskkond. Kuigi ta ei ole paljude käesolevas raamatus esitatud ideedega nõus, meeldiks mulle mõelda, et mingis mõttes on minu kirjatööd jäänud truuks tema tööde intellektuaalsele vaimule.

Käesoleva raamatu esimene versioon (mille pealkiri oli tollal “Teel teadvuseteooria poole” [​*Toward a Theory of Consciousness*​]) valmis 1992. ja 1993. aastal kuus kuud kestnud tihedas töös. Selle aja paiku oli mul Indiana ülikoolis hulk kasulikke diskussioone paljude inimestega: kõigi CRCC töötajatega, eriti Bob Frenchi ja Liane Gaboraga, ning paljude inimestega teistest teaduskondadest, seal­hul­gas Mike Dun­ni, Rob Gold­sto­ne’i, Anil Gup­ta, Jim Hett­me­ri, Jer­ry Se­lig­ma­ni ja Tim van Gelderiga. Tänan ka teadvusalase väitlusklubi liikmeid paljude meeldivate vestluste eest “Nick’si”[[1]](#footnote-2) tagatoas esmaspäeva pärastlõunati.

Innustav keskkond oli ka Washingtoni ülikoolis, kus ma õppisin kaheaastase McDonnelli stipendiumi toel filosoofiat, neuroteadust ja psühholoogiat; seal sain ma käesolevat raamatut ​**[lk.vi]**​ lõpetades omal nahal tunda ka Zenoni paradoksi. Olen tänulik James S. McDonnelli fondile toetuse eest, kõigile minu teadvusalaste aspirantuuriseminaride osalistele arutelude eest, mis aitasid raamatut viimistleda, ning paljudele inimestele, sealhulgas Mor­ten Chris­tian­se­ni­le, An­dy Clar­ki­le, Ja­son Cle­ven­ge­ri­le, Peg­gy Des­Au­tel­si­le, Pe­pa To­ri­bi­o­le ja Tad Zawidzkile, vestluste ja kommentaaride eest.

Paar viimase aastaga on mulle nii kõnes kui kirjas antud tohutult palju kasulikke mõtteid selle raamatu sisu kohta. Paljude teiste seas väärivad tänu Jon Bax­ter, Ned Block, Alex Byrne, Fran­cis Crick, Dan Den­nett, Eric Diet­rich, Avi Elit­zur, Matt­hew El­ton, Owen Fla­na­gan, Stan Frank­lin, Lia­ne Ga­bo­ra, Gu­ven Gii­zel­de­re, Chris Hill, Ter­ry Hor­gan, Steve Horst, Frank Jack­son, Jaeg­won Kim, Chris­tof Koch, Mar­tin Lec­key, Dave Lei­sing, Ker­ry Le­ven­berg, Joe Le­vine, Da­vid Lew­is, Bar­ry Loe­wer, Bill Ly­can, Paul Mc­Cann, Da­ryl Mc­Cul­lough, Brian Mc­Laugh­lin, Tho­mas Met­zin­ger, Ro­bert Mil­ler, And­rew Milne, John O’Lea­ry-Haw­thorne, Jo­seph O’Rourke, Cal­vin Ost­rum, Rhett Sa­vage, Aa­ron Slo­man, Leo­pold Stu­ben­berg ja Red Watson. Neid, kellele ma võlgnen tänu huvitavate jutuajamiste eest teadvuse üle üldiselt, on liiga palju, et neid siin ükshaaval loetleda. Eraldi tahaksin avaldada tänu Norton Nelkinile, kellelt ma sain oma käsikirja tagasi rohkete kasulike märkustega veidi aega enne seda, kui ta suri lümfoomi. Me jääme temast veel kauaks puudust tundma.

Oma üldisemate filosoofiliste vaadete kujunemise eest võlgnen tänu paljudele inimestele. Esialgsed seisukohad teadvuse kohta arendasin ma välja suuresti iseseisvalt, kuid sellealase kirjanduse lugemine on neid oluliselt rikastanud. Peaaegu iga kord, kui meile tuleb uus idee, avastame peagi, et keegi on seda väljendanud juba enne meid. Uuemaaegsetest mõtlejatest on eriti Thomas Nagel, Frank Jackson ja Joseph Levine aidanud juhtida tähelepanu raskustele teadvuse käsitlemisel; nende töö katab umbes sama territooriumi nagu minu raamatu esimesed peatükid. Mitmes punktis kattub minu töö ka Ned Blocki, Robert Kirki ja Michael Lockwoodi töödega. Metafüüsiline mudel, mida ma arendan 2. peatükis, toetub paljus Terry Horgani, Saul Kripke, David Lewise jt. töödele, ning Frank Jackson on minust sõltumatult välja töötanud sarnase mudeli, mida ta tutvustas oma suurepärastes loengutes John Locke’i​le pühendatud loengusarja raames​[[2]](#footnote-3) 1995. aastal. Ikka ja jälle on mulle pakkunud innustust ja väljakutseid Daniel Dennetti, Colin McGinni, John Searle’i ja Sydney Shoemakeri ideed.

Kõige rohkem võlgnen tänu järgmistele isikutele: Gregg Rosenbergile meeldejäävate vestluste ja väärtusliku tagasiside eest; Lisa Thomasele sombisid käsitleva raamatu ja moraalse toe eest; Sharon Wahlile asjatundliku toimetajatöö ja sooja sõpruse eest; ning eelkõige oma vanematele — kõigile kolmele — abi ja toetuse eest. Alalise inspiratsiooni eest tänan ka kõiki oma kvaale ja nende sünnikeskkonda.

Kui ma olin just raamatut lõpetamas, sattus mulle restoranis õnneküpsis​[[3]](#footnote-4) sõnadega “Sinu elu on tulvil vaimustavaid saladusi”. See on seniajani tõepoolest nii olnud, ja ma olen selle eest väga tänulik.

## ​**[lk.xi]**​ Sissejuhatus: Võtame teadvust tõsiselt

Teadvus on tohutu mõistatus ning võib-olla üldse kõige suurem kõigist tõketest meie teel maailma teadusliku mõistmise poole. Füüsikateadus ei ole veel valmis, kuid on kaunis selge; bioloogia on lahti mõistatanud paljud iidsed saladused elu olemuse ümbert. Nende valdkondade mõistmises on lünki, kuid me ei pea nende täitmist võimatuks — meil on mingi ettekujutus, kuidas nende probleemide lahendus võiks välja näha, tarvis on vaid üksikasjad paika saada.

Isegi vaimuteaduses on saavutatud palju edu. Uuemad tööd kognitiiv- ja neuroteaduse alal aitavad meil inimese käitumist ja selle tõukejõude üha paremini mõista. Muidugi ei ole meil veel kuigi palju üksikasjalikke kognitsiooniteooriaid, kuid üksikasjad pole kättesaamatus kauguses.

Seevastu teadvus on sama mõistatuslik nagu ennegi. Asjaolu, et käitumise aluseks olevate põhjuslike protsessidega kaasneb subjektiivne siseelu, on endiselt läbinisti salapärane. Meil on küllaldaselt põhjust uskuda, et teadvus kasvab välja füüsilistest süsteemidest, näiteks ajust, kuid meil pole kuigivõrd aimu, kuidas see toimub või miks teadvus üldse eksisteerib. Kuidas saab füüsiline süsteem, näiteks aju, olla ka läbielamise subjekt? Miks peaks olema olemas “tunne olla” niisugune süsteem? Tänapäeva teaduslikud teooriad peaaegu ei puudutagi tõeliselt raskeid küsimusi teadvuse kohta. Vähe sellest, et meil pole üksikasjalikku teooriat; meil pole isegi vähimatki aimu, missugune on teadvuse koht looduse üldises korras.

Viimastel aastatel on teadvuse kohta ilmunud palju raamatuid ja artikleid, ja võiks tunduda, et me liigume edasi. Ent lähemalt vaadates selgub, et enamik neist töödest vaatab teadvuse kõige raskematest probleemidest sootuks mööda. Sageli käsitletakse neis teadvuse nii-öelda lihtsaid probleeme: Kuidas aju töötleb keskkonnast saabuvaid ​**[lk.xii]**​ stiimuleid? Kuidas ta seob informatsiooni ühtseks tervikuks? Kuidas me moodustame raporteid oma seesmiste seisundite kohta? Need on kõik olulised küsimused, kuid nendele vastamine ei lahenda rasket probleemi: miks kaasneb kogu selle töötlusega elamuslik siseelu? Mõnikord vaadatakse sellest küsimusest täiesti mööda, mõnikord jäetakse ta paremaid aegu ootama, mõnikord aga teatatakse, et lahendus on juba olemas. Ent kõigil juhtudel jääb lugejale tunne, et keskne küsimus on ikka sama mõistatuslik nagu ennegi.

Me ei peaks selle mõistatuse üle ahastama, vaid võtma teadvuse probleemi kui üht meie aja kõige põnevamat intellektuaalset väljakutset. Kuna teadvus on nii fundamentaalne nähtus, mida me samas nii halvasti mõistame, siis võib selle probleemi lahendus avaldada väga sügavat mõju meie arusaamisele universumist ja iseendist.

Mina olen teadvuse küsimuses optimist: ma arvan, et ühel päeval on teadvuse teooria meil viimaks käes, ja siin raamatus ma püüangi seda leida. Kuid teadvuse kallale ei saa asuda tavaliste meetoditega: selleks, et üldse kuhugi jõuda, tuleb kõigepealt vaadata näkku kõigele, mis probleemi nii raskeks teeb. Alles siis saame asuda teooria enda kallale, ilma silmaklappideta ja hea ettekujutusega eesseisvast ülesandest.

.

Siin raamatus ei anna ma teadvuse probleemile lõplikku lahendust, vaid püüan ainult ohjad pihku haarata. Ma püüan probleemide olemuses selgusele jõuda, näitan, et neuro- ja kognitiivteaduse standardsed meetodid nende lahendamiseks ei sobi, ning püüan seejärel edasi liikuda.

Oma teadvusekäsitlust välja töötades olen ma püüdnud järgida mõningaid kitsendusi. Esimene ja kõige olulisem neist on, et ma püüan võtta teadvust tõsiselt. Kõige kergem on luua teadvuse “teooriat”, eitades teadvuse olemasolu või defineerida seletamist vajav fenomen ümber millekski, mida ta ei ole. Tavaliselt saadakse tulemuseks elegantne teooria, probleem aga ei kao. Siin raamatus lähtun ma kõikjal eeldusest, et teadvus on olemas ning et ei saa leppida probleemi ümbersõnastamisega, mis taandab kõik mingite kognitiivsete või käitumuslike funktsioonide seletamisele. Seda ma mõtlengi teadvuse tõsiselt võtmise all.

Mõnikord öeldakse, et teadvus olevat illusioon. Mina ei saa aru, mida selline väide õieti tähendadagi võiks. Mulle tundub, et maailmas ei ole midagi, mille olemasolus me oleksime kindlamad kui omaenda teadvuselamuses. Mõnikord olen ma kõigest väest püüdnud end veenda, et tegelikult ei ole selles midagi, et teadvuselamus on tühi, pelk illusioon. Selles arusaamas on midagi ahvatlevat, ja filosoofid ongi püüdnud seda kõigil aegadel ära kasutada, kuid lõpuks on tulemus osutunud alati täiesti mitterahuldavaks. Ma aistin midagi oranži, lahustun täielikult oranži värvuse aistingus, ja seejuures toimub midagi — siin on midagi, mis nõuab seletust ka veel siis, kui kõik eristus- ja tegevusprotsessid on ära seletatud — nimelt *elamus* [​*experience*​].

Tõsi, ma ei saa tõestada, et siin on veel üks probleem — just sellepärast, et ma ei saa tõestada, et teadvus on olemas. Me tunneme teadvust ​**[lk.xiii]**​ otsesemalt kui midagi muud, seega oleks siin kohatu midagi “tõestama” hakata. Parim, mida ma teha saan, on esitada oma argumente, kus vähegi võimalik, ja püüda kummutada vastaspoole omi. Ei saa salata, et mingist hetkest alates tuleb seejuures hakata apelleerima intuitsioonile; kuid kõik argumendid toetuvad ju nagunii kusagil intuitsioonile, ja ma olen püüdnud oma argumentides oma intuitiivsetele aimustele võimalikult selgesti osutada.

Seda olukorda võiks nimetada teadvuse uurimise Suureks Vastasseisuks. Kui te arvate, et “lihtsatele” probleemidele vastamisega on ära seletatud kõik, mis seletamist vajab, siis tuleb teil välja ühtmoodi teooria; kui te arvate, et lisaks kõigele sellele on olemas veel ka “raske” probleem, siis teistsugune. Mingist punktist edasi on raske selle vastasseisu üle argumenteerida ning poleemika mandub tihti rusikaga lauale põrutamiseks. Minu arvates on siin ilmselgelt veel midagi, mis nõuab seletust, teised aga ütlevad, et midagi muud enam ei ole, ning jäävad sellega rahule. (Mitteametlike küsitluste andmetel on esimese seisukoha pooldajaid kaks–kolm korda rohkem, kusjuures eri valdkondade teadlastel ja üliõpilastel on see suhtarv üsna ühesugune.) Võib-olla tuleb meil selle põhimõttelise vastasseisuga lihtsalt leppida.

Raamat võib pakkuda intellektuaalset huvi ka lugejale, kes arvab, et siin mingit erilist probleemi ei ole, aga päriselt mõeldud on ta lugejatele, kes tajuvad probleemi kogu oma olemusega. Praeguseks on meil küllalt hea ettekujutus sellest, mismoodi peaks teooria välja nägema eeldusel, et probleemi ei ole. Mina püüan käesolevas töös uurida, kuhu me välja jõuame, kui eeldada, et probleem on olemas. Õigupoolest seisnebki raamatu põhiline tees selles, et igaüks, kes võtab teadvust tõsiselt, peaks viimaks jõudma positsioonile, mida ma allpool kirjeldama hakkan.

Teine kitsendus, mida ma olen järginud, seisneb selles, et ma võtan tõsiselt ka teadust. Ma ei ole püüdnud seada kahtluse alla olemasolevaid teaduslikke teooriaid valdkondades, kus neil on autoriteeti. Samas ei ole ma kartnud minna nendega vastuollu seal, kus teadlaste arvamused on sama alusetud nagu kõigi teiste omad. Näiteks ei ole ma püüdnud vaielda vastu seisukohale, et füüsiline maailm on põhjuslikult kinnine või et käitumist on võimalik seletada füüsika mõistetes; aga kui füüsik või kognitiivteadlane väidab, et ka teadvust saab seletada füüsika mõistetega, siis on see vaid lootus, millele praegused teooriad tuge ei paku, ja küsimus jääb lahtiseks. Ühesõnaga, ma olen püüdnud arendada oma ideid nii, et need oleksid kooskõlas tänapäeva teadusega, kuid samas pole ma püüdnud piirduda ideedega, mis on tänapäeva teadlaste seas populaarsed.

Kolmas kitsendus seisneb selles, et ma pean teadvust looduslikuks fenomeniks, mille üle valitsevad loodusseadused. Seega peaks mingi korrektne teadvuseteooria siiski olemas olema, ükskõik, kas me suudame seda leida või mitte. Sellele, et teadvus on looduslik fenomen, on raske vastu vaielda: teadvus on looduse väga laialt levinud osa, mida võib kohata kõikjal meie oma inimliigis ja väga tõenäoliselt ka paljudel teistel liikidel. Ja meil on kõigiti alust uskuda, et looduslike fenomenide üle kehtivad kõik fundamentaalsed loodusseadused; oleks väga kummaline, kui see poleks nii teadvuse puhul. See ei tähenda, et teadvust puudutavad loodusseadused oleksid samasugused nagu teiste valdkondade seadused, ega isegi, et need oleksid tingimata just füüsikaseadused. Siin võib seadustel olla sootuks teistsugune kuju.

​**[lk.xiv]**​ Teadvuse probleem paikneb teaduse ja filosoofia ebamugaval piirialal. Mina ütleksin, et see on täiesti legitiimne teaduslik probleem: teadvus on samasugune looduslik fenomen, nagu seda on liikumine, elu ja kognitsioon, ning nõuab seletust samamoodi nagu needki. Kuid teadvust ei saa uurida tavapäraste teaduslike meetoditega. Igapäevasel teaduslikul metodoloogial ei ole kerge teadvusest kinni haarata, ja seda mitte viimases järjekorras raskuste tõttu, mida valmistab juba selle fenomeni vaatlemine. Niipea kui esimese isiku perspektiivist väljuda, läheb andmete hankimine raskeks. See ei tähenda, nagu poleks väliste andmetega midagi peale hakata, kuid me ei saa asuda põhjendama oma andmete relevantsust senikaua, kuni me pole jõudnud koherentse filosoofilise mõistmiseni. Seega on teadvuse probleem küll teaduslik probleem, kuid võib-olla niisugune, mida kallale tuleb esialgu asuda filosoofia meetoditega.

Järeldusi, milleni ma raamatus jõuan, peavad mõned võib-olla ebateaduslikeks. Nimelt väidan ma, et teadvusele ei leidu reduktiivset seletust, ning esitan koguni argumente mingis vormis dualismi kasuks. Kuid niisugune asi kuulub teadusliku uurimise protsessi juurde. Selgub, et üht liiki seletused ei tööta, ja seetõttu tuleb appi võtta ka muusuguseid seletusvõtteid. Kõik, mida ma siin ütlen, on tänapäeva teaduse tulemustega ühitatav; ma ei püüa olemasolevat pilti looduslikust maailmast pea peale pöörata, vaid seda laiendada. See laiendus annab võimaluse luua teadvuse kohta naturalistlik teooria, mis oleks vastasel korral võimatu. Minule tundub, et ebateaduslik oleks hoopis teadvuse probleemide ignoreerimine; teaduse vaim nõuab, et me vaataksime neile otse näkku. Nendele, kes usuvad, et teadus eeldab tingimata materialismi, ütlen ma: elame, näeme.

Tuleb tähendada, et raamatu järeldused on just nimelt järeldused selle sõna kõige rangemas mõttes. Minu temperament kallutab mind tugevasti materialistlike, reduktiivsete seletuste poole ning mul puuduvad selged vaimulikud ja usulised kalduvused. Palju aastaid lootsin ma jõuda materialistliku teooriani, ja kui ma sellest lootusest viimaks loobuma pidin, sündis see üsna vastumeelselt. Lõpuks sai mulle enda arvates selgeks, et mitte keegi, kes tahab teadvust tõsiselt võtta, ei saa neid järeldusi vältida. Materialism annab meile kauni ja veenva maailmapildi, ent teadvuse seletamiseks tuleb kasutada ka muid ressursse.

Nüüdseks olen ma nende järeldustega peaaegu rahul. Neil ei tundu olevat mingeid ohtlikke järelmeid ja nad annavad võimaluse mõtelda ja teoretiseerida teadvuse üle viisil, mis on teistest võimalustest peaaegu igas suhtes üle. Teadusliku maailmapildi laienemisel on seejuures olnud positiivne mõju, vähemalt minule: universumist on saanud minu jaoks huvitavam koht.

.

Raamatul on neli osa. [Esimeses osas](#ORG.I) annan ma probleemidest ülevaate ning kirjeldan mudelit, mis võimaldab nende kallal töötada. [1. peatükk](#ORG.1) on sissejuhatus, kus ma selgitan, mis on teadvus, panen paika piirid, mis lahutavad seda mõistet mõnedest naabermõistetest, näitan, mille poolest teadvus tõsist huvi pakub, ja annan esialgse ülevaate teadvuse peentest seostest vaimu teiste aspektidega. ​**[lk.xv]**​ [2. peatükk](#ORG.2) arendab välja metafüüsilise seletusmudeli, mille piires hargneb suur osa edasistest mõttearendustest. Mida tähendab ühe või teise fenomeni reduktiivne seletamine; mida tähendab, et üks või teine fenomen on füüsiline? 2. peatükk arutleb nende küsimuste üle, rakendades peamiselt *superventsuse* mõistet. Nagu ma seal näidata püüan, on kõigiti põhjust arvata, et peaaegu kõigile maailma nähtustele võib leida reduktiivse seletuse, kuid võib selguda, et teadvus on selles suhtes erand.

Nende eeltööde järel tuleb [teine osa](#ORG.II), mis koondab tähelepanu teadvuse mitteredutseeritavusele. [3. peatükk](#ORG.3) püüab tõestada, et tavapärased reduktiivsed seletamismeetodid ei sobi teadvuse seletamiseks. Ühtlasi kritiseerin mitmesuguseid reduktiivseid kontseptsioone, millega on välja tulnud neuroteaduse, kognitiivteaduse ja muude valdkondade uurijad. Peamine ei ole siin negatiivne tulemus, vaid järeldus, et teadvuseteooria, millega võiks rahule jääda, peaks olema mingit uut liiki mittereduktiivne teooria. [4. peatükk](#ORG.4) astub järgmise sammu, tõendades, et materialism on väär ja et tõene saab olla vaid mingitliiki dualism, ning näitab, kuidas mittereduktiivne teadvuseteooria peaks üldjoontes välja nägema. [5. peatükk](#ORG.5) on peamiselt apologeetiline: ma käsitlen seal mõningaid näilisi raskusi, millele minu vaateviis põrkub ja mis tulenevad muu hulgas teadvuse enda ja teadvust puudutavate otsustuste omavahelistest suhetest, ning püüan näidata, et need raskused ei ole ületamatud.

[Kolmandas osas](#ORG.III) hakkan liikuma positiivse teadvuseteooria poole. Igas peatükis arendan ma välja selle teooria ühe komponendi. [6. peatükis](#ORG.6) on vaatluse all teadvuse ja kognitiivsete protsesside vaheline koherentsus ning terve hulk süstemaatilisi seoseid nende kahe vahel. Nende seoste abil püüan ma analüüsida ja põhjendada neuroteaduse ja kognitiivteaduse keskset rolli inimese teadvuse seletamisel. [7. peatükis](#ORG.7) tuleb arutusele teadvuse seos süsteemi funktsionaalse organisatsiooniga; ma püüan mõtteliste katsete varal põhjendada väidet, et teadvus on *organisatsiooniline invariant* — s.t. kõigil sobiva funktsionaalse organisatsiooniga süsteemidel on ühesugused teadvuselamused, ükskõik, olenemata süsteemi koosseisust. [8. peatükk](#ORG.8) uurib, mismoodi võiks välja näha fundamentaalne teadvuseteooria, ning pakub välja oletuse, et selles oleks oluline koht teadvuse tihedal seosel informatsiooniga. See on palju spekulatiivsem peatükk kui kõik teised, ent kui me soovime edasi liikuda, siis me praeguses olukorras vist ilma mõningase spekuleerimiseta läbi ei saa.

Kaks viimast peatükki on magustoit. Neis rakendan ma eelmiste osade tulemusi tehisintellekti ja kvantmehhaanika kesksetele küsimustele. [9. peatükis](#ORG.9) püüan ma esitada argumente *tugeva tehisintellekti teesi* kasuks, mis ütleb, et sobival viisil implementeeritud arvutiprogrammis sünnib teadvusega vaim. [10. peatükis](#ORG.10) tuleb käsitlemisele segane küsimus, kuidas tuleks tõlgendada kvantmehhaanikat; püüan seal kasutada eelmistes peatükkides arendatud ideid teadvuse kohta, et toetada kvantmehhaanika “kollapsivaba” interpretatsiooni.

Pole võimatu, et kõige rohkem vastukaja äratab negatiivne materjal. Kuid minu tegelik eesmärk on positiivne: ma tahan näha töötavat teadvuseteooriat. Kui ma alles hakkasin filosoofiaga tutvuma, pani mind imestama, et suurem osa poleemikast ​**[lk.xvi]**​ teadvuse üle koondus küsimustele, kas probleem on olemas või mitte ning kas teadvus on füüsiline või mitte, kuid teooriate loomise kohustus oli jäetud otsekui ainult ühe osapoole kanda. Mulle näis, et “teooriaid” on esitanud ainult need autorid, kes (minu arusaamist mööda) ei võtnud teadvust tõsiselt. Nüüdseks oskan juba minagi nautida ontoloogilise debati finesse nagu iga teinegi, kuid minu peamiseks eesmärgiks on endiselt detailse teooria loomine. Kui mõnest selle raamatu ideest peaks olema teistele abi veel parema teooria loomisel, siis on mu püüdlused end ära tasunud.

Ma kavandasin seda raamatut kui tõsist filosoofia-alast tööd, kuid püüdsin kirjutada nii, et ta oleks jõukohane ka mittefilosoofidele. Minu mõttelises auditooriumis on kogu aeg istunud ka see üliõpilane, kes ma olin ise kümme aastat tagasi, ja ma loodan, et raamat, mis mul nüüd valmis sai, oleks talle meeldinud. Mõned raamatu jaod käsitlevad filosoofia tehnilisi peensusi. Need on tähistatud tärniga (\*) ja lugeja võib nad julgesti vahele jätta. Kõige tehnilisem materjal on [2.](#ORG.2) ja [4.](#ORG.4) peatükis. [Esimese peatüki 4. jagu](#ORG.1.4) ning teise peatüki [2.](#ORG.2.2) ja [3.](#ORG.2.3) jagu, samuti [5. peatüki viimane jagu](#ORG.5.8) käsitlevad filosoofilise semantika keerukusi. Teistest tärniga tähistatud lõikudest võiks lugeja vähemalt pilgu üle lasta, et saada mingi ettekujutus, millest jutt käib. Suure osa eriti tehnilisest materjalist ja kommentaaridest filosoofia-alase kirjanduse kohta paigutasin märkustesse raamatu lõpus. Ühel tehnilisel mõistel on aga ülioluline koht kogu raamatus — nimelt *superventsuse* mõistel, mida tutvustab [2. peatüki algus](#ORG.2.1). See hirmuäratava nimetusega mõiste väljendab väga loomulikku ideed, mis, kui see endale korralikult selgeks teha, asetab kõik kesksed probleemid omadele kohtadele. Suure osa [2. peatüki edasisest materjalist](#ORG.2.2) võib esimesel lugemisel vahele jätta, kuid hiljem võib lugejal tulla soov selle juurde tagasi pöörduda, et jõuda selgusele lugemise käigus tekkinud küsimustes.

Lugeja, kes tahab saada raamatu sisust võimalikult kiire ülevaate ega soovi süveneda erialastesse peensustesse, lugegu alustuseks läbi [1. peatükk](#ORG.1), sirvigu vajalike taustteadmiste saamiseks läbi [2. peatüki esimesed jaod](#ORG.2.1), seejärel lugegu (vajadusel [1. jagu](#ORG.3.1) üle lehitsedes) läbi kogu [3. peatükk](#ORG.3), mis käsitleb keskseid argumente reduktiivsete seletuste vastu, ning 4. peatüki [esimene](#ORG.4.1) ja [viimane](#ORG.4.6) jagu, kus tulevad jutuks kesksed dualismi puudutavad kaalutlused. Tasub lugeda läbi ka [6. peatüki](#ORG.6) algusosa, et saada umbkaudne ettekujutus positiivse vaateviisi üldisest kontuuridest. Positiivsest materjalist endast võib ehk kõige eneseküllasemana mainida [7. peatükki](#ORG.7), mis tänu oma hõlpsasti arusaadavatele mõttelistele eksperimentidele ränist ajudega on ühtlasi kõige lõbusam; lugejale aga, kes naudib pööraseid ja vigurlikke spekulatsioone, võiks pakkuda naudingut [8. peatükk](#ORG.8). [9.](#ORG.9) ja [10.](#ORG.10) peatükk peaksid sobima igaühele, keda vastavad probleemid huvitavad.

.

Paar filosoofilist märkust. Teadvust käsitlev filosoofiline kirjandus on üsna süsteemitu: eri voolud kõnelevad sarnastest küsimustest pealtnäha täiesti sõltumatult, üksteisega dialoogi astumata. Et selles segadikus mõnevõrra korda luua, olen pakkunud välja ühtsustava mudeli, mis toob selgust mitmesugustesse metafüüsilistesse ja seletuslikesse probleemidesse. Suurt osa kirjanduses peetavatest diskussioonidest on võimalik kadudeta selle mudeli keelde ​**[lk.xvii]**​ tõlkida, ja ma loodan, et see struktuur toob nähtavale sügavad seosed mitmesuguste küsimuste vahel.

Ühes suhtes on raamat vist tavatu: ma ei arutle siin kuigi palju filosoofiast tuttava identsuse (näiteks vaimsete ja füüsiliste seisundite identsuse) mõiste üle, eelistades sellele superventsuse mõistet. Ma olen leidnud, et arutlused, kus väited on üles ehitatud identsuse mõistele, mitte ei heida peamistele küsimustele valgust, vaid ajavad neid pigem veelgi segasemaks ning teevad kõige olulisematest raskustest möödavaatamise väga hõlpsaks. Seevastu superventsus annab, nagu mulle tundub, ideaalse mudeli põhiliste küsimuste käsitlemiseks. Et aga oma filosofeerimises mitte laiali valguda, tuleb pöörata terast tähelepanu sellele, kust tuleb superventsusseose jõud: kas selle tagab loogiline paratamatus, looduslik paratamatus või miski muu? Enam-vähem kõik autorid on nõus, et teadvus on füüsika suhtes mingis mõttes superventne; küsimus on vaid, kui tihe see seos on. Autorid, kes neid modaalsusküsimusi ignoreerivad, lähevad tavaliselt mööda ka kõige raskematest küsimustest teadvuse kohta. Lugejad, kes on modaalsetele mõistetele skeptilised, jäävad küllap skeptiliseks ka kõigile minu mõttekäikudele, kuid minu arvates ei ole nende küsimuste püstitamiseks ühtki teist rahuldavat viisi.

.

Üks selle raamatu kallal töötamise rõõme on olnud näha, kuidas teadvuse probleem puutub kokku sügavate küsimustega paljudes teistes teaduse ja filosoofia valdkondades. Kuid samas teevad probleemi ulatus ja sügavus ta ka ränkraskeks. Ma saan liigagi hästi aru, et peaaegu kõige üle, millest ma raamatus kõnelen, oleks saanud arutleda veelgi põhjalikumalt, ning et paljudes kohtades olen ma vaid pealispinda riivanud. Kuid ma loodan, et mul on vähemasti õnnestunud näidata, et teadvuse probleemiga on võimalik edasi liikuda, eitamata teadvuse olemasolu ja taandamata teda millekski, mida ta ei ole. Probleem on paeluv ja tulevik on põnev.

### [epigraaf]

​**[lk.xviii]**​ *Ei.* Xia peatus ning pöördus aeglaselt Joey poole. Tema mündikarva silmad läksid pärani lahti. *Sul ei ole siin ohutu olla.* Õudusest kahvatu näoga vaatas Xia maja poole. *Mine kohe koju. Enne kui hilja. Ja otsi mulle vastumürk.*​

​*Mis vastumürk?*​

Xia kadus kadakate taha, kuid tema viimased sõnad plahvatasid Joey peas nagu paukpadrun: *Vastumürk sombimürgile.*​

Dian Curtis Regan, Minu klassiõde on sombi​[[4]](#footnote-5)

# ​**[lk.1]**​ I OSA: Põhialused

## ​**[lk.3]**​ 1. peatükk. Kaks ettekujutust vaimust

### 1. Mis on teadvus?

Teadvuselamus on kõige tavalisem ja samas kõige salapärasem asi maailmas. Meil ei ole mitte millestki otsesemat teadmist kui teadvusest, kuid pole kaugeltki selge, kuidas ühitada seda teadmist kõigi meie teiste teadmistega. Miks teadvus eksisteerib? Mida ta teeb? Kuidas on võimalik teadvuse sünd hallis ainekämbus? Me tunneme teadvust palju lähemalt kui kõike muud, kuid mõistame kõike muud palju paremini kui teadvust.

Teadvus võib olla hämmastavalt intensiivne. Ta on kõigist fenomenidest kõige elusam, ja pole midagi, mis oleks meie jaoks reaalsem. Kuid teadvus võib meid oma õhkõrnusega ka meeleheitele ajada: on ju hästi teada, kui raske on kõneainet täpselt määratleda, kui jutt on teadvuselamustest. “Rahvusvaheline psühholoogiasõnaraamat” [​*The In­ter­na­tion­al Dictionary of Psychology*​] ei tee katsetki teadvust otseselt iseloomustada:

*Teadvus*​: tajude, mõtete ja tunnete olemasolu; teadlikkus. Terminit pole võimalik defineerida teisiti kui terminite kaudu, millest pole võimalik aru saada, senikaua kui pole teada, mida tähendab teadvus. Teadvust aetakse sageli segi eneseteadvusega. Ta eeldab ainult teadlikkust välisest maailmast. Teadvus on paeluv, kuid heitlik fenomen: pole võimalik täpselt öelda, mis ta on, mida ta teeb või miks ta evolutsiooni käigus tekkis. Teadvuse kohta pole kirjutatud midagi lugemisväärset. (Sutherland 1989)

Peaaegu igaüks, kes on teadvuse üle põhjalikult järele mõelnud, saab neist meeleoludest mingil määral aru. Teadvus on nii tabamatu asi, et isegi säärasele arglikule defineerimiskatsele annaks vastu vaielda: subliminaalsete tajude ja mitteteadvuslike mõtete olemasolu näitab, et tajud ja mõtted ei pea tingimata olema teadvuslikud. Teadvuse otsustava tähtsusega elemendiks, ​**[lk.4]**​ vähemalt selle sõna kõige huvitavamas mõttes, on *elamus* [​*experience*​]. Kuid see pole definitsioon, vaid kõige enam ehk selgitus.

Katsed defineerida teadvuselamust elementaarsemate mõistete kaudu ei anna mingeid tulemusi. Sama hästi võiks püüda defineerida fundamentaalsemate mõistete kaudu mateeriat või ruumi. Me ei suuda teha midagi paremat, kui esitada näiteid ja kirjeldusi teadvuse enda tasemel. Kuid kirjeldusi ei saa pidada tõelisteks definitsioonideks, kuna neis on kaudselt alati loogiline ring. Siiski võivad kirjeldused aidata täpsustada, millest käib jutt. Ma eeldan, et igal lugejal on teadvuselamused olemas. Loodetavasti saab lugeja meie kirjelduste põhjal aru, et just nendest me räägimegi.

Meie uurimisobjekti on ehk kõige sobivam nimetada *elamuse subjektiivseks kvaliteediks*​. Kui me tajume, mõtleme ja tegutseme, käivitub meis terve pööris põhjuslikke infotöötlusprotsesse. Need ei kulge tavaliselt pimedas, vaid neil on ka seesmine aspekt: on olemas tunne olla kognitiivne agent. See seesmine aspekt ongi teadvuselamus. Teadvuselamused varieeruvad väga laias ulatuses: kirgastest värvustajudest vaevu aimuvate lõhnaaistinguteni; lõikavast valust tabamatu mõtteni “keele peal”; igapäevastest helidest ja lõhnadest hingepõhjani vapustavate muusikaelamusteni; banaalsest kihelusest ränkraske, sügava eksistentsiaalse ahastuseni; piparmündi spetsiifilisest maitsest omaenda iseduse kõikehõlmava läbielamiseni. Kõigil neil elamustel on omaenda selgelt tuntav kvaliteet. Kõik nad mängivad vaimu siseelus silmapaistvat rolli.

Me võime öelda, et sellel või teisel olendil on teadvus, kui — et kasutada väljendit, mille tegi kuulsaks Thomas Nagel[[5]](#footnote-6) — on olemas *tunne olla* [​*something it is like to be*​] see olend. Analoogiliselt võib öelda, et vaimuseisund on teadvuslik, kui on olemas tunne olla selles vaimuseisundis. Seda võib öelda ka teisiti: vaimne seisund on teadvuslik, kui temaga seostub mingi kvalitatiivne tunne, elamuslik kvaliteet. Neid kvalitatiivseid tundeid nimetatakse ka *fenomenaalseteks kvaliteetideks* või lühidalt *kvaalideks* [​*qualia*​].[[6]](#footnote-7) Teadvuse seletamine tähendabki nende fenomenaalsete kvaliteetide seletamist. Just selles seisneb keha–vaimu probleemi tõeliselt raske osa.

Miks peaks teadvuselamus üldse eksisteerima? Tal on keskne tähtsus subjektiivse vaatepunkti jaoks, kuid objektiivselt vaadates on teadvus midagi täiesti ootamatut. Objektiivsest vaatepunktist võib pikalt-laialt jutustada, kuidas väljade, lainete ja osakeste ülipeente vastastoimete tulemusena aegruumilises muutkonnas​[[7]](#footnote-8) kujunevad ülikeerulised süsteemid, nagu seda on aju. Põhimõtteliselt ei ole nende süsteemide keerukas infotöötluses või arukates reaktsioonides stiimulitele mingit sügavat filosoofilist salapära, seda pole isegi säärastes keerulistes võimetes nagu õppimine, mälu ja keel. See kõik avaldab muljet, kuid midagi metafüüsiliselt rabavat siin ei ole. Seevastu teadvuselamuse olemasolu on niiviisi vaadates midagi hoopis uut. Ainult nonde teiste andmete põhjal ei oleks seda saanud ennustada.

​**[lk.5]**​ Niisiis, teadvuse olemasolu on üllatav. Kui me teaksime ainult füüsika fakte, või isegi ka fakte keerukate süsteemide dünaamikast ja neis toimuvast infotöötlusest, ei oleks meil mingit veenvat alust postuleerida teadvuselamuse olemasolu. Kui meil puuduksid otsesed tõendid esimese isiku positsioonilt, tunduks see hüpotees alusetuna, võib-olla koguni müstilisena. Ja ometi me teame otseselt, et teadvuselamus on olemas. Küsimus on, kuidas ühitada seda meie kõigi muude teadmistega.

Teadvuselamus on üks osa looduslikust maailmast ja nõuab seletust nagu iga teinegi looduslik fenomen. Peamisi seletust nõudvaid asju on siin vähemalt kaks. Esimene ja kõige kesksem on teadvuse olemasolu ise. Miks on teadvuselamus üldse olemas? Kui ta kasvab välja füüsilistest süsteemidest, nagu see nähtavasti on, siis kuidas see toimub? Siit hargneb edasi ka spetsiifilisemaid küsimusi. Kas teadvus on ise füüsikaline nähtus või ainult miski, mis füüsikaliste süsteemidega kaasneb? Kui laialdaselt teadvust looduses esineb? Kas teadvuselamusi on ka näiteks hiirtel?

Teiseks nõuab seletust teadvuselamuste spetsiifiline iseloom. Kui eeldada, et teadvuselamus on olemas, siis miks on igal üksikul läbielamisel just niisugune konkreetne eripära? Miks on mul silmi avades ja oma kabinetis ringi vaadates just niisugune paljutahuline elamus? Või sügavamal tasemel: miks ma näen punast värvust just nii-, aga mitte teistsugusena? Tundub olevat mõeldav, et vaadates midagi punast, näiteks roosi, oleks minu värvuselamus niisugune, nagu ta on tegelikult siis, kui ma vaatan midagi sinist. Miks on elamus just niisugune? Miks on meie elamus punasest aistingust[[8]](#footnote-9) just niisugune, nagu ta tõepoolest on, aga mitte niisugune, nagu mõni hoopis teine aisting, näiteks see, mida kutsub esile trompetiheli?

Kui võtta klaveril neljanda oktavi C[[9]](#footnote-10), käivitub keerukas sündmuste ahel. Õhus tekivad helivõnked ja minu kõrva saabub helilaine. Sisekõrv lahutab võnkumise sagedusteks ja saadab kuulmiskorteksisse[[10]](#footnote-11) mingi signaali. Seal toimub edasine töötlus: signaali mitmesuguste aspektide eristamine ja liigitamine ning lõpuks reageerimine. Põhimõtteliselt ei ole seda kõike kuigi raske mõista. Kuid miks peaks sellega kaasnema elamus? Ja miks peaks sellega kaasnema just niisugune elamus, just niisuguse iseloomuliku mahlaka tooni ja tämbriga? Need on kaks keskset küsimust, millele me teadvuseteoorialt vastust ootame.

Ideaalis ootaksime teadvuseteoorialt vähemalt seda, et ta näitaks meile kätte tingimused, mille korral füüsikaliste protsesside tagajärjel sünnib teadvus, ja et ta nende protsesside kohta, mille puhul see toimub, ütleks täpselt, missugused on nendega kaasnevad elamused. Ja veel peaks teooria täpselt seletama, kuidas teadvus tekib — me tahame, et see ei jääks maagiliseks nähtuseks, vaid oleks arusaadav. Kokkuvõttes ootame, et teooria lubaks meil näha teadvust loodusliku maailma lahutamatu osana. Praegu on meil võib-olla raske ette kujutada, kuidas säärane teooria peaks välja nägema, ent senikaua kui meil teooriat ei ole, ei saa öelda, et me mõistame teadvust lõpuni.

​**[lk.6]**​ Enne kui jätkata, tuleb teha üks terminoloogilise märkus. “Teadvus” on mitmetähenduslik termin, mis võib tähistada üpris mitut erinevat fenomeni. Mõnikord tähistatakse selle sõnaga mõnd kognitiivset võimet, näiteks introspektsioonivõimet või suutlikkust raporteerida oma vaimuseisunditest. Mõnikord on “teadvus” lihtsalt “ärkveloleku” sünonüüm. Mõnikord seostub ta inimese võimega oma tähelepanu koondada või oma käitumist tahteliselt juhtida. Mõnikord tähendab “teadlikkus millestki” [​*to be conscious of something*​] lihtsalt selle “millegi” teadmist. Kõik need on tavapärased uusused, kuid kõik nad viitavad fenomenidele, mis ei kuulu siinse raamatu ainesse, ja kõigi nende seletamisega on ka oluliselt vähem raskusi. Nendest alternatiivsetest teadvusemõistetest tuleb pikemalt juttu allpool, ent esialgu pean ma selle sõnaga silmas ainult läbielamiste subjektiivset kvaliteeti — tunnet olla kognitiivne agent.

On veel teisigi termineid ja väljendeid, mis tähistavad umbes sedasama fenomenide klassi, mida tähistab sõna “teadvus” oma keskses tähenduses. Niisugused on näiteks “elamus”, “läbielamine”, “kvaal”, “fenomenoloogia”, “fenomenaalne”, “subjektiivne elamus” ja “tunne olla”. Kui mitte arvestada grammatilist külge, siis seisnevad nende väljendite erinevused peamiselt peentes tähenduslikes nüanssides. Väljend “mul on teadvus” on selles mõttes ligikaudselt sünonüümne väljenditega “mul on kvaalid”, “mul on subjektiivseid läbielamisi” jne. Erinevused selle fenomenideklassi sees on tühised. Nagu “teadvus”, nii pole ka paljud neist teistest väljenditest täiesti ühemõttelised, kuid mina ei tarvita neid siin raamatus kusagil mingites muudes tähendustes. Kõneldes fenomenist, mis on selle raamatu keskseks teemaks, tarvitan ma kõiki neid väljendeid, kuid kõige täpsemad neist on “teadvus” ja “elamus”, ja need terminid hakkavadki siin edaspidi korduma.

#### Väike teadvuselamuste kataloog

Teadvuselamuste jälgimine on põnev tegevus. Elamusi on tohutult erisuguseid ja igal neist on omaenda iseloom. Järgnevalt esitame eelteoreetilise ja impressionistliku teadvuselamuse vormidest väikese, kaugeltki mitte täieliku loendi, mida ei tuleks võtta liiga tõsiselt ega filosoofiliselt, kuid mis loodetavasti aitab koondada tähelepanu raamatu peateemale.

##### Visuaalsed elamused.

Visuaalse elamuse paljude vormide seas on teadvuselamuse kanooniliste näidistena oma puhtuse ja pealtnäha täieliku sõnulseletamatuse tõttu eriti oluline koht värvusaistingutel. Mõned värvuselamused on väga ilmekad ning sobivad seepärast eriti hästi meie tähelepanu juhtimiseks teadvuse müsteeriumile. Näiteks praegusel hetkel näen ma raamaturiiulil ere-sügavlillade kaantega raamatut, seinal on foto sõnajalgadest, mille roheline toon on peaaegu sürrealistlik, ja aknast paistab ​**[lk.7]**​ eredate vilkuvate punaste, roheliste, oranžide ja siniste tuledega jõulukuusk. Ent hämmastada võib iga värvus, niipea kui me talle tähelepanu pöörame ja tema olemuse üle juurdlema hakkame. Miks peaks mul olema seda värvust nähes just niisugune tunne? Miks peaks mul teda nähes üldse mingi tunne olema? Kuidas kirjeldada selle värvuselamuse olemust inimesele, kes pole kunagi midagi sarnast läbi elanud?

Visuaalsel elamusel on teisigi vorme — on olemas kuju-, suurus-, heledus- ja tumeduselamused. Üks eriti peen aspekt on sügavuselamus. Lapsepõlves nägi mul üks silm suurepäraselt, teine väga halvasti. Kuna üks silm nägi hästi, siis oli pilt maailmast selge ja terav ning kahtlemata kolmemõõtmeline. Ühel päeval sain ma aga prillid, ja muutus oli ilmselge. Pilt ei läinud palju teravamaks kui enne, kuid maailm oli korraga justkui “kolmemõõtmelisem”: sügavused olid kuidagi sügavamad ja kogu maailm külluslikum. Et aru saada, millest ma räägin, võib lugeja ühe silma korraks kinni katta. Varem oleksin öelnud, et minu nägemissügavuses ei ole midagi parandada — pilt maailmast tundus juba niigi nii kolmemõõtmeline, nagu vähegi olla sai. Muutus oli peen, peaaegu tabamatu, kuid samas täiesti rabav. Muidugi võib siin nüüd hakata intellektuaalselt heietama, kuidas binokulaarne nägemine võimaldab mõlemast silmast saabuva informatsiooni ühendamise teel saada informatsiooni kauguste kohta ning aitab seeläbi ka tegevust tõhusamalt juhtida; ent millegipärast ei anna kogu see põhjuslik jutt ikkagi edasi läbielamisega kaasnevat tunnet. Miks peaks säärane muutus infotöötlusprotsessides muutma nii põhjalikult ka elamust? See jäi mulle mõistatuseks kümneaastaselt ja hämmastab endiselt ka praegu.

##### Kuulmiselamused.

Helid on visuaalsetest kujutistest mõnes suhtes isegi kummalisemad. Kujutise struktuur on tavaliselt otseses vastavuses maailma struktuuriga, helid võivad aga olla pealtnäha üsna iseseisvad. Minu telefonile tuleb kõne, aparaadi mehhanismis hakkab midagi võnkuma, õhus tekib keerukas laine, see jõuab minu kõrva — ja mingi moel kuulen ma, otsekui nõiaväel, helinat. Helina kvaliteedis ei ole justkui midagi, mis vastaks ühelegi struktuurile maailmas, kuid ma tean päris kindlasti, et heli tuleb telefoni valjuhääldist ja et ta sünnib mingi kindla kujuga lainetusest. Kuid miks peaks laine selline kuju, või isegi need närviimpulsid, sünnitama just niisuguse kvaliteedi?

Muusikaelamused ongi vahest kõige külluslikumad helielamused, kuigi ka kõneelamused ei jää neist palju maha. Muusika võib meid üle ujutada ja täielikult endasse neelata, ümbritseda meid nii, nagu meid ümbritseb visuaalne väli. Teised kuulmiselamused seda üldiselt ei suuda. Muusikaelamust võib analüüsida, lahutades tajutud helid nootideks ja toonideks ning leida nende vahel keerukaid suhteid, kuid muusikaelamuses on justkui midagi enamat. Akord kutsub esile ​**[lk.8]**​ ühtse kvalitatiivse elamuse, mida ei anna juhuslikult võetud noodid. Vana klaver ja kauge oboe võivad üheskoos luua ootamatult painava elamuse. Selle üle juurdlema hakkates küsime endalt taas: miks peaks nii- ja niisuguste helide kuulmisega kaasnema just niisugune tunne?

##### Kompimiselamused.

Ülikülluslikku elamuskvaliteetide ruumi pakuvad ka tekstuurid: võrrelge näiteks sameti pehmet pinda külma metalli või higise käe või habemetüükas lõua tekstuuriga. Igal neist on oma ainulaadne kvaliteet. Taas hoopis teistsuguseid puute-elamusi pakuvad vesi, suhkruvatt või teise inimese huuled.

#####  Haistmiselamused.

Kujutlege kopituse lõhna vanas riidekapis, roiskuvate toidujäätmete lehka, värskelt niidetud rohu hõngu, ahjusooja leiva aroomi. Lõhnaaistingute mitmekesisuse, tabamatuse ja sõnulseletamatu iseloomu tõttu on haistmine kõigist meeltest mingis mõttes kõige salapärasem. Ackerman (1990) nimetab haistmist “tummaks, sõnatuks meeleks”. Kõigis aistingutes on midagi tabamatut, kuid teistel meeltel on siiski omadusi, mis kirjeldamist mõnevõrra hõlbustavad. Nägemis- ja kuulmiselamustel on keerukas kombinaatorne struktuur, mida on võimalik kirjeldada. Elamused, mida annavad kompimis- ja maitsmismeel, tekivad tavaliselt otsesel kokkupuutel mingi objektiga, ja ajapikku on viidete alusel neile objektidele kujunenud rikkalik sõnavara nende elamuste kirjeldamiseks. Seevastu lõhnadel ei ole kuigivõrd ilmset struktuuri, nad hõljuvad meie ümber sageli pealtnäha ilma ühegi objektita, olles lihtsalt ja elementaarselt kohal meie meelte muutkonnas​. (Loomad tuleksid nende kirjeldamisega võib-olla paremini toime (vt. joonis 1.1).) See elementaarsus tuleneb osalt võib-olla sellest, et protsessid, mille varal meie haistmisretseptorid eri liiki molekule eristavad, toimivad “luku ja võtme” põhimõttel. See, et mingit kindlat liiki molekulid põhjustavad just niisuguse aistingu, on pealtnäha meelevaldne, kuid seda nad teevad.

**Joonis 1.1.** Lõhnaelamuse tabamatus ja tabatavus. (Calvin ja Hobbes © Watterson. Distributed by Universal Press Syndicate. Reprinted with permission. All rights reserved)​[[11]](#footnote-12)

— Nuh-nuh … Keegi teeb lõket. Nii hea on haista külmal talvepäeval lõkkesuitsu lõhna.

— Kas pole kummaline, et lõhnad puudutavad meid nii sügavalt, aga ometi me ei oska neid kirjeldada?

— Ei tea midagi. Näiteks sellel lõkkel seal on kõrev, surkjas lõhn.

— Seda oligi arvata, et loomadel lõhnade jaoks sõnu jätkub.  — Tsipa ruhane ka, õhk on kuivavõitu.

##### ​**[lk.9]**​ Maitseelamused.

Uurimistöödest psühhofüüsika alal on teada, et maitseaistingutel on ainult neli sõltumatut dimensiooni: magus, hapu, mõru ja soolane. Kuid ühenduses meie haistmismeelega mahutab see neljamõõtmeline ruum väga suurel hulgal kõikvõimalikke elamusi, karriga aedoa​salati[[12]](#footnote-13), halvaa või piparmündikommide maitsest küpsete virsikute maitseni.

##### Kuuma- ja külmaelamused.

Lämmatavalt kuumal ja niiskel suvepäeval on meie elamused kvalitatiivselt hoopis teistsugused kui pakaselisel talvepäeval. Mõelge ka kuumaaistingule nahal tule läheduses ning kõrvetavale külmaaistingule väga külma jääd puudutades.

##### Valu.

Valu on teadvuselamuse kanooniline näide, mida filosoofid eriliselt armastavad — võib-olla sellepärast, et valuaistingud moodustavad täiesti erilise kvalitatiivsete elamuste klassi ning kuigi valu tajutakse tavaliselt mingis kindlas kehaosas, on neid raske seada otsesesse vastavusse ühegi struktuuriga maailmas või meie kehas. Seepärast võib tunduda, et valu on isegi subjektiivsem kui enamik meelte-elamusi. Valuelamusi on väga erinevaid, alates äkilistest lõikavatest valuhoogudest, ägedatest kõrvetustest ja teravatest pistetest ja lõpetades tuima, näriva valuga.

##### Muud kehalised aistingud.

Valu on vaid kõige silmatorkavam mitmesugustest kindlate kehapiirkondadega seostuvatest aistingutest. Sääraste aistingute hulka kuuluvad veel peavalu (mis moodustab võib-olla omaette valude klassi), näljanäpistus, sügelus, kõdi ja urineerimisvajadusega kaasnev elamus. Paljudel kehalistel aistingutel on oma täiesti ainulaadne kvaliteet, mis teeb neid erinevaks kõigist teistest elamustest: mõelge kasvõi orgasmile või elektrikondi äralöömisele. Elamusi seostub veel ka lihastajuga — võimega teada oma keha asendit ja ruumilist asukohta.

##### Kujutluspildid.

Sukeldudes veelgi sügavamale — elamustesse, mis ei seostu konkreetsete objektidega meie keskkonnas või kehas, vaid sünnivad nii-öelda seesmiselt —, jõuame kujutluspiltideni. Sageli on visuaalsed pildid, mida me oma fantaasias loome, fenomenoloogiliselt väga värvikad, kuigi kaugeltki mitte nii detailsed kui pildid, mis tekivad otsesel visuaalsel tajumisel. Silmi sulgedes või kissitades võime näha huvitavaid värvilisi mustreid, pärast eredavärvilise pildi vaatamist tugevaid järelkujutisi. Analogiliselt võib meie fantaasia luua ka auditoorseid, isegi taktiilseid, olfaktoorseid ja maitselisi “pilte”, kuigi need on vähem püsivad ja nendega kaasnev kvalitatiivne tunne on enamasti nõrgem.

##### Teadvuslikud mõtted.

Mõnede meie mõtete ja uskumustega ei kaasne mingit erilist kvalitatiivset tunnet, teistega aga küll — ​**[lk.10]**​ eelkõige selgesõnaliste, konkreetsete mõtetega, mida me endamisi mõtleme, ning mitmesuguste muude mõtetega, mis meie teadvusvoogu mõjutavad. Ühe või teise konkreetse mõttega kaasnevat kvalitatiivset tunnet on sageli raske tabada, kuid ta on kahtlemata olemas. On olemas tunne mõelda neid mõtteid.

Kui ma mõtlen näiteks lõvile, siis hoovab minu fenomenoloogilisse teadvusse otsekui midagi lõvilist: tunne mõelda lõvile erineb mingil tabamatul moel tundest mõelda Eiffeli tornile. Veelgi märgatavam ja tugevam fenomenaalne hõng kaasneb kognitiivsete hoiakutega, näiteks soovidega — need avaldavad meile otsekui mingit fenomenoloogilist “tõmmet” —, ja mälestustega, näiteks nostalgia- või kahetsuselamustega.

##### Emotsioonid.

Emotsioonidega kaasnevad sageli väga selged elamused. Särtsakas rõõm, sügav ja kurnav depressioon, hõõguv vihahoog, kahetsuse melanhoolia — kõik need võivad avaldada teadvuselamusele sügavat mõju, kuigi vähem spetsiifilist kui kindla ruumipiirkonnaga seonduvad elamused, näiteks aistingud. Senikaua kui emotsioon kestab, tungib ta kõigisse meie teadvuselamustesse ja annab neile värvi.

Teised, põgusamad tunded jäävad mõjuvuselt emotsioonide ja vaimu ilmsemate kognitiivsete aspektide vahepeale. Mõelge näiteks mõnusööstu hetkel, mil te anekdoodi mõttest aru saate. Veel võib näiteks tuua pinevustunde õudusfilmi vaadates või mõnd tähtsat sündmust oodates. Samasse klassi kuulub ka närvilisusega mõnikord kaasnev ärevustunne.

##### Minatunne.

Mõnikord me tunneme, et teadvuselamuses on lisaks kõigile neile spetsiifilistele elementidele veel midagi  — otsekui mingi taustmüra, mis kuulub kuidagi loomupäraselt teadvuse juurde ja kestab ka siis, kui kõik teised komponendid puuduvad. See fenomenoloogiline Mina on nii sügav ja tabamatu, et tundub mõnikord illusoorsena, kuna temas ei ole midagi peale spetsiifiliste elementide nagu need, millest me ülal mõningaid loetlesime. Siiski tundub selles fenomenoloogilises Minas midagi olevat, kuigi selle kohta on väga raske midagi täpsemat öelda.

.

Ülaltoodud kataloog hõlmas paljusid baasjuhtumeid, kuid teist sama palju jäi sellest ka välja. Näiteks ei maininud ma seal unenägusid, erutust ja väsimust, joovet, samuti paljusid eriskummalisi elamusi, mida saab esile kutsuda farmakoloogilisel teel. Palju väga värvikaid elamusi tekib kahe või ennama ülalkirjeldatud komponendi ühinemisel. Mainisin juba efekti, mille loovad üheskoos lõhn ja maitse, ent sama mainimisväärseteks näideteks on elamused, mis sünnivad muusika ja emotsioonide väga tihedatest ja peentest vastastoimetest. Jätsin kõrvale ka teadvuselamuse ühtsuse — kõigi nende elamuste liitumise ainsa subjekti ühtseks elamuseks. Nagu minatunne, nii näib ka see ühtsus ​**[lk.11]**​ mõnikord illusoorsena — teda on ilmselgelt raskem tabada kui ühtki spetsiifilist elamust —, kuid intuitsioon ütleb väga selgesti, et see ühtsus on olemas.

Teadvuselamuste kirevast mitmekesisusest allpool paraku enam nii põhjalikult juttu ei tule. Lihtne värvusaisting pakub teadvuselamuse filosoofiliste mõistatuste uurijale sama sügavaid probleeme kui elamus, mille me saame Bachi koraali kuulates. Need sügavad küsimused puudutavad kõiki neid vorme niivõrd ühtviisi, et elamuse konkreetne iseloom ei olegi nii oluline. Kuid siiski peaks see teadvuselamuste rikkale mitmekesisusele heidetud põgus pilk aitama meil teemale tähelepanu koondada ning annab meile ka mõningase varu näiteid, mida on hea abstraktsemaid mõttekäike jälgides meeles pidada.[[13]](#footnote-14)

### 2. Fenomenaalne ja psühholoogiline vaimukäsitlus

Vaim [​*mind*​] ei ammendu teadvuselamusega. Seda võib mõista, kui panna tähele, et kuigi tänapäeva kognitiivteadusel pole olnud peaaegu mitte midagi öelda teadvuse kohta, on tal olnud palju öelda vaimu kohta üldiselt. Teda huvitavad vaimu teised aspektid. Kognitiivteadus tegeleb peamiselt käitumise seletamisega, ja vaim huvitab teda, kui üldse, vaid käitumise seesmise alusena ning vaimuseisundid seisunditena, mille abil saab käitumist põhjuslikult seletada. Need seisundid võivad olla niihästi teadvuslikud kui mitteteadvuslikud. Kognitiivteaduse seisukohalt on kõik nii- või teistsugust käitumist põhjustavad seesmised seisundid ühtviisi vaimu seisundid, olgu nad teadvuslikud või mitte.

Selle kõige aluseks on kaks täiesti erinevat vaimu mõistet. Esimene neist on *fenomenaalne* vaimumõiste. See tähendab, et vaimu all mõeldakse teadvuselamust ja vaimuseisundite all teadvuslikult läbi elatavaid seisundeid. See on vaimu kõige mõistatuslikum aspekt, ja just sellele aspektile koondan ma siin oma tähelepanu. Kuid see aspekt ei ammenda vaimu. Teiseks on olemas ​*psühholoogiline* vaimumõiste: arusaam vaimust kui käitumise põhjuslikust või seletuslikust alusest. Vaimuseisundid on selle arusaama järgi seisundid, mis mängivad õiget põhjuslikku rolli käitumise kujundamises või vähemasti sobivat rolli käitumise seletamisel. Psühholoogilise käsitluse jaoks pole kuigi oluline, kas ühe või teise vaimuseisundiga kaasneb teadvuslik kvaliteet või mitte. Oluline on vaid seisundi roll kognitiivses majapidamises.

Fenomenaalse arusaama järgi iseloomustab vaimu see, mis tunne subjektil on, psühholoogilise kontseptsiooni järgi aga see, mida ta teeb. Nende kahe käsitlusviisi vahel pole mingit vastuolu. Kumbki neist ei ole vaimu ainuõige analüüs. Nad käsitlevad kaht erinevat fenomeni, mis on mõlemad täiesti reaalsed.

​**[lk.12]**​Allpool kõnelen ma mõnikord vaimu fenomenaalsest ja psühholoogilisest “aspektist”, mõnikord “fenomenaalsest vaimust” ja “psühholoogilisest vaimust”. Praeguses varases etapis ei taha ma ette otsustada küsimust, kas fenomenaalne vaim ja psühholoogiline vaim ei osutu ehk üheks ja samaks asjaks. Pole võimatu, et iga fenomenaalne seisund on ühtlasi psühholoogiline seisund, kuna ta mängib olulist rolli käitumise põhjuslikes mehhanismides ja käitumise seletamises, või et kõigil psühholoogilistel seisunditel on tihe seos fenomenaalse maailmaga. Esialgu on oluline vaid kontseptuaalne erinevus nende kahe käsitlusviisi vahel: fenomenaalne seisund on seisund, mis seostub subjekti jaoks mingi kindla tundega; psühholoogiline seisund on seisund, mis mängib mingit sobivat põhjuslikku rolli. Neid kaht kontseptsiooni ei tohiks ühte patta panna, vähemalt mitte uurimise algstaadiumis.

Konkreetset mentaalset mõistet on tavaliselt võimalik analüüsida kas kui fenomenaalset mõistet, kui psühholoogilist mõistet või kui nende kahe kombinatsiooni. Näiteks aistingut selle sõna peamises tähenduses on parem vaadelda kui fenomenaalset mõistet: aistida tähendab olla seisundis, millega kaasneb mingi tunne. Seevastu õppimist ja mälu on võib-olla parem vaadelda kui psühholoogilisi mõisteid. Õppimine tähendab esimeses lähenduses käitumuslike võimete sobivat kohandamist vastuseks keskkonnast saabuvatele kindlat liiki stiimulitele. Üldiselt iseloomustab vaimu fenomenaalseid omadusi tunne, mis subjektil selle omadusega seostub, psühholoogilisi omadusi aga nende roll käitumise põhjuslikes mehhanismides ja/või käitumise seletamises.

Loomulikult on termini “psühholoogiline” seesugune kasutusviis tinglik, tulenedes psühholoogia samastamisest kognitiivteadusega ülalöeldud mõttes. Igapäevane “psühholoogilise seisundi” mõiste on vististi avaram ja võib vägagi hästi hõlmata ka fenomenaalseid elemente. Kuid sellest, kuidas ma neid sõnu tarvitan, ei sõltu allpool midagi.

#### Pisut ajalugu

Vaimu fenomenaalsete ja psühholoogiliste aspektide samastamisel on pikk ajalugu. Osalt võib selle eest vastutavaks pidada René Descartes’i. Oma kurikuulsa doktriiniga vaimu läbipaistvusest iseendale jõuab ta väga lähedale mentaalsete ja fenomenaalsete aspektide samastamisele. Descartes’i järgi on iga vaimus aset leidev sündmus *cogitatio*​, mingi elamuse sisu. Sellesse klassi paigutas ta tahteaktid​, kavatsused ja igatliiki mõtted. Oma vastulauses “​Neljandatele vastuväidetele​” kirjutas ta:

Mis puudutab minu väidet, et vaimus, kuivõrd see on mõtlev asi, ei saa olla midagi sellist, millest ta teadlik ei oleks, siis see tundub mulle endastmõistetav. Sest vaimus, kui teda sellisena vaadelda, ei ole midagi, mis ei oleks mõte või ei sõltuks mõttest. See, mis ei ole mõte ega ​**[lk.13]**​ sõltu mõttest, ei kuulu vaimule kui mõtlevale asjale; ja meil ei saa olla ühtegi mõtet, millest me ei oleks teadlikud hetkel, mil ta meisse ilmub.

Kuigi Descartes psühholoogilisi ja fenomenaalseid aspekte tegelikult ei samastanud, eeldas ta ometi, et kõigil psühholoogilistel fenomenidel, mis väärivad vaimsete fenomenide nime, on teadvuslik aspekt.[[14]](#footnote-15) Mitteteadvusliku vaimuseisundi mõiste oleks Descartes’i õpetuse järgi vastuoluline.

Vaimu kahe aspekti eristamise eest võlgneme tänu rohkem psühholoogiliste teooriate kui filosoofia arengule. Veel sada aastat tagasi olid psühholoogid, nagu näiteks Wilhelm Wundt ja William James, äratuntavalt kartesiaanid, uurides käitumise põhjusi introspektsiooni teel ja toetudes psühholoogilisi teooriaid arendades introspektiivsetele andmetele. Sel moel andsid nad psühholoogiliste küsimuste üle otsustamisel otsustava hääle fenomenoloogiale. Kuid edasise arengu käigus hakati psühholoogilisi nähtusi peagi tunnustama iseseisva uurimisvaldkonnana.

Esmajoones Sigmund Freud ja tema kaasaegsed põhjendasid veenvalt ideed, et paljud vaimu tegevused on alateadlikud ning et inimesel võib olla mitteteadvuslikke uskumusi ja soove. Asjaolu, et säärane arusaam tundus neile loogilisena, tunnistab juba iseenesest, et nad analüüsisid mõtlemist mittefenomenaalselt. Tundub, et Freud käsitas neid mõisteid põhjuslikena. Näiteks soove mõisteti, seda otse välja ütlemata, enam-vähem kui seisundeid, mis kutsuvad esile mingit liiki käitumisakte seoses oma objektidega. Analoogiliselt mõisteti põhjuslikku rolli täitvatena ka uskumusi. Muidugi ei pannud Freud neid analüüse sõnadesse, kuid on ilmne, et tema mõistestik toetub mingitele sellesarnastele vaateviisidele. Otsesõnu aga ütles ta, et seisund ei pea tingimata jõudma teadvusse, et temaga saaks seletada käitumist, ning et selleks, et midagi saaks pidada sooviks või uskumuseks, ei pea tal olema teadvuslikku kvaliteeti. Arusaam vaimust, millele need järeldused toetuvad, on fenomenaalsetest mõistetest sõltumatu.

Sama ajal paiku kerkis psühholoogias esile ka biheivioristlik mõttesuund, mis hülgas introspektsionistliku traditsiooni täielikult. Töötati välja uus, “objektiivne” psühholoogiline seletusviis, milles ei olnud kohta teadvusel. Selle seletusviisi edu ei olnud täielik, kuid ta andis kinnitust seisukohale, mille järgi käitumist on võimalik psühholoogiliselt seletada ka fenomenaalset aspekti ignoreerides. Biheivioristide teoreetilised seisukohad ei olnud ühtsed: mõned uurijad tunnistasid teadvuse olemasolu, kuid ei pidanud seda psühholoogilistes seletustes oluliseks, mõned aga eitasid teadvuse olemasolu täielikult. Paljud läksid veelgi kaugemale ja eitasid üldse igasuguste vaimuseisundite olemasolu. Põhjenduseks öeldi, et käitumist saab seletada ainuüksi välistele andmetele toetudes ning et seesmised seisundid on seejuures irrelevantsed. Sügavam põhjus võis olla selles, et kõigil vaimuga seostuvatel mõistetel oli küljes põlu alla langenud fenomenaalsuse hõng.

​**[lk.14]**​ Igal juhul sai nende kahe mõttesuuna mõjul ortodoksseks arusaam, et käitumise seletamine ei sõltu mitte kuidagi fenomenaalsetest mõistetest. Nüüdseks on biheiviorismi asemele astunud informaatiline kognitiivteadus, kuid sama ortodoksia on jäänud üldjoontes kehtima. Kuigi selle arengu käigus anti mingi roll taas ka seesmistele seisunditele, mõnikord isegi vaimuseisundite nime all, ei olnud neis enam midagi spetsiifiliselt fenomenaalset. Neid seisundeid soostuti tunnistama just sellepärast, et neid sai kasutada käitumise seletamiseks; kõigist nendega kaasnevatest fenomenaalsetest kvaliteetidest vaadati parimal juhul lihtsalt mööda. Seega sattus tähelepanu keskpunkti vaimu psühholoogiline käsitlusviis.

Filosoofias formuleeris uue vaateviisi, kus rõhk ei olnud enam fenomenaalsetel, vaid psühholoogilistel aspektidel, Gilbert Ryle (1949), kes kinnitas, et kõik meie vaimsed mõisted saab taandada nendega seostuvatele käitumisviisidele või -soodumustele.[[15]](#footnote-16) Selles seisukohas — loogilises biheiviorismis — võib ära tunda paljude tänapäeva psühholoogilise filosoofia ortodokssete seisukohtade eelkäija. Muu hulgas formuleeris see suund kõige selgemini vaimsete mõistete seose käitumise põhjuslike mehhanismidega.

Ryle ei esitanud oma teooriat kui lihtsalt mõningate vaimsete mõistete analüüsi, vaid kavatses hõlmata sellega kõik vaimsed mõisted. Paljud on leidnud ja ka mina leian, et sellise vaateviisiga ei ole meie fenomenaalsete mõistete, näiteks aistingute ja ka teadvuse enda analüüsimisel midagi peale hakata. Paljude autorite arvates oli ilmne, et kui me kõneleme fenomenaalsetest seisunditest, siis ei mõtle me nende all kindlasti mitte käitumist või mõnd käitumissoodumust. Siiski andis Ryle’i analüüs meile paljutõotava meetodi paljude teiste vaimsete mõistete, näiteks uskumuste, naudingute, soovide, teeskluse ja mäletamise käsitlemiseks.

Lisaks raskustele fenomenaalsete seisunditega ei puudunud Ryle’i vaateviisil ka tehnilised raskused. Esiteks on loomulik pidada vaimu seisundeid käitumise põhjustajateks; ent kui nad on (erinevalt seesmistest seisunditest) ise käitumuslikud seisundid või soodumused, siis on raske näha, kuidas nad saaksid seda teha. Teiseks on Chisholm (1957) ja Geach (1957) väitnud, et ühtki vaimuseisundit ei saa defineerida ainult mingi käitumissoodumuste kogumi kaudu, kõigist teistest vaimuseisunditest sõltumatult. Näiteks kui ma usun, et väljas sajab vihma, siis oleneb minu käitumissoodumus sellest, kas ma soovin märjaks saada või mitte. See tähendab, et mingi kindla vaimuseisundiga kaasnevate käitumissoodumuste iseloomustamisel tuleb arvesse võtta ka teisi vaimuseisundeid.

Nendest raskustest oskas leidlikult üle saada *funktsionalism*​, mille töötas välja David Lewis (1966) ning arendas eriti põhjalikult edasi David Armstrong (1968).[[16]](#footnote-17) Selle vaate kohaselt määratleb vaimse seisundi täielikult tema põhjuslik roll: stiimulid, mis seisundit tavaliselt põhjustavad, käitumisviisid, mida ta tavaliselt põhjustab, ning vastastoimed teiste vaimuseisunditega. Nii tegi funktsionalism vaimuseisunditest täiesti seesmised seisundid, ​**[lk.15]**​ mistõttu neile sai nüüdsest hakata omistama sobivat põhjuslikku suhet käitumisega (mis andis vastuse esimene kahest ülaltoodud vastuväitest) ning andis võimaluse arvestada vaimsete seisundite määratlemisel nende omavahelisi vastastoimeid (vastuseks teisele vastuväitele).

Selle vaateviisi järgi saab meie vaimseid mõisteid analüüsida *funktsionaalselt*​, nende aktuaalsete või tüüpiliste põhjuste ja tagajärgede kaudu. Konkreetsete vaimsete mõistete analüüsimine säärasel viisil ei ole kaugeltki triviaalne ülesanne; Armstrong (1968) pakub välja terve hulga analüüse, kuid need on kõik väga lünklikud. Põhimõtteliselt aga võib funktsionalism anda paljudele meie vaimsetele mõistetele, vähemalt kuivõrd need mängivad rolli käitumise seletamisel, päris vastuvõetavaid tõlgendusi. Näiteks õppimist saab analüüsida kui käitumuslike võimete adapteerimist vastuseks keskkonnast saabuvatele stiimulitele. Või keerulisem näide: uskumust, et väljas sajab vihma, võib väga umbkaudselt analüüsida kui seisundit, mis tekib tavaliselt siis, kui väljas sajab vihma; mis kutsub esile käitumise, mis on omal kohal siis, kui sajab vihma; mis on mingites kindlates loogilistes seostes teiste uskumuste ja soovidega; jne. Üksikasjade väljatöötamiseks on siin veel tublisti ruumi, kuid üldine suund on paljude uurijate arvates õige.

Kuid nagu Ryle, nii ka Armstrong ja Lewis ei esitanud oma teooriat kui üksikute vaimsete mõistete analüüsi, vaid nende kavatsuse järgi pidi sellest saama kõigi vaimsete mõistete analüüs. Nende väitel on sellisel viisil võimalik analüüsida teiste hulgas ka elamuse, aistingu, teadvuse jt. mõisteid. Selline fenomenaalsete aspektide sulandamine psühholoogilistesse oli minu arvates sama suur viga nagu see, mille tegi Descartes, kui ta sulandas psühholoogilised aspektid fenomenaalsetesse. See on lihtsalt fenomenaalsete aspektide ebaõige analüüs. Kui me küsime endalt kellegi kohta, kas tal on hetkel mingi värvuselamus, siis ei ole küsimus selles, kas ta võtab vastu keskkonnast saabuvaid stiimuleid ja töötleb neid mingil kindlal viisil, vaid meid huvitab, kas tal on värvusaistingule vastav elamus, ja see on hoopis teine küsimus. Võimalus, et mingi aisting mängib põhjuslikku rolli, ilma et temaga kaasneks elamus, on kontseptuaalselt vastuolutu.

Teisiti öeldes, seesugune fenomenaalsete mõistete analüüs jätab vastamata küsimusele, miks peaks see probleem üldse kellelegi korda minema.[[17]](#footnote-18) Teaduslik seletamine, kuidas mingi seisund võib mängida nii- või teistsugust põhjusliku rolli, võib muidugi põrkuda tehnilistele raskustele, kuid mingit erilist mõistatust selles ei ole. Mõistatus on, miks peaks olema olemas tunne olla selles seisundis — miks peaks seisundil olema fenomenaalne kvaliteet. Miks mängib seisund põhjuslikku rolli ja miks tal on fenomenaalne kvaliteet — need on kaks täiesti erinevat küsimust. Funktsionalistlik analüüs siin erinevust ei näe ja jääb seepärast mitterahuldavaks.

Võtan selle küsimuse allpool palju põhjalikuma uurimise alla, esialgu aga võime tähendada, et kuigi funktsionalism ei saa rahuldavalt hakkama fenomenaalsete mõistete analüüsiga, suudab ta anda hea analüüsi teistele vaimsetele mõistetele, näiteks õppimisele ja mälule, võib-olla ka uskumustele. Nendel juhtudel ​**[lk.16]**​ analoogilisi raskusi ei teki. Selles, et mingi süsteem on võimeline õppima, ei ole sugugi rohkem salapära kui selles, et ta suudab kohandada oma käitumist keskkonnast saabuvatele stiimulitele; õigupoolest ongi need kaks enam-vähem üks ja sama. Ja kui me küsime endalt, kas keegi mingis kindlas olukorras midagi õppis, siis on küllap mõistlik öelda, et meid huvitab, kas temaga on toimunud mõni muutus, mille tagajärjel ta suudab seesuguste olukordadega paremini toime tulla. Õppimise mõiste põhjalik analüüs oleks muidugi palju üksikasjalikum kui siinne esimene lähendus, kuid sama mudel võimaldab analüüsi ka edasi arendada.

Tõepoolest, funktsionalistlik käsitlus vastab täpselt psühholoogiliste omaduste ülaltoodud [definitsioonile](#id_22psühholoogiline). Enamik mittefenomenaalseid vaimseid omadusi kuulubki sellesse klassi ja seepärast on nad ka funktsionaalselt analüüsitavad. Muidugi võib iga funktsionalistliku analüüsi puhul alati vaielda üksikasjade üle. Jäävad ka mõned olulised põhimõttelised küsimused, näiteks küsimus keskkonna rollist psühholoogiliste omaduste iseloomustamisel ning küsimus, kas otsustav ühenduslüli psühholoogiliste omaduste ja käitumise vahel on põhjuslik või seletuslik või mõlemat korraga. Need on aga juba üksikasjad, mis pole siinkohal kuigi olulised. Oluline on, et mittefenomenaalseid vaimuseisundeid iseloomustab suuresti nende roll meie kognitiivses majapidamises.

Kogu ülaltoodud arutluse moraal on, et vaimul on reaalselt nii psühholoogilisi kui ka fenomenaalseid aspekte ning et need on kaks ise asja. Esimeses lähenduses võib öelda, et fenomenaalsed mõisted vaatavad vaimule esimese, psühholoogilised kolmanda isiku aspektist. See, kuidas me vaimu käsitleme, oleneb oluliselt sellest, kummad vaimu aspektid meile huvi pakuvad. Kui meid huvitab vaimu roll käitumise kujundamises, koondub meie tähelepanu psühholoogilistele omadustele. Kui meid huvitab aga vaimuseisundite teadvuslik läbielamine, keskendume fenomenaalsetele omadustele. Mitte kumbagi aspekti, ei fenomenaalset ega psühholoogilist, ei tohi teise kaudu olematuks defineerida. On mõeldav, et mingi sügava analüüsi varal oleks võimalik tuua nähtavale fundamentaalne ühenduslüli fenomenaalsete ja psühholoogiliste aspektide vahel. Kuid sellise lüli leidmine ei oleks triviaalne ülesanne, mille saaks lihtsalt mingi kokkuleppega ära sätestada. Sulandades fenomenaalsed aspektid psühholoogilistesse, senikaua kui meil puudub nende kohta asja põhjani tungiv seletus, lamestaksime teadvuselamuse probleemi; sulandades psühholoogilised aspektid fenomenaalsetesse, jätaksime vaimu rolli käitumise seletustes liiga napiks.

### 3. Vaimsete mõistete kaksikelu

On vist mõistlik öelda, et psüühika psühholoogiline ja fenomenaalne külg üheskoos ammendavad kogu vaimu. See tähendab, et iga vaimne omadus on kas fenomenaalne ​**[lk.17]**​ omadus, psühholoogiline omadus või mingi nende kahe kombinatsioon. Vaimu omadusi, mis silmanähtavalt kisendavad seletuse järele, tundub olevat kaks: teadvuselamuse mitmesugused vormid ja käitumise põhjuslikud mehhanismid. Mingit kolmandat *explanandum*​it ei paista kusagil, ja kaks esimest tõendusmaterjali allikat — elamused ja käitumine — ei anna alust uskuda mingite kolmandat liiki omaduste olemasolusse, mis poleks ei fenomenaalsed ega funktsionaalsed. (Väikese erandi moodustavad võib-olla relatsioonilised omadused, millest [peagi](#id_24relational) samuti juttu tuleb). Kindlasti on olemas ka teisi vaimuseisundite klasse, millest me sageli kõneleme — intentsionaalsed seisundid, emotsionaalsed seisundid jne. —, kuid neid on usutavasti võimalik taandada psühholoogilistele või fenomenaalsetele seisunditele või mingile nende kahe kombinatsioonile.

Asja teeb keeruliseks see, et paljud igapäevased vaimsed mõisted ei lahterdu puhtalt ühe liigi alla neist kahest, vaid neil on nii fenomenaalne kui ka psühholoogiline komponent. Heaks näiteks võib tuua valu. Selle mõistega nimetatakse sageli üht eriti ebameeldivat fenomenaalsete kvaliteetide liiki, ja sellisel juhul tõuseb esiplaanile fenomenaalne mõiste. Kuid seesama termin võib tähistada ka psühholoogilist mõistet: siis mõistetakse selle all üldiselt seisundit, mis kaldub organismi kahjustama, kutsuma esile vastumeelsusreaktsioone jne. Mõlemal nimetatud aspektil on valu tavapärases mõistes keskne koht. Selle kohta võib öelda, et valu mõiste lahkneb kaheks, fenomenaalseks mõisteks ja psühholoogiliseks mõisteks, või et need on üheainsa sisurikka mõiste kaks komponenti.

Küsides, kumb on valu juures olulisem, kas fenomenaalne kvaliteet või funktsionaalne roll, mässime end paljudesse keerdsõlmedesse. Näiteks: kas kujuteldaval süsteemil, mis rahuldaks kõiki funktsionaalseid kriteeriume, kuid millel puuduks teadvuselamus, oleks päriselt valus? Võib-olla on ahvatlev vastata sellele küsimusele eitavalt, aga mida teha valuga, mis kestab päev otsa, kuid vaheaegadega, mil valu teadvusse ei jõua? Ei ole mõtet hakata siin midagi üht- või teistpidi ette otsustama. Semantilisest otsusest, kas mingi fenomenaalne kvaliteet on tõesti nii hädavajalik, et lugeda üht või teist seisundit valuks, ei olene tegelikult midagi olulist. Õigem oleks tunnistada, et valu mõistel on mitu komponenti, ja nende vahel selge sõnaga vahet teha, kõneldes näiteks fenomenaalsest valust ja psühholoogilisest valust. Meie igapäevases arusaamas ühinevad need kaks mõistet arvatavasti mingis kindlas vahekorras, kuid filosoofiline arutelu kulgeb selgemini, kui teha nende vahele selge vahe.

Põhjus, miks fenomenaalseid ja psühholoogilisi omadusi sageli ühte patta pannakse, on ilmne: sellepärast, et nad käivad enamasti käsikäes. Kui kudede kahjustumine põhjustab mingi protsessi, mis toob kaasa vastumeelsusreaktsiooni, realiseerub tavaliselt ka mingi fenomenaalne kvaliteet. See tähendab, et psühholoogilise valuga käib tavaliselt kaasas fenomenaalne valu. See, et selle protsessiga kaasneb fenomenaalne kvaliteet ​**[lk.18]**​, ei ole kontseptuaalne tõde, kuid see on fakt maailma kohta. Kui need omadused esinevad igapäevastes olukordades koos, siis on loomulik, et neid seovad omavahel ka meie argimõisted.

Taolist kaksikelu elavad paljud vaimsed mõisted. Näiteks tajumise mõistet võib mõista puht psühholoogiliselt — viitavana protsessile, mille käigus kognitiivsüsteem võtab keskkonnast saabuvaid stiimuleid vastu niisugusel viisil, et seisundid, mis temas selle tagajärjel tekivad, mängivad teatavat rolli kognitiivsete protsesside suunamisel. Kuid sama mõistet võib tõlgendada ka fenomenaalselt, viitavana tajutava sisu teadvuslikule läbielamisele. Viimase tõlgenduse vastu räägib subliminaalse taju võimalikkus, kuid me võime öelda, et säärased tajud on tajud vaid sõna nõrgemas tähenduses. Ent seegi on vaid terminoloogiline küsimus. Kui me tahame, et meie jutt oleks selge, võime lihtsalt iga kord kokku leppida, kas me kõneleme psühholoogilisest omadusest, fenomenaalsest omadusest või mingist nende mõlema kombinatsioonist.

Siiski kalduvad mõned neist kahepalgelistest mõistetest rohkem fenomenaalsuse, mõned rohkem psühholoogilisuse poole. Võtame näiteks aistingu mõiste, mis on tihedas suguluses taju mõistega ja millel on samuti nii fenomenaalne kui ka psühholoogiline komponent. Fenomenaalne komponent tõuseb aistingutes palju selgemini esile kui tajudes, nagu näitab asjaolu, et mitteteadvuslike tajude idee on tundub meile mõttekamana kui mitteteadvuslike aistingute idee. Ent ka siin ei ole asi päris ühene: sõnal “taju” on ka tähendusvarjund, mis eeldab teadvuslikku läbielamist, ja sõnal “aisting” on ka varjund, mis seda ei eelda; kuid need varjundid tunduvad olevat palju vähem olulised kui nende alternatiivid. Sõna “taju” on võib-olla loomulikum tarvitada psühholoogilise terminina, sõna “aisting” aga fenomenaalsena. Sellisel juhul võiksime aistingut pidada otsekui taju fenomenaalseks vasteks.

Heaks testiks selle kohta, kas mingi vaimne mõiste ​*M* on pigem psühholoogiline või pigem fenomenaalne, on küsimus: kas mingi nähtus võiks olla ​*M* erijuht, kui temaga ei kaasne ühtki eriliselt fenomenaalset kvaliteeti? Kui jah, siis on ​*M* pigem psühholoogiline mõiste. Kui ei, siis on ta fenomenaalne mõiste või vähemalt nende kahe kombinatsioon, milles on kesksel kohal fenomenoloogiline külg. Viimast võimalust ei saa välistada, sest on mõisteid, mis hõlmavad nii fenomenaalset kvaliteeti kui ka sobivat kognitiivset rolli; näiteks sõna “aisting” kesksel tähendusel ongi selles mõttes nähtavasti mõlemasugune iseloom. Kuid vähemalt saab niiviisi teha vahet mõistetel, millel on ka fenomenoloogiline külg, mõistetest, millel seda ei ole.

Esitades ülaltoodud testküsimuse näiteks õppimise kohta, näeme, et see on suurelt jaolt psühholoogiline mõiste. Esimeses lähenduses tähendab õppimine lihtsalt kognitiivsete võimete mingil kindlal viisil kohandamist mitmesuguste uute olukordade ja stiimulitega. Selleks, et üks või teine kognitiivne protsess oleks õppimine, ei pea tal olema mingit erilist fenomenaalset kvaliteeti; see võib küll olemas olla, kuid mitte see ei tee temast õppimist. Õppimisel võib olla mingi kerge fenomenaalne varjund, tulenevalt seosest mõnede teiste mõistetega nagu näiteks uskumused, ​**[lk.19]**​ millest tuleb juttu [allpool](#PARA.24-10). Ent see varjund on parimal juhul väga nõrk. Õppimist seletades tuleb eelkõige seletada, kuidas süsteemil õnnestub sobival viisil adapteeruda. Midagi sarnast võib öelda ka kategoriseerimise ja mälu kohta, mis on nähtavasti suuremas osas psühholoogilised mõisted, kuna neis on kesksel kohal kognitiivne roll.

Emotsioonide puhul on fenomenaalne aspekt palju ilmsem. Rõõmule või kurbusele mõeldes tuleb meile meelde ka mingit kindlat laadi teadvuselamus. Pole aga siiski päris selge, kas selleks, et mingi seisund oleks emotsioon, peab tal tingimata olema ka fenomenaalne aspekt, sest kahtlemata kaasneb emotsiooniga ka mingi tugev psühholoogiline omadus. Nagu ennegi, nii ka siin ei ole meil tarvis selles küsimuses midagi ette ära otsustada. Me võime kõnelda lihtsalt emotsioonide psühholoogilistest ja fenomenaalsetest aspektidest ja täheldada, et nende kahega on ammendatud kõik emotsioonide seletust nõudvad aspektid.

Kõige keerukama juhu moodustavad nn. ​*propositsionaalsed hoiakud*​ — s.t. hoiakud propositsioonidele maailma kohta —, näiteks uskumused. Kui ma näiteks usun, et Bob Dylan on tulemas kontsertreisile Austraaliasse, siis olen ma teatava propositsiooniga Bob Dylani kohta nõus; kui ma seda aga loodan, siis on mul samale propositsioonile teistsugune hoiak. Nende vaimuseisundite keskseks eripäraks on semantiline aspekt ehk *intentsionaalsus*​: see, et nad ütlevad midagi olukorra kohta maailmas. See tähendab, et uskumustel on semantiline sisu: minu äsja mainitud uskumuse umbkaudseks sisuks on (kuigi selle üle võib diskuteerida) propositsioon “Dylan on tulemas kontsertreisile Austraaliasse”.

Uskumusi peetakse kõige sagedamini psühholoogilisteks omadusteks. Selle arusaama kohaselt tähendab mingi propositsiooni tõesusse uskumine kõige esimeses lähenduses seisundit, kus inimene käitub nii, nagu oleks sobiv käituda juhul, kui see propositsioon oleks tõene, seisundit, mida selle propositsiooni tõesus arvatavasti esile kutsuks, ning seisundit, kus arutlemise kognitiivne dünaamika kajastab selle seisundi sobivaid vastastoimeid teiste uskumuste ja soovidega. Kuid uskumuse funktsionaalsed kriteeriumid on üpris raskesti tabatavad, ja mitte keegi ei ole veel esitanud nende kriteeriumide ligilähedaseltki täielikku analüüsi. Sellegipoolest on põhjust arvata, et kirjeldatud arusaam hõlmab suure osa uskumuse olulistest tahkudest. Enam-vähem sama ütleb ka idee, et uskumus on midagi *seletuskonstrukti* taolist: me omistame teistele inimestele uskumusi peamiselt selleks, et seletada nende käitumist.

Mõni ütleb võib-olla, et see pilt pole täielik ning et uskumuse mõiste alla kuulub veel mingi hädavajalik komponent lisaks vastavale psühholoogilisele protsessile. Eelkõige on pildist välja jäänud uskumise elamuslikud aspektid, milleta, nagu mõni uurija väidab, uskumust olla ei saa. Näiteks Searle (1990a) on väitnud, et uskumuse intentsionaalne sisu sõltub täielikult uskumusega kaasnevast teadvusseisundist või siis teadvusseisundist, mida uskumus võib esile kutsuda. Ilma teadvuseta on olemas vaid “justkui”-intentsionaalsus.[[18]](#footnote-19) [[19]](#footnote-20)

​**[lk.20]**​ On selge, et uskumustega kaasnevad sageli teadvuselamused: on olemas tunne midagi päriselt (s.t. teadvuslikult) uskuda, ja mitteteadvuslik uskumus võib teadvuslikku uskumust enamasti vähemalt esile kutsuda. Otsustav küsimus on aga, kas see teadvuslik kvaliteet ongi see, mis teeb seisundist uskumuse ja annab uskumusele tema sisu. Mõne uskumuse puhul tundub see võib-olla usutavamana kui teiste puhul: näiteks võib vist öelda, et teadvuslik kvaliteet on vajalik selleks, et uskuda midagi tõeliselt omaenda läbielamiste kohta, ning võib ka olla, et mingit kindlat liiki elamusi eeldavad mõned pertseptuaalsed uskumused välismaailma kohta (näiteks: võib-olla on selleks, et uskuda, et mingi ese on punane, tarvilik, et ta kutsuks esile punaseid elamusi?). Teistel juhtudel ei ole see nii ilmne: näiteks kui ma pean Don Bradmanit kõigi aegade suurimaks kriketimängijaks, siis oleks mul tõenäoliselt sama uskumus ka siis, kui mul seostuksid temaga hoopis teistsugused teadvuselamused. Uskumuste fenomenoloogia on üsna ebaselge ja pole kerge mõista, kuidas võiks üks või teine fenomenaalne kvaliteet teha mainitud uskumusest uskumuse Bradmani kohta. Tundub, et uskumuse sisu määravad ära pigem seos uskumuse ja Bradmani vahel ning roll, mida see seos mängib minu kognitiivses süsteemis.

Võib asuda ka nõrgemale positsioonile ja öelda, et kuigi ükski konkreetne uskumus ei eelda ühtki konkreetset fenomenaalset kvaliteeti, peab subjektil selleks, et üldse midagi uskuda, olema vähemalt võime midagi teadvuslikult läbi elada.[[20]](#footnote-21) Ideel, et teadvusliku siseeluta olendil ei saaks olla tõelisi uskumusi, vaid parimal juhul pseudouskumusi, ei puudu veenvus. Kuid see teeks fenomenaalse kvaliteedi rolli intentsioonilistes mõistetes üsna kasinaks. Kõige olulisemad mingiks konkreetseks uskumuseks tarvilikud tingimused jäävad fenomenaalsest sfäärist välja. Uskumustest saaks igasuguse fenomenaalse komponendi isegi täiesti välja jätta; siis jääks alles pseudouskumuse mõiste, mis kõigis kõige olulisemates aspektides sarnaneks uskumusega, kuid milles ei figureeriks teadvuse mõiste. Tõepoolest tundubki, et säärased pseudouskumused suudaksid täita suurema osa uskumuse mõiste seletuslikust rollist.

Olgu sellega kuidas on, siinkohal ei püüa ma neis rasketes küsimustes intentsionaalsuse ja teadvuse kohta veel midagi kindlalt otsustada. Võib mainida, et on olemas vähemalt *deflatiivne*​ uskumuse mõiste, mis on puhtalt psühholoogiline ega haaku mitte kuidagi teadvuselamusega; kui mingi olend on niisuguses psühholoogilises seisundis, siis sarnaneb tema seisund nii mõneski suhtes uskumusega, kui mitte arvestada fenomenaalseid aspekte. On olemas ka uskumuse *inflatiivne* mõiste, mille kohaselt tõeline uskumine eeldab teadvuselamust ning võib-olla veel enamatki — et selleks, et mingit propositsiooni tõeliselt uskuda, on tarvilikud mingit kindlat liiki teadvuselamused. Minu eesmärkide seisukohalt pole kuigi oluline, missugune neist on uskumuse mõistetest on “õige”.

Oluline on siin see, et uskumustel pole ühtki omadust, mis poleks kas fenomenaalne või psühholoogiline. Võib-olla tuleb ​**[lk.21]**​ siin teha väike mööndus ning lisada ka *relatsiooniline* element uskumuste puhuks, mis sõltuvad lisaks subjekti seesmisele seisundile ka olukorrast tema keskkonnas. Näiteks on väidetud, et selleks, et uskuda, et vesi on märg, peab subjektil olema mingi kindel suhe oma keskkonnas leiduva veega. Tavaliselt eeldatakse, et see suhe on põhjuslik, ja seda asjaolu saaks kajastada juba vastava psühholoogilise omaduse määratluses, mainides selles asjaomaseid põhjuslikke rolle, mis ulatuvad inimese peast väljapoole, tema keskkonda. Sellisel juhul ei ole tarvis enam ühtki uut komponenti. Igatahes ei esita tähelepanek, et mõnel psüühilisel seisundil võib lisaks psühholoogilisele ja fenomenaalsele komponendile olla ka relatsiooniline komponent, meile mingeid erilisi nõudmisi. Uusi sügavaid müsteeriume sellest juurde ei tule.

Näitamaks, et intentsionaalsetel seisunditel peale fenomenaalsete ja psühholoogiliste/relatsiooniliste aspektide mingeid muid süvaaspekte ei ole, paneme tähele, et kõik seletust nõudvad fenomenid​[[21]](#footnote-22) jagunevad kahte klassi: fenomenideks, mis on meile kättesaadavad kolmanda isiku perspektiivist, ja fenomenideks, millele pääseb ligi esimese isiku perspektiivist. Esimesse klassi kuuluvad fenomenid taanduvad viimselt käitumisele, suhetele keskkonnaga jne., ja neid võib liigitada psühholoogilisteks ja relatsioonilisteks fenomenideks. Teise klassi fenomenid taanduvad uskumisega kaasnevatele elamustele — näiteks sellele, kuidas meie mõisted fenomenaalses maailmas avalduvad — ning kuuluvad seega ühe osana teadvuse probleemi juurde, toomata juurde mingit uut müsteeriumi. Uskumise põhjused igas konkreetses (sh. sealhulgas semantilise sisu, uskumuse objekti jne.) aspektis tulenevad alati kas ühte või teise klassi kuuluvatest fenomenidest; peale nende kahe ei ole ühtki muud iseseisvat fenomenide klassi, mis meilt seletust ootaksid.

Selles võib selgusele jõuda ka teisiti — pannes tähele, et pärast indiviidi psühholoogiliste, fenomenaalsete ja relatsiooniliste omaduste fikseerimist ei jää nähtavasti enam ühtki sõltumatult varieeritavat vaimset omadust. Ehkki tundub olevat täiesti võimalik kujutleda minuga psühholoogiliselt identset, kuid teistsuguste elamustega olendit, ei suuda me analoogiliselt kujutleda olendit, kes oleks minuga kolmes nimetatud omaduste liigis identne, ent erinev oma uskumustelt. Kontseptuaalses ruumis pole sellisele võimalusele lihtsalt kohta. Intentsionaalsed mõisted on mingis mõttes vähem fundamentaalsed kui psühholoogilised ja fenomenaalsed mõisted, kuna neid ei saa varieerida viimastest sõltumatult.[[22]](#footnote-23)

Kõik, mida ma ütlesin siin uskumuste kohta, kehtib samavõrra ka teiste intentsionaalsete seisundite, näiteks soovide, lootuste jne. kohta. Kõigil neil seisunditel on psühholoogiline ja fenomenaalne aspekt ning meil ei ole mingit vajadust otsustada, kumb neist on primaarne, kuigi saaks esitada veenvaid argumente psühholoogilise analüüsi kasuks. Oluline on see, et neil seisunditel pole ühtki aspekti, mis poleks kas psühholoogiline või fenomenaalne (sisaldades võib-olla ka relatsioonilist ​**[lk.22]**​ komponenti). Psühholoogilised ja fenomenaalsed aspektid ühtekokku ongi vaimu kesksed aspektid.

#### Fenomenaalsete ja psühholoogiliste omaduste koosolu

On inimvaimu fakt, et alati, kui realiseerub mingi fenomenaalne omadus, realiseerub ka vastav psühholoogiline omadus. Teadvuselamus ei kulge vaakumis. Elamusega kaasnevad alati kognitiivsed protsessid, millest ta tõenäoliselt mingis mõttes välja kasvabki. Näiteks aistingu puhul toimub alati vastav infotöötlus, mida võib ehk nimetada tajumiseks. Või kui me elame teadvuslikult läbi rõõmu, siis kaasneb sellega tavaliselt mingi seesmine seisund, mis mängib rõõmule vastavat funktsionaalset rolli. Ei saa välistada loogilist võimalust, et leidub läbielamisi, mille taga ei ole mingit põhjuslikku mehhanismi. Kuid tundub olevat empiiriline fakt, et need kaks käivad käsikäes.

Säärase koosolu tõttu võib mõni arglikum mõtleja hakata kahtlema, kas siin üldse mingit reaalset vahet on. Vähemasti kontseptuaalne erinevus on siin siiski selgelt olemas, kuigi asjaomaste mõistete mahud näivad kattuvat. Võib küsida, kuidas seletada fenomenaalset kvaliteeti, ja võib küsida, kuidas seletada põhjuslikku rolli, ja need on erinevad küsimused.

Sellegipoolest kajastab fenomenaalsete ja psühholoogiliste omaduste koosolu midagi olulist meie fenomenaalsetes mõistetes. Meil puudub sõltumatu keel fenomenaalsete kvaliteetide kirjeldamiseks. Me oleme näinud, et neis on midagi tabamatut. Kuigi rohelisus on omaette aisting, millel on omaenda külluslik seesmine iseloom, ei oska me tema kohta kuigi palju öelda peale selle, et ta on roheline. Fenomenaalsetest kvaliteetidest rääkides ei saa me kõnealuseid omadusi kirjeldada teisiti kui nendega kaasnevate väliste omaduste või nende põhjusliku rolli kaudu. Keel, milles me kõneleme fenomenaalsetest kvaliteetidest, on tuletis meie mittefenomenaalsest keelest. Või, nagu ütles Ryle, aistingute jaoks ei ole “puhtaid” sõnu.

Ülaltoodud teadvuselamuste kataloogi vaadates näeme, et ma ei kirjeldanud elamusi kordagi nende seesmiste kvaliteetide kaudu, vaid väljenditega nagu “ahjusooja leiva aroom”, “mustrid, mida näeb silmi sulgedes” jne. Isegi sellise väljend nagu “roheline aisting” kirjeldab oma referenti tegelikult alles millegi välise kaudu. Me omandame selle väljendi ju ostensiivselt, õppides rakendama teda elamustele, mida kutsuvad esile rohi, puud jne. Ja üldse, niivõrd kui meie fenomenaalsetest kategooriatest on üldse võimalik teistele teateid anda, toimub see alati kas nendega seonduvate tüüpiliste väliste korrelaatide või nendele vastava psühholoogilise seisundi kaudu. Näiteks rõõmu fenomenaalsest kvaliteedist kõneldes seostatakse sõna “rõõm” tegelikult mingi põhjusliku rolliga — olukorraga, kus ​**[lk.23]**​ inimene tunneb, et kõik on hästi, hüppab rõõmust vms. See ongi ehk Wittgensteini kuulsa ütluse “Seesmine protsess vajab väliseid kriteeriume” üks võimalikke tõlgendusi.[[23]](#footnote-24)

Fenomenaalsete mõistete sõltumine põhjuslikest kriteeriumidest on viinud mõned autorid (nende hulgas aeg-ajalt Wittgensteini ja Ryle’i) oletusele, et meie vaimsete mõistete taga polegi midagi peale nendega seonduvate põhjuslike kriteeriumide. Säärasel arvamusel ei puudu veenvus: kui fenomenaalsele omadusele osutatakse alati psühholoogilise omaduse kaudu, miks siis mitte oletada, et need kaks ongi üks ja sama omadus? Kuid sellele kiusatusele tuleb vastu seista. Kõneldes rohelisest aistingust, ei kõnele me lihtsalt “seisundist, mida põhjustavad rohi, puud jne.”, vaid fenomenaalsest kvaliteedist, mis tavaliselt tekivad, kui rohi ja puud meis mingi seisundi põhjustavad. Kui põhjuslik analüüs siin üldse võimalik on, siis peab see olema midagi taolist nagu “fenomenaalsete seisundite liik, mida põhjustavad rohi, puud jne.”.[[24]](#footnote-25) [[25]](#footnote-26). Selles mõistes sisalduva fenomenaalse elemendi tõttu ei saa teda analüüsida ainuüksi funktsionaalselt.

Ja üldse, kui me osutame mingile fenomenaalsele omadusele psühholoogilise omaduse​ ​*P* kaudu, siis ei kattu fenomenaalne mõiste lihtsalt mõistega “​*P*​”, vaid mõistega “seesugune teadvuselamus, mis tavaliselt kaasneb *P*​-ga”. Oluline on just see, et fenomenaalse kvaliteedi või teadvuselamuse mõistet ennast ei saa defineerida psühholoogiliste mõistete kaudu, vaid temas on, nagu nägime, midagi elementaarset. Kui elamuse või fenomenaalse kvaliteedi mõistet oleks võimalik analüüsida funktsionaalselt, siis saaks funktsionaalselt analüüsida ka konkreetseid fenomenaalseid omadusi, kuid sellise analüüsi puudumisel niisugust järeldust teha ei saa.

Me ei saa samastada mõistet “fenomenaalne ​*P*​” mõistega “psühholoogiline ​*P*​” põhjustel, mida me juba teame: need on kaks erinevat mõistet, sest siin on kaks eri *explanandum*​it. Kuigi “fenomenaalset *P*​-d” saab määratleda kui “elamust, mis tavaliselt kaasneb psühholoogilise ​*P*​-ga”, on vägagi hästi ja vastuolutult võimalik ette kujutada olukorda, kus fenomenaalne *P* on olemas, kuid temaga ei kaasne psühholoogilist *P*​-d, ja ümberpöördult. Rolls-Royce’i logot võib jämedalt defineerida kui logot, mida võib tavaliselt näha Rolls-Royce’i autodel, kuid see ei tähenda, et Rolls-Royce’i logo on Rolls-Royce’i auto.

Kõige selle põhjal võime näha, kui kasin on meie spetsiifiliselt fenomenaalne sõnavara psühholoogilise sõnavaraga võrreldes, ning mõista, miks on fenomenaalseid omadusi nii sageli pandud ühte patta psühholoogilistega. Enamikus igapäevastes asjades ei ole see erinevus oluline: kui me ütleme kellegi kohta, et ta on rõõmus, siis ei pea me tingimata silmas pidama just üht kahest — kas fenomenaalset kvaliteeti või funktsionaalset rolli —, sest need käivad tavaliselt käsikäes. Kuid filosoofilistes mõttekäikudes ja eriti seletustes on nende segiajamine ränk viga. Meil võib olla ahvatlus seda teha, ​**[lk.24]**​ kuna sellest eristusest loobudes muutub teadvuselamuse seletamine korraga väga lihtsaks. Kuid täpselt samal põhjusel jääb niiviisi saadav seletus ka täiesti mitterahuldavaks. Teadvuse probleemist ei ole võimalik lahti saada ainuüksi verbaalsete võtetega.

### 4. Kaks keha–vaimu probleemi

Vaimsete omaduste jagunemise tõttu fenomenaalseteks ja psühholoogilisteks jaguneb keha–vaimu probleem kaheks osaks: kergeks ja raskeks. Vaimu psühholoogilistest aspektidest tuleneb terve hulk tehnilisi probleeme kognitiivteaduses ja hulk huvitavaid keerdküsimusi filosoofias, kuid sügavaid metafüüsilisi saladusi neis ei ole. Küsimus “Kuidas võib füüsiline süsteem olla võimeline õppima ja mäletama?” ei ole sama akuutne kui vastav küsimus aistingute või üldse teadvuse kohta. Põhjus on selge. Õppimine ja mälu on käesoleva peatüki [3. jaos](#PARA.24-8) toodud analüüsi järgi funktsionaalsed omadused, mida iseloomustavad nende põhjuslikud rollid, mistõttu küsimus “Kuidas võib füüsilisel süsteemil olla nii- ja niisugune psühholoogiline omadus?” on samaväärne küsimusega “Kuidas võib mingil füüsilise süsteemi seisundil olla nii- või niisugune põhjuslik roll?”. See on küsimus füüsilisi süsteeme uurivatele teadustele, ja sellele vastamiseks tuleb lihtsalt jutustada, kuidas füüsikalise süsteemi organisatsioon võimaldab tal reageerida keskkonnast saabuvatele stiimulitele ja käituda sobival viisil. Kuigi sellega seoses kerkib tohutuid tehnilisi probleeme, saab neile lahenduse leida selgepiirilise uurimisprogrammi raames. Metafüüsilisi probleeme on siin suhteliselt vähe.

See ei tähenda, nagu ei valmistaks psühholoogilised omadused üldse mitte mingeid filosoofilisi raskusi. Märkimisväärseid raskusi on näiteks juba nende mõistete korrektse analüüsimisegagi. Kuigi peaaegu kõik on nõus, et need on funktsionaalsed mõisted, võib küsimuses, kuidas nimelt peaks vajalik funktsionaalne analüüs toimuma, olla kaunis suuri lahkarvamusi. Eriti viljakat pinda vaidlusteks pakuvad mõned intentsionaalsed omadused, näiteks uskumused ja soovid. Näiteks ei mõisteta tänini kuigi hästi, milles nimelt seisneb ühe või teise intentsionaalse seisundi sisu. Tehnilisi raskusi valmistab ka küsimus, kuidas nimelt saavad selletaolised kõrgtaseme konstruktid mängida reaalset põhjuslikku rolli käitumise kujundamisel, eriti kui neid mingis osas konstitueerivad keskkonna omadused või kui pole rangeid seadusi, mis seoksid psühholoogilisi seisundeid käitumisega. Veel on kognitiivteaduste alustes vastuseta osalt empiirilised küsimused, kuidas need omadused olemasolevates kognitiivsetes süsteemides realiseeruvad, või kas nad isegi üldse realiseeruvad.

Kõik need on tõsised probleemid, kuid oma olemuselt on nad pigem keerdküsimused kui müsteeriumid. Olukord on analoogiline olukorraga bioloogia ​**[lk.25]**​ filosoofias, kus ei ole pakilist elu–keha probleemi, küll aga rohkesti tehnilisi probleeme, mis puudutavad evolutsiooni, valikut, kohanemist, kohastumusi ja liike. Analoogiliselt bioloogiaga, kus on ammu vabanetud suuremast osast näilistest metafüüsilistest müsteeriumidest, võib öelda, et psühholoogias on keha–vaimu probleem põhimõttelselt lahendatud. Jäänud on hulk väiksemaid tehnilisi probleeme, millega rutiinse teadusliku ja filosoofilise analüüsi käigus arvatavasti pikapeale toime tullakse.

Hoopis teine lugu on vaimu fenomenaalsete aspektidega. Siin on keha ja vaimu probleem sama nõutukstegev, nagu ta on alati olnud. Füüsika ja kognitiivteaduste muljetavaldav areng ei ole heitnud oluliselt valgust küsimusele, kuidas ja miks kaasnevad kognitiivsete funktsioonidega teadvuselamused. Vaimu mõistmises on progress toimunud peaaegu ainuüksi käitumise seletamise pinnal ega ole üldse puudutanud teadvuselamuse küsimust.

Soovi korral võime näha psühholoogiliste ja fenomenaalsete nähtuste eristamises mitte niivõrd keha–vaimu probleemi lõhestamist, vaid selle lahutamist kaheks eraldi osaks​[[26]](#footnote-27) Keha–vaimu probleemi kõige raskemaks osaks on küsimus, kuidas saab füüsilisest süsteemist kasvada välja teadvuselamus. Keha ja teadvuselamuse vahelise ühenduslüli võib lahutada kaheks osaks: ühenduslüliks füüsiliste ja psühholoogiliste nähtuste vahel ning ühenduslüliks psühholoogiliste ja fenomenaalsete nähtuste vahel. Nagu ülal nägime, on meil nüüd üsna hea ettekujutus sellest, kuidas füüsikalisel süsteemil võivad olla psühholoogilised omadused, ja seega on psühholoogiline keha–vaimu probleem lahendatud. Jääb veel küsimus, miks ja kuidas kaasnevad nende psühholoogiliste omadustega fenomenaalsed omadused — miks kaasneb näiteks kõigi valuga seonduvate stiimulite ja reaktsioonidega valuelamus. Jackendoffi (1987) eeskujul võime nimetada seda ülejääki *vaimu–vaimu probleemiks*​. Praegused füüsikalised seletused ei vii meid psühholoogilisest vaimust kaugemale. Endiselt on selgusetuks jäänud ühenduslüli psühholoogilise vaimu ja fenomenaalse vaimu vahel.[[27]](#footnote-28)

Pole võimatu, et ühenduslüli fenomenaalsete ja füüsikaliste nähtuste vahel on sõltumatu ühenduslülist psühholoogiliste ja füüsikaliste nähtuste vahel ning probleemi säärane lahutus seega võimatu, kuid see ei tundu kuigi tõenäolisena. Fenomenaalsete ja psühholoogiliste omaduste ülaltäheldatud selge korrelatsioon viib mõttele, et nende vahel on mingi sügav seos. Järgmistes peatükkides püüan näidata, et see seos on äärmiselt tugev nig et küsimuse lahutamine osadeks on keha ja vaimu probleemi käsitlemisel kasulik strateegia. Ja see tähendab, et teadvuselamuse mõistmiseks on oluline mõista seost psühholoogiliste ja fenomenaalsete nähtuste vahel.

### 5. Kaks teadvuse mõistet

Kuna nii paljudel vaimu puudutavatel terminitel on kahetine iseloom, siis pole üllatav, et isegi sõnal “teadvus” on nii fenomenaalne kui ka psühholoogiline ​**[lk.26]**​ tähendus. Siiani olen ma kõnelnud teadvusest selle sõna fenomenaalses mõttes, mis hõlmab vaimu kõiki ülalmainitud fenomenaalseid aspekte. Selles mõttes tähendab teadvus lihtsalt mingite fenomenaalsete kvaliteetide realiseerumist. See ongi sõna “teadvus” põhiline tähendus, või igatahes tähendus, mille seletamine valmistab kõige suuremaid raskusi. Kuid see ei ole selle termini ainus tähendus. Teadvuseks nimetatakse ka mitmesuguseid psühholoogilisi omadusi, näiteks raporteerimisvõimet või informatsiooni introspektiivset kättesaadavust. Seesugused psühholoogilised omadused võib paigutada *psühholoogilise teadvuse* rubriiki, erinevalt *fenomenaalsest teadvusest*​, mida ma peamiselt käsitlenud olen.

See kahemõttelisus võib teadvuse üle arutlemisel palju segadust põhjustada. Sageli juhtub, et mõni autor pakub teadvusele välja uue seletuse, püstitades alustuseks fenomenaalse teadvuse probleemi kogu oma raskuses, kuid viimaks annab seletuse vaid mõnele psühholoogilise teadvuse aspektile, näiteks introspektsioonivõimele. Tema seletusel võib iseenesest mingi väärtuski olla, kuid lugejale jääb tunne, et autor ei ole oma lubadust lõpuni täitnud.

#### Psühholoogilise teadvuse erikujud[[28]](#footnote-29)

Psühholoogilisi mõisteid, mida võidakse tähistada terminiga “teadvus”, on terve hulk. Nende hulka kuuluvad näiteks järgmised.

.

<<PARA.28-3>​>​*Ärkvelolek.* Mõnikord ütleme, et inimene on teadvusel, ja mõtleme selle all, et ta parajasti ei maga. On mõistlik oletada, et meil on elamusi ka magades, ja seega on selge, et see tähendus ei ühti fenomenaalse teadvuse mõistega. On vägagi tõenäoline, et ärkvelolekut on võimalik funktsionaalselt analüüsida, esimeses lähenduses vahest võimena töödelda informatsiooni maailma kohta ja seda ratsionaalselt kasutada.

<<PARA.28-4>​>​*Introspektsioon.* See on protsess, mille varal me võime saada teadlikuks oma seesmiste seisundite sisust. Kui te küsite minult minu vaimuseisundite kohta, leian ma vastuse introspektsiooni abiga. Ligipääsul oma vaimsetele seisunditele on tavapärases teadvusemõistes oluline koht, ja see on vähemasti osaliselt funktsionaalne mõiste. Seda võiks analüüsida kui subjekti ratsionaalsete protsesside suutlikkust kohasel viisil käidelda informatsiooni oma seesmiste seisundite kohta ning tema võimena seda informatsiooni kohasel viisil kasutada.

<<PARA.28-5>​>​*Raporteerimisvõime.* See on meie suutlikkus anda teistele teateid oma psüühiliste seisundite sisust. Raporteerimisvõime eeldab introspektsioonivõimet, kuid on kitsam mõiste, kuna eeldab keeleoskust. Selline teadvusemõiste on sageli olnud keskseks huviobjektiks filosoofia ja psühholoogia operatsionalistliku suuna esindajatele.

​**[lk.27]**​ *Eneseteadvus.* See termin tähistab meie võimet iseendast mõtelda, meie teadlikkust enda olust omaette indiviidina, kes erineb teistest indiviididest. Minu eneseteadvust võiks ehk analüüsida kui minu ligipääsu oma minamudelile või kui representatsiooni, mis mingil viisil seostub minuga. Võib vabalt olla, et eneseteadvus on olemas ainult inimesel ja mõnel üksikul loomaliigil.

​*Tähelepanu.* Me ütleme sageli, et keegi on millestki teadlik [​*conscious of*​] parajasti sel ajal, kui ta sellele tähelepanu pöörab — see tähendab, kui tunduv osa tema kognitiivsetest ressurssidest tegeleb vastava informatsiooniga. Miski võib olla meie fenomenaalses teadvuses ka siis, kui me seda tähele ei pane. Nii on see näiteks perifeerse nägemise puhul.

​*Käitumise tahteline juhtimine.* Me nimetame käitumisakti teadvuslikuks, kui ta on sooritatud tahtlikult — teisiti öeldes, kui teole on eelnenud mõte, mille mingil elemendil on teoga sobivat liiki põhjusik seos.

​*Teadmine.* Kui me igapäevases keeles ütleme, et keegi on mingist faktist teadlik [​*conscious*​], siis tähendab see, et ta teab seda fakti, ja öelda, et ta on millestki teadlik, tähendab täpselt sama mis öelda, et ta teab seda. Teaduslikes arutlustes teadvuse kohta kerkib see sõnatarvitus harva esile, kuid igapäevases uususes ei ole see tähendus arvatavasti harvem kui kõik teised.

.

Need kõik on suuresti funktsionaalsed mõisted, nagu võib aru saada, kui hakata mõtlema, kuidas me võiksime püüda seletada fenomene, mida nad tähistavad. Näiteks tähelepanu võiks proovida seletada kognitiivsete protsessidega, mis koondavad ressursse mingile ühele olemasoleva informatsiooni aspektile. Introspektsiooni seletamiseks võiks püüda seletada protsesse, mis teevad subjekti asjaomasel viisil tundlikuks oma seesmistele seisunditele. Analoogiline on lugu ka teiste omaduste seletamisega: kõigil juhtudel näib asja olemust tabavat just funktsionaalne seletusviis.

Ent kuigi neil mõistetel on psühholoogiline tuum, kaasnevad paljude või kõigiga neist fenomenaalsed seisundid. Näiteks on olemas kindlat liiki fenomenaalne seisund, mis kaasneb eneseteadvusega. Mingi fenomenaalne seisund kaasneb ka introspektsiooniga, tähelepanuga ja käitumise tahtelise juhtimisega. Mõnikord peetaksegi sõnadega “introspektsioon” ja “eneseteadvus”, samuti teiste ülalkäsitletud kaheaspektiliste terminitega, silmas fenomenaalseid seisundeid. See võib põhjustada segadust. Mõnikord väidetaksegi, et selleks, et mõnd protsessi saaks nimetada tõeliseks introspektsiooniks, tähelepanuks jne., peab tal olema fenomenaalne aspekt. Nagu ennegi, käib ka siin vaidlus suuresti sõnade üle. On selge, et igaühega neist mõistetest seostub fenomenaalseid kui ka psühholoogilisi omadusi. Kes ei taha tähistada psühholoogilist omadust nii auväärse terminiga nagu “tähelepanu”, võib selle asemel öelda “pseudotähelepanu”. Ent kuidas me neid omadusi ka ei nimetaks, põhimõttelistes filosoofilistes probleemides ei muutu sellest midagi.

​**[lk.28]**​ Psüühilistes protsessides, millele need tähendused viitavad, figureerivad tavaliselt nii fenomenaalsed kui ka psühholoogilised omadused, kuid nagu teistegi vaimsete mõistete puhul, nii ka siin ei tohiks neid kaht omavahel segi ajada. Samuti tuleb olla ettevaatlik, et mitte ajada nende sõnade fenomenaalseid tähendusi segi fenomenaalse teadvuse endaga.

#### Teadvus ja teadlikkus

Nägime, et emotsioonielamustega, eneseteadvuselamusega, aistinguelamustega jne. kaasnevad alati mingid psühholoogilised omadused. On loomulik oletada, et mingi psühholoogiline omadus kaasneb ka elamuse, fenomenaalse teadvuse endaga. Ma arvan, et niisugune omadus on tõepoolest olemas; seda võiks nimetada *teadlikkuseks*​. See on psühholoogilise teadvuse kõige üldisem liik.

​*Teadlikkust* võib jämedates joontes määratleda kui seisundit, kus meil on ligipääs mingile informatsioonile ja võimalus kasutada seda informatsiooni oma käitumise juhtimiseks. Me võime olla teadlikud näiteks mingist objektist oma ümbruses, või oma keha või vaimu seisundist. Teadlikkusega mingist informatsioonist kaasneb üldiselt võime selle informatsiooni põhjal oma käitumist suunata. See on selgelt funktsionaalne mõiste. Igapäevases keelepruugis on “teadlikkus” sageli sünonüümne “teadvusega”, kuid mina reserveerin sõna “teadlikkus” ainuüksi äsjakirjeldatud funktsionaalse mõiste tähistamiseks.

Üldiselt tundub, et kõikjal, kus on fenomenaalne teadvus, on ka teadlikkus. Minu fenomenaalse elamusega kollasest raamatust minu kõrval kaasneb minu funktsionaalne teadlikkus raamatust ning minu teadlikkus kollasest värvusest. Minu valuelamusega kaasneb minu teadlikkus millestki ebameeldivast, mis mõjub mulle tavaliselt nii, et ma seda edaspidi võimaluse korral väldin. Iga teadvuselamusega kaasneb teadlikkus — seda tõestab fakt, et teadvuselamused on *raporteeritavad*​: kui mul on mingi elamus, siis võin ma kõnelda faktist, et mul see elamus on. Ma ei pruugi elamusele tähelepanu pöörata, kuid soovi korral olen ma võimeline teda tähele panema ja temast kõnelema. Raporteerimisvõime olemasolust järeldub otseselt, et ma olen ülalöeldud mõttes teadlik. Loomulikult võib loomadel ja ka inimestel, senikaua kui nad veel kõnelda ei oska, olla teadvuselamusi ka ilma raporteerimisvõimeta, ent usutavasti on ka nendel olenditel teadlikkus mingil määral siiski olemas. Teadlikkus ei eelda raporteerimisvõimet, kuigi kõnevõimelistel olenditel käivad need kaks enamasti käsikäes.

Teadvusega kaasneb alati teadlikkus, kuid teadlikkusega äsjaöeldud mõttes ei pea tingimata kaasnema teadvus. Näiteks võib inimene olla mingist faktist teadlik ka siis, kui sellega ei kaasne mingit konkreetset fenomenaalset elamust. Kuid teadvuse mõistet saab arvatavasti kitsendada nii, et ta osutub oma mahult kattuvaks või peaaegu kattuvaks ​**[lk.29]**​ fenomenaalse teadvuse mõistega. Ma ei hakka seda ideed siinkohal edasi arendama, kuid võtan ta lähemalt arutluse alla [6. peatükis](#ORG.6.3).

Teadlikkuse mõiste hõlmab sõna “psühholoogiline teadvus” kõiki või peaaegu kõiki ülalloetletud psühholoogilisi teadvusemõisteid. Introspektsiooni võib määratleda kui teadlikkust mingist seesmisest seisundist. Tähelepanu võib määratleda kui eriti selget teadlikkust mingist objektist või sündmusest. Eneseteadvust võib mõista kui teadlikkust iseendast. Käitumise tahtelise juhtimisega pole asi nii lihtne, kuigi seda võib osaliselt analüüsida kui tähelepanu, mida me pöörame oma käitumisele. Ärkvelolekut võib laias laastus iseloomustada kui seisundit, milles inimene on võimeline oma keskkonda mingil määral ratsionaalselt käsitama, ja seega eeldab ka ärkvelolek mingitliiki teadlikkust.

Idee funktsionaalsest teadvusemõistest, mida saab määratleda ligipääsu kaudu, on üksikasjalikult välja töötanud Block (1995), kes arutleb erinevuse üle *fenomenaalse teadvuse* ja *ligipääsuteadvuse* vahel. Blocki arusaam ligipääsuteadvusest on küllalt täpses vastavuses teadlikkuse mõistega, nagu mina seda siin kirjeldasin. (Sellest vastavusest tuleb lähemalt juttu [6. peatükis](#PARA.125-11).) Üsna sarnasel viisil teeb ka Newell (1992) selget vahet “teadlikkusel” ja “teadvusel”. Teadlikkuse kohta ütleb ta, et see on “subjekti võime seada oma käitumine sõltuvusse mingist teadmisest” ning kirjeldab seejärel üksikasjalikult, mille poolest erineb niiviisi mõistetud teadlikkus teadvusest, mis on tema sõnul mittefunktsionaalne fenomen. Sarnastest eristustest on kirjutanud ka teised filosoofid ja kognitiivteadlased.[[29]](#footnote-30) [[30]](#footnote-31) [[31]](#footnote-32)

#### Teadvuse seletamine ja teadlikkuse seletamine

Nagu psühholoogilised omadused üldse, nii ka teadlikkus ei sea meie ette kuigi palju metafüüsilisi probleeme. Probleemid, mis kerkivad seoses teadvuse psühholoogiliste erikujudega, on sama raskusjärku nagu need, mida esitavad meile mälu, õppimine ja uskumused. Muidugi ei ole ka teadlikkuse mõiste kristallselge ja võib seega nõuda veel küllalt tõsist filosoofilist analüüsimist. Tohutult palju on teha ka kognitiivteaduses, mis uurib, kuidas looduslikud ja tehislikud kognitiivsed süsteemid saaksid funktsioneerida teadlikult. Kuid nende uurimisprogrammide sihiseaded on üpris selged. Pole erilist põhjust kahelda, et kognitiivteaduse normaalne areng need probleemid kohase filosoofilise analüüsi toel pikapeale lahendab.

Kuivõrd just teadvus on vaimuteadusele tõeliselt raske probleem, on meie arutlustes kesksel kohal fenomenaalne teadvus. Siin on probleemide raskusjärk hoopis teine. Isegi kui meil oleks teadvusega süsteemide talitluse füüsikaline ja informaatiline külg juba seletatud, oleks ikka veel seletamata, miks on süsteemil teadvuselamused. Mõned autorid vaidlevad sellele väitele muidugi vastu, ja nende argumentidest teen ma allpool pikemalt juttu. Siinkohal võib aga piirduda tähelepanu juhtimisega *prima facie* erinevustele probleemide vahel, mis kerkivad seoses ühelt poolt teadvuse fenomenaalsete ja ​**[lk.30]**​ teiselt poolt tema psühholoogiliste erikujudega. Just fenomenaalse teadvuse probleem on see, mis meile muret teeb.

Paraku heidetakse fenomenaalne ja psühholoogiline teadvusemõiste kõigist oma erinevustest hoolimata kirjanduses sageli ühte patta. Igapäevakeeles ei ole see segiajamine kuigi oluline, sest need teadlikkus ja fenomenaalne teadvus käivadki tavaliselt käsikäes. Ent kui hakata neid seletama, siis on nende kontseptuaalsel erinevusel otsustav tähtsus. Niipalju kui “teadvusele” on kuigivõrdki rahuldavaid seletusi pakutud, seletavad need tavaliselt vaid psühholoogilist aspekti. Fenomenaalne aspekt jääb üldjuhul täielikult vaatluse alt välja.

Paljud viimaseaegsed teadvuse filosoofilised analüüsid on käsitlenud peamiselt mittefenomenaalseid aspekte. Rosenthal (1996) kinnitab, et vaimne seisund on teadvuslik parajasti siis, kui selle seisundi kohta on olemas kõrgemat järku mõte. See analüüs võib ehk aidata seletada introspektiivset teadvust ja võib-olla ka teisi teadlikkuse aspekte, kuid tundub, et ta ei seleta fenomenaalset elamust.[[32]](#footnote-33) Analoogiliselt tegeleb ka suur osa Dennetti raamatust (1991) üksikasjaliku kognitiivse mudeli visandamisega, mida ta pakub välja kui teadvuse seletust. Tundub aga, et see mudel kirjeldab peamiselt subjekti võimet oma mentaalsest seisundist verbaalselt raporteerida. Seega võiks Dennetti mudel anda seletuse raporteerimisvõimele, introspektiivsele teadvusele ja võib-olla veel mõnele teadlikkuse aspektile, kuid mudelis pole midagi, mis seletaks fenomenaalset teadvust. (Dennett sellise otsusega muidugi ei nõustuks.)

Armstrong (1968), kelle jaoks teadvus on takistus tema funktsionalistliku vaimuteooria väljatöötamise teel, vaatleb selle mõiste analüüsimiseks teatavat hüpoteetilist eneseskaneeerimismehhanismi. Võib-olla aitab tema teooria mõista eneseteadvust ja introspektiivset teadvust, kuid fenomenaalse elamuse probleemi jätab see täiesti kõrvale. Armstrong (1981) kõneleb nii pertseptuaalsest kui ka introspektiivsest teadvusest, ent need huvitavad teda vaid kui teadlikkuse erikujud; probleemide üle, mis kerkivad seoses elamuse fenomenaalsete kvaliteetidega, ta ei arutle. Seega hiilib tema funktsionalistlik teooria teadvuse mõiste ähmasust ära kasutades kõrvale just sellest, mis on teadvuses tõeliselt problemaatiline.

Teised autorid on “teadvuse” teemal kirjutades mõelnud selle all peamiselt eneseteadvust või introspektiivset teadvust. Van Gulick (1988), kes ütleb, et teadvust tuleks määratleda kui “refleksiivset metapsühholoogilist informatsiooni”, on parimal juhul andnud nende psühholoogiliste mõistete analüüsi ja möönab tegelikult ka ise, et säärane analüüs ei tarvitse puudutada fenomenaalseid aspekte. Ka Jaynesi (1976) põhjalik teadvuseteooria käsitleb ainult meie teadlikkust omaenda mõtetest. Ta ei ütle midagi tajumisega kaasnevate fenomenide kohta, mistõttu tema teooriat ei saa pidada isegi üldiseks teadlikkuse teooriaks, ammugi mitte fenomenaalse teadvuse teooriaks. Hofstadteril (1979) on ​**[lk.31]**​ teadvuse kohta öelda mõndagi huvitavat, ent introspektsioon, vaba tahe ja minatunne huvitavad teda rohkem kui elamus *per se*​.

Niipalju kui psühholoogid teadvuse üle üldse arutlenud on, ei ole nad enamasti teinud kuigi hoolikalt vahet fenomenaalsete ja psühholoogiliste mõistete vahel. Tavaliselt käsitlevad psühholoogilised uurimistööd mõnd üksikut teadlikkuse aspekti, näiteks introspektsiooni, tähelepanu või eneseteadvust. Isegi teadvuse psühholoogilistel aspektidel on, vähemalt viimase ajani, olnud psühholoogias otsekui halb maine. Võib-olla tuleb see nende mõistete mõningasest ebaselgusest ja raskustest, mis kerkivad seoses sääraste kõrgtaseme fenomenidega nagu introspektsioon. Võib oletada, et suuremalt jaolt tuleb see halb maine sellest, et teadvuseks nimetatakse ka fenomenaalset teadvust, mistõttu teadvuse uurijad tunnevad endid otsekui kuriteo kaasosalistena.

Mõnikord võib kuulda, et psühholoogiline uurimistöö olevat viimastel aastatel hakanud tulema tagasi teadvuse juurde. Tegelikult tundub asi olevat selles, et teadlased on hakanud aktiivselt uurima teadvuse psühholoogilisi aspekte ning et seejuures ei ole uuritavate fenomenide kohta kardetud kasutada sõna “teadvus”. Fenomenaalne teadvus jääb aga enamasti kõrvale. Sellest võib ehk arugi saada. On ju lihtne mõista, kuidas eksperimentaalpsühholoogia meetodid võivad kaasa aidata mitmesugust liiki teadlikkuse mõistmiseni, kuid pole kerge näha, kuidas nad võiksid seletada fenomenaalset elamust[[33]](#footnote-34) .

Kognitiivsed mudelid sobivad hästi teadvuse psühholoogiliste aspektide seletamiseks. Ettekujutus füüsilisest süsteemist, mis on võimeline jälgima oma seesmisi seisundeid, ratsionaalselt käitlema keskkonnast saabuvat informatsiooni või koondama tähelepanu kord ühele, kord teisele objektile, ei too kaasa suuri metafüüsilisi probleeme. On kaunis selge, et funktsionaalse teooriani, mis suudaks neid võimeid seletada, on võimalik jõuda, kuigi selleks võib kuluda kümneid või sadu aastaid. Tõeliselt raske probleem on aga fenomenaalse teadvuse probleem, mida senised psühholoogilise teadvuse seletused ei puuduta.

Siitpeale mõtlen ma sõnaga “teadvus” jälle ainuüksi fenomenaalset teadvust. Kui mul on tarvis kasutada psühholoogilisi mõisteid, kõnelen ma “psühholoogilisest teadvusest” või “teadlikkusest”. Ja peamiselt huvitab mind just fenomenaalne teadvus.

## 2. peatükk. Superventsus ja seletamine

Missugune on teadvuse koht looduse süsteemis? Kas teadvus on füüsiline nähtus? Kas teadvust saab seletada füüsika mõistetes? Nende küsimuste kallale asumiseks tuleb luua mingi mudel; käesolevas peatükis konstrueeringi ma säärase mudeli. Selle keskseks elemendiks on *superventsuse* mõiste; allpool ma defineerin selle ja rakendan teda, selgitamaks reduktiivse seletamise ideed. Selle mõiste abiga visandan ma enamiku kõrgtaseme fenomenide seosest füüsikaliste faktidega pildi, mis hõlmab nähtavasti kõike — võib-olla peale teadvuselamuse.

### 1. Superventsus

Laialt levinud ettekujutuse kohaselt on meie universumi kõige fundamentaalsemateks faktideks füüsikalised faktid, ja kõik teised faktid sõltuvad nendest. Kui mõista sõna “sõltuvad” küllalt nõrgas mõttes, on see vist peaaegu triviaalne tõde; kui aga mõista seda tugevas mõttes, on see väide vaieldav. Kõrgtaseme faktide ja madaltaseme faktide vahel on üldse väga mitmesuguseid sõltuvussuhteid, ja sõltuvussuhete tüüp, mis kehtib ühes valdkonnas, näiteks bioloogias, ei pruugi kehtida mujal, näiteks teadvuselamuste vallas. Filosoofiline arusaam superventsusest annab meile mudeli, mille raames saab arutleda kõigi nende suhete üle ühetaolisel viisil.

Superventsuse mõiste abil saab formaliseerida intuitiivse idee, et üks faktide kogum võib täielikult determineerida teise faktide kogumi.[[34]](#footnote-35) Näiteks määratlevad maailma füüsikalised faktid nähtavasti ka kõik bioloogilised faktid, selles mõttes et kui fikseerida kõik maailma füüsikalised faktid, ei jää bioloogilistele faktidele enam muutumisruumi. (Kui kõik füüsikalised faktid on fikseeritud, siis on ühtlasi fikseeritud, missugused objektid on elusolendid.) Enam-vähem seda me silmas peamegi, kui ütleme, ​**[lk.33]**​ et bioloogilised omadused on füüsikaliste omaduste suhtes superventsed. Üldkujul tähendab superventsus suhet kahe omaduste kogumi vahel: B-omaduste, mida me peame intuitiivselt kõrgtaseme omadusteks, ja A-omaduste, mis on fundamentaalsemad, madaltasemelised omadused.

Meid huvitavatel juhtudel on A-omadusteks tavaliselt füüsikalised omadused, või täpsemalt öeldes need fundamentaalsed omadused, mida peaks sellisteks lõpuleviidud füüsika. Nende hulka kuuluksid arvatavasti mass, laeng, asukoht ajas ja ruumis, mitmesuguste väljade ajalis-ruumiline jaotus, mitmesuguste jõudude ilmingud, mitmesuguste lainete kuju jne. Nende omaduste täpne olemus ei ole oluline. Kui füüsikas peaks midagi radikaalselt muutuma, võib selguda, et asjaomaste omaduste klass erineb oluliselt sellest, mida ma äsja kirjeldasin, kuid minu argumentides ei muutu sellest midagi. A-omaduste hulka ei kuulu säärased kõrgtaseme omadused nagu mahlakus, nätskus, kaelkirjaksus jms., kuigi ka need on mingis mõttes füüsikalised omadused. Edaspidi pean ma füüsikalistest omadustest kõneldes vaikimisi silmas fundamentaalseid omadusi. Mõnikord nimetan ma neid selguse huvides ka *mikrofüüsikalisteks omadusteks* või *füüsikalisteks madaltaseme omadusteks*​.

Maailma ​*A-faktid* ja *B-faktid* on faktid, mis puudutavad A-omaduste ja B-omaduste realiseerumist ja jaotust.[[35]](#footnote-36) [[36]](#footnote-37) Seega hõlmavad füüsikalised faktid maailma kohta kõiki fakte füüsikaliste omaduste realiseerumise kohta aegruumilises muutkonnas. Samuti on kasulik leppida kokku, et maailma füüsikaliste faktide hulka kuuluvad ka tema füüsika põhiseadused. Mõne kontseptsiooni järgi tulenevad need seadused deterministlikult juba konkreetsete füüsikaliste faktide kogumist, kuid seda ei saa pidada endastmõistetavaks tõeks.

Superventsuse definitsiooni üldine kuju on järgmine:

B-omadused on A-omaduste suhtes *superventsed*​, kui kaks olukorda, mis on identsed oma A-omadustelt, ei saa olla erinevad oma B-omadustelt.

Näiteks bioloogilised omadused on füüsikaliste omaduste suhtes superventsed, kuivõrd kaks võimalikku olukorda, mis on füüsikaliselt identsed, on alati ka bioloogiliselt identsed. (Ma kasutan siinkohal sõna “identne” eristamatuse, mitte numeerilise​[[37]](#footnote-38) identsuse tähenduses. Selles mõttes võivad kaks erinevat lauda olla füüsikaliselt identsed). Äsjases üldises formuleeringus figureerivaid mõisteid täpsustades on võimalik jõuda spetsiifilisemate superventsusmõisteteni. Olenevalt sellest, kas me peame olukordade all silmas indiviide või terveid maailmu, saame vastavalt *lokaalse* ja *globaalse* superventsuse mõiste; olenevalt sellest, kuidas me tõlgendame võimalikkuse mõistet, saame *loogilise* superventsuse ja *loodusliku* superventsuse mõiste ja võib-olla veel teisigi. Järgnevalt asungi neid eristusi selgitama.

#### Lokaalne ja globaalne superventsus

B-omadused on A-omaduste suhtes *lokaalselt superventsed*​, kui iga indiviidi A-omadustest tulenevad deterministlikult sama indiviidi B-omadused. Või teisiti öeldes: kui kahes ​**[lk.34]**​ indiviidis realiseeruvad ühesugused A-omadused, siis realiseeruvad neis alati ka ühesugused B-omadused. Näiteks objekti kuju on tema füüsikaliste omaduste suhtes lokaalselt superventne: kui kahel suvalisel objektil on ühesugused füüsikalised omadused, siis on neil vältimatult ka ühesugune kuju. Seevastu väärtus ei ole füüsikaliste omaduste suhtes lokaalselt superventne: “Mona Lisa” füüsiliselt täpne koopia ei ole nii väärtuslik kui “Mona Lisa” originaal. Füüsikaliste omaduste suhtes lokaalselt superventseteks ei saa üldjuhul pidada objekti omadusi, mis ei sõltu ainult tema füüsilistest omadustest, vaid mingil moel ka tema kontekstist — tema ümbrusest ja ajaloost. “Mona Lisa” originaal on väärtuslikum kui koopia, sest nende ajalooline kontekst on erinev: originaal on Leonardo da Vinci enda maalitud, koopia mitte.[[38]](#footnote-39) [[39]](#footnote-40)

Seevastu B-omaduste globaalne superventsus A-omaduste suhtes tähendab seda, et kogu maailma B-faktid kohta tulenevad üheselt A-faktidest, või teisiti öeldes, ei ole kaht võimalikku maailma, millel oleksid ühesugused A-omadused, kuid erinevad B-omadused.[[40]](#footnote-41) [[41]](#footnote-42) [[42]](#footnote-43) Maailma all tuleb siinkohal mõista kogu universumit; erinevad võimalikud maailmad vastavad erinevatele võimalikele asjaseisudele universumis.

Lokaalne superventsus eeldab globaalset superventsust, kuid mitte ümberpöördult. Näiteks on mõeldav, et bioloogilised omadused on globaalselt superventsed füüsikaliste omaduste suhtes selles mõttes, et igasugune maailm, mis oleks füüsiliselt identne meie maailmaga, oleks temaga ka bioloogiliselt identne. (Siinkohal tuleb teha väike mööndus, mida ma mainin [varsti](#ORG.2.1.3).) Kuid tõenäoliselt ei ole nende superventsus lokaalne. Kahtlemata võivad kaks füüsiliselt identset organismi mingite bioloogiliste omaduste poolest erineda. Näiteks võib ühel olla keskkonnaerinevuste tõttu parem kohastumus kui teisel. Ning on isegi mõeldav, et kaks füüsiliselt identset, kuid erineva evolutsioonilooga organismi võivad kuuluda kahte eri liiki.

Globaalse ja lokaalse superventsuse vaheline erinevus ei loe teadvuselamustest kõneldes kuigi palju, sest kui teadvus on üldse füüsikaliste omaduste suhtes superventne, siis on ta seda arvatavasti lokaalselt. Kui kaks olendit on füüsiliselt identsed, siis ei välista keskkonna ja ajaloolise konteksti erinevused nende elamuste identsust. Loomulikult võib kontekst elamusi kaudselt mõjutada, kuid sellisel juhul avaldab ta mõju seesmisele struktuurile, nagu see toimub tajumise puhul. Hallutsinatsioonid, illusioonid jms. fenomenid annavad tunnistust sellest, et otseselt annab elamusele kuju seesmine struktuur, mitte kontekst.

#### Loogiline ja looduslik superventsus

Meie jaoks olulisem on erinevus *loogilise*​ (ehk ​*kontseptuaalse*​​) ja pelgalt *loodusliku*​ (ehk *noomilise*​ või *empiirilise*​) superventsuse vahel.

​**[lk.35]**​ B-omadused on A-omaduste suhtes *loogiliselt superventsed*​, kui ei ole kaht loogiliselt võimalikku olukorda, mis oleksid identsed A-, kuid erinevad B-omadustelt. Loogilisest võimalikkusest tuleb pikemalt juttu [siinsamas peatükis allpool](#ORG.2.5.0.1). Esialgu võib selle all üldjoontes mõista võimalikkust kõige laiemas mõttes; see on enam-vähem sama mis mõeldavus, ainult et ilma kitsendusteta, mida seavad meie maailma seadused. Kasulik on mõelda loogiliselt võimalikust maailmast kui maailmast, mille loomine oleks (hüpoteetiliselt!) olnud Jumala võimuses, kui see oleks olnud tema tahtmine.[[43]](#footnote-44) Jumal ei oleks saanud luua maailma, kus emarebased oleksid isased, kuid ta oleks võinud luua maailma, kus telefonid lendaksid. Mingi väite loogilise võimalikkuse üle otsustamisel on kitsendused üldjuhul kontseptuaalsed. Isase emarebase mõiste on vastuoluline, seega on isased emarebased loogiliselt võimatud; lendava telefoni mõiste on küll veidi ebaharilik, kuid kontseptuaalseid ebakõlasid selles ei ole, ja seega on lendavad telefonid loogiliselt võimalikud.

Tuleb rõhutada, et me ei defineeri loogilist superventsust tuletatavuse kaudu mingis formaalse loogika süsteemis, vaid loogiliselt võimalike maailmade (ja indiviidide) mõiste kaudu, mis taolistest formaalsetest kaalutlustest ei sõltu. Sellist võimalikkust nimetatakse filosoofilises kirjanduses sageli “mitterangelt [​*broadly*​] loogiliseks” võimalikkuseks, vastupidiselt “rangelt [​*strictly*​] loogilisele” võimalikkusele, mille definitsioon sõltub formaalsetest süsteemidest.[[44]](#footnote-45)

Globaalsel tasemel on bioloogilised omadused füüsikaliste omaduste suhtes loogiliselt superventsed. Isegi Jumal ei oleks saanud luua maailma, mis oleks füüsiliselt identne meie omaga, kuid erineks sellest bioloogiliselt. Bioloogiliste faktide sõltumatuks varieerimiseks lihtsalt puudub loogiline ruum. Niipea kui me oleme fikseerinud kõik maailma füüsikalised faktid — sealhulgas faktid iga viimsegi osakese paiknemise kohta ajas ja ruumis —, on fikseeritud ka kõigi maailmas leiduvate objektide makroskoopiline kuju, nende liikumis- ja talitlemisviis ning kõik nende füüsikalised vastasmõjud. Kui selles maailmas leidub elus känguru, siis leidub igas selle maailmaga identses maailmas temaga füüsiliselt identne känguru, ja see känguru on automaatselt elus.

Me võime ette kujutada, et mingi hüpoteetiline ülivõimekas olend — näiteks Laplace’i deemon, kes teab iga osakese asukohta universumis — suudaks kõiki mikrofüüsikalisi fakte teades nende põhjal otsekohe “välja arvutada” kõik bioloogilised faktid. Mikrofüüsikalistest faktidest piisab talle maailma mikroskoopilise struktuuri ja dünaamika ajalis-ruumilise mudeli konstrueerimiseks, ja selle mudeli abil tuletaks ta otsekohe kogu makroskoopilise struktuuri ja dünaamika. See informatsioon annab talle kõik, mida ta vajab, et otsustada, missugused süsteemid on elusad, missugused kuuluvad samasse liiki jne. Kui ta tunneb bioloogilisi mõisteid ja tal on täpne mikrofüüsikaliste faktide kirjeldus, ei vaja ta enam mingit muud informatsiooni.

​**[lk.36]**​Kui B-omadused on A-omaduste suhtes loogiliselt superventsed, siis võib öelda, et B-omadused *tulenevad*​ A-omadustest. (Me ütleme, et üks fakt *tuleneb* teisest, kui on loogiliselt võimatu, et ta kehtiks, ilma et kehtiks teine fakt.) Säärastel juhtudel saab Laplace’i deemon lugeda A-faktide kirjeldusest välja B-faktid, kui ta valdab selleks tarvilikke B-mõisteid. (Loogilise superventsuse mitmesuguste mõistmisviiside vahelisest seosest kõnelen ma käesolevas peatükis allpool veel palju põhjalikumalt; siinne arutlus on peamiselt illustratiivne.) Mingis mõttes võib öelda, et kui loogiline superventsus kehtib, siis ei ole selleks, et B-faktid oleksid niisugused, nagu nad on, tarvis midagi muud, kui et kõik A-faktid oleksid niisugused, nagu nad on.

Kuid superventsust võib olla ka ilma loogilise superventsuseta. Superventsuse nõrgem variant leiab aset siis, kui kaks omaduste kogumit on looduses omavahel süstemaatilises ja täiuslikus korrelatsioonis. Näiteks sõltub gaasi ühe mooli rõhk süstemaatiliselt tema temperatuurist ja ruumalast kooskõlas seadusega *pV*​ = ​*KT*​, kus *K* on konstant. (Kujutan praegu näitlikkuse mõttes olukorda nii, nagu kõik gaasid oleksid ideaalsed gaasid.) Aktuaalses maailmas määrab gaasi mooli rõhu alati üheselt tema temperatuur ja ruumala: pole empiiriliselt võimalik, et kahel erineval ühemoolilisel gaasikogusel oleks üks ja sama temperatuur ja ruumala, kuid erinev rõhk. Sellest järeldub, et gaasi ühe mooli rõhk on mingis mõttes superventne tema temperatuuri ja ruumala suhtes. (Põhjustel, mis kohe selguvad, on A-omaduste klass, mida ma siinses näites vaatlen, palju kitsam kui kõigi füüsikaliste omaduste klass.) Kuid selline superventsus on nõrgem kui loogiline superventsus. Loogiliselt on võimalik, et ühel moolil gaasil oleks sama temperatuuri ja ruumala juures teistsugune rõhk; selleks kujutlegem näiteks maailma, kus gaasikonstant​ ​*K* on suurem või väiksem. Säärase korrelatsiooni olemasolu on lihtsalt looduse fakt.

See oli näide ühe omaduse looduslikust superventsusest teiste omaduste suhtes: rõhk on looduslikult superventne temperatuuri ja ruumala suhtes ning omaduse suhtes olla üks mool gaasi. Üldiselt on B-omadused A-omaduste suhtes looduslikult superventsed, kui kahes ühesuguste A-omadustega looduslikult võimalikus olukorras on ka B-omadused alati ühesugused.

Looduslikult võimalik olukord on olukord, mis võib reaalselt, ainsatki loodusseadust rikkumata looduses aset leida. See on palju rangem kitsendus kui pelk loogiline võimalikkus. Näiteks olukord, kus gaasikonstant oleks teistsugune, on loogiliselt võimalik, kuid ei saaks kunagi aset leida reaalses maailmas ega ole seega looduslikult võimalik. Kõigis looduslikult võimalikes olukordades on kahes võrdse temperatuuri ja ruumalaga gaasimoolis ühesugune rõhk.

Intuitiivselt vastab looduslik võimalikkus sellele, mida me peame reaalseks *empiiriliseks*​​ võimalikkuseks: looduslikult võimalik olukord on olukord, mis võiks sobivatel tingimustel kujuneda reaalses maailmas. Lisaks aktuaalsetele olukordadele kuuluvad siia ka kontrafaktilised olukorrad, mis oleksid võinud maailma ajaloos aset leida, kui rajatingimused oleksid olnud teistsugused, või mis võivad aset leida tulevikus, olenevalt sellest, mida tulevik toob. Näiteks miilikõrgune pilvelõhkuja ​**[lk.37]**​ oleks peaaegu kindlasti looduslikult võimalik, kuigi aktuaalses maailmas niisugust veel ehitatud ei ole. Looduslikult on isegi võimalik (kuigi äärmiselt ebatõenäoline), et ahv tipiks klaviatuuril “Hamleti” teksti. Võib ka öelda, et looduslikult võimalik olukord on olukord, mis on kooskõlas meie maailma loodusseadustega.[[45]](#footnote-46) Sel põhjusel nimetatakse looduslikku võimalikkust mõnikord ka *noomiliseks*​ või *nomoloogiliseks*​ võimalikkuseks[[46]](#footnote-47) — kreekakeelsest sõnast *nomos*​, mis tähendab seadust.

Olukordi, mis on loogiliselt võimalikud, kuid mitte looduslikult võimalikud, on tohutult palju. Sellesse klassi kuuluvad kõik olukorrad, mis rikuvad meie maailma loodusseadusi: näiteks universum, kus puuduks gravitatsioon või kus funamentalsetel konstantidel oleksid teistsugused väärtused. Selliseid olukordi kohtab sageli ulmekirjanduses, kus kõneldakse antigravitatsiooniseadmetest, igiliikuritest jne. Neid on lihtne ette kujutada, kuid meie maailmas ei hakka neid peaaegu kindlasti mitte kunagi olema.

Teistpidi on iga looduslikult võimalik olukord ka loogiliselt võimalik. Looduslike võimaluste klass on seega loogiliste võimaluste klassi alamklass. Selle erinevuse kohta võib tuua järgmise näite: nii kuupmiil kulda kui ka kuupmiil uraan-235 on vististi loogiliselt võimalikud, kuid looduslikult võimalik on teadaolevalt ainult esimene neist — kuupmiili uraan-235 (stabiilne) eksisteerimine on meie maailmas võimatu.

Looduslik superventsus kehtib siis, kui kõigis ühesuguste A-omaduste jaotustega olukordades on ka B-omaduste jaotus ühesugune — teisiti öeldes, kui mingi olukorra A-faktidest tulenevad loodusliku paratamatusega B-faktid. Nii on see siis, kui ühesuguste A-omaduste kogumitega kaasnevad alati samad B-omadused, ja kui see korrelatsioon ei ole lihtsalt kokkusattumuslik, vaid seaduspärane — see tähendab, kui ühesuguste A-omaduste ilmnedes, ükskõik kus ja ükskõik millal, ilmnevad alati ühesugused B-omadused. (Filosoofilises keeles: sõltuvus peab kehtima ka kontrafaktilistes olukordades.) See koosolu ei pea tingimata paika pidama igas loogiliselt võimalikus olukorras, kuid ta peab pidama paika igas looduslikult võimalikus olukorras.

On selge, et loogiline superventsus eeldab looduslikku superventsust. Kui kahes suvalises loogiliselt võimalikus olukorras, mille A-omadused ühtivad, on ka B-omadused ühesugused, siis on ka kahes suvalises looduslikult võimalikus olukorras B-omadused samad. Kuid, nagu nägime gaasiseaduse näitel, mitte tingimata ümberpöördult: gaasi mooli temperatuur ja ruumala määravad rõhu kõigis looduslikult võimalikes, kuid mitte kõigis loogiliselt võimalikes olukordades, ja seega sõltub rõhk temperatuurist ja ruumalast looduslikult, kuid mitte loogiliselt. Olukordades, kus peab paika looduslik superventsus ilma loogilise superventsuseta, ütlen ma edaspidi, et kehtib *pelgalt looduslik* [​*mere natural*​] superventsus.

Põhjustel, mis [varsti](#ORG.2.1.3) selguvad, on raske leida juhtumeid, kus mingil füüsikaliste omaduste hulgal kehtib looduslik superventsus ilma loogilise superventsuseta; kuid üheks heaks näiteks võib võtta jällegi teadvuse. Tundub olevat väga tõenäoline, et teadvus on füüsikaliste omaduste suhtes looduslikult (lokaalselt või globaalselt) superventne, kuivõrd looduslikus maailmas on kahel füüsiliselt identsel olendil ka kvalitatiivselt identsed elamused. Ei ole aga hoopiski ilmne, et teadvus on füüsikaliste omaduste suhtes ka loogiliselt superventne. ​**[lk.38]**​ Igatahes tundub olend, kes oleks füüsiliselt identne mõne olendiga, kellel on teadvus, kuid kellel teadvuselamused täielikult puuduksid või oleksid teistsugused, paljudele autoritele loogiliselt võimalikuna. (On ka neid, kes ei ole sellega nõus, kuid see on siin praegu ainult näide.) Kui nii, siis on teadvuselamus füüsikaliste omaduste suhtes looduslikult, kuid mitte loogiliselt superventne. Füüsikalise struktuuri ja elamuse vahelise paratamatu seose tagavad ainult loodusseadused, mitte mõni loogiline või kontseptuaalne jõud.

Meie eesmärkideks on väga tähtis teha vahet loogilise ja loodusliku superventsuse vahel.[[47]](#footnote-48) Intuitiivselt võib seda vahet mõista järgmiselt. Kui B-omadused on A-omaduste suhtes loogiliselt superventsed, siis ilmuvad B-faktid automaatselt ja tasuta, niipea kui kui Jumal (hüpoteetiliselt) loob maailma, kus leiavad aset mingid kindlad A-faktid. Kui aga B-omadused on A-omaduste suhtes superventsed puht looduslikult, siis peab Jumal pärast seda, kui ta on kehtestanud A-faktid, tegema veel tööd, et tagada B-faktid: ta peab kehtestama seaduse, mis seostab A-faktid B-faktidega. (Laenasin selle kujundliku seletuse Kripkelt (1972).) Niipea kui säärane seadus on kehtima pandud, tulevad A-faktidega automaatselt kaasa vastavad B-faktid; kuid põhimõtteliselt oleks olnud võimalik, et see nii ei oleks.

Mõnikord võib kuulda ka *metafüüsilisest* superventsusest, mille aluseks ei ole ei loogiline ega looduslik paratamatus, vaid “​*tout court*​ paratamatus” või, nagu mõnikord öeldakse, “metafüüsiline paratamatus”. (Lähtekohaks on siin Kripke (1972) arutlused aposterioorsest paratamatusest). Allpool püüan näidata, et metafüüsiliselt võimalikud maailmad on lihtsalt loogiliselt võimalikud maailmad (ning et metafüüsiline võimalikkus on aposterioorse semantilise varjundiga loogiline võimalikkus), kuid esialgu võib julgesti eeldada, et on olemas metafüüsilise superventsuse mõiste, mida tuleb määratleda analoogiliselt sellega, nagu me äsja määratlesime loogilise ja loodusliku superventsuse. Mõnikord räägitakse ka *nõrgast* superventsusest, kuid see tundub olevat liiga nõrk mõiste, et väljendada ühtki vähegi huvitavat sõltuvussuhet omaduste vahel.[[48]](#footnote-49)

Loogilise ja loodusliku superventsuse erinevus on sõltumatu globaalse ja lokaalse superventsuse erinevusest. Võib kõnelda nii globaalsest loogilisest superventsusest kui ka lokaalsest loogilisest superventsusest, kuigi mind huvitab allpool sagedamini esimene. Kui ma räägin loogilisest superventsusest ilma lisatäpsustusteta, pean ma silmas globaalset loogilist superventsust. Samuti on täiesti võimalik rääkida nii globaalsest kui ka lokaalsest looduslikust superventsusest; kuid looduslikud superventsussuhted, mis meid siin huvitavad, on üldjuhul lokaalsed, või vähemalt saab neid käsitleda lokaalsetena — sel lihtsal põhjusel, et tõenditeks, mis kinnitavad looduslikke superventsussuhteid, on tavaliselt mingite omaduste kimpude vahelised lokaalsed reeglipärased vastavused.[[49]](#footnote-50) [[50]](#footnote-51)

#### Üks loogilise superventsuse raskusi\*

Loogilise superventsuse mõistega seoses tuleb mainida tehnilist raskust, mille allikaks on asjaolu, et loogiliselt võimalik on ka maailm, mis oleks füüsiliselt ​**[lk.39]**​ identne meie maailmaga, kuid sisaldaks peale selle midagi mittefüüsilist, mida meie maailmas ei ole, näiteks ingleid, ektoplasmat või kummitusi. On võimalik mõtelda maailmast, mis oleks täpselt samasugune nagu meie maailm, kuid milles oleks lisaks ka mingi hulk ingleid, kes hõljuvad mingis ektoplasmast koosnevas mittefüüsilises sfääris. Kui need inglid paljuneksid ja evolutsioneeruksid, siis oleksid neil ka omaenda bioloogilised omadused. Neil võiks väga hästi olla kõikvõimalikke uskumusi ja keeruka struktuuriga ühiskond.

Probleem, mille need näited meie ette seavad, on ilmne. Inglitega maailm on meie maailmaga füüsiliselt identne, kuid erineb sellest bioloogiliselt. Kui inglimaailm on loogiliselt võimalik, siis ei ole bioloogilised omadused füüsikaliste omaduste suhtes meie definitsiooni järgi​[[51]](#footnote-52) superventsed. Meie aga tahaksime kindlasti öelda, et bioloogilised omadused on füüsikaliste omaduste suhtes superventsed — kui mitte inglimaailmas (eeldades, et aktuaalses maailmas ingleid ei ole!), siis vähemasti meie maailmas. Intuitiivselt tundub ebasoovitav, et juba ainuüksi inglimaailma pelk loogiline võimalikkus takistaks bioloogiliste omaduste deterministlikku tulenemist füüsikalistest omadustest meie maailmas.

Selle raskuse tõttu leiavad mõned autorid (näiteks Haugeland 1982, Petrie 1987), et loogiline võimalikkus ja paratamatus on liiga tugevad, et neid võiks pidada superventsussuhete määratlemisel sobivaks võimalikkuse ja paratamatuse liigiks, ning et selle asemel tuleks töötada nende nõrgemate vormidega, võib-olla näiteks loodusliku võimalikkuse ja paratamatusega. Kuid nii kaotaks mõtte ülalkirjeldatud väga kasulik vahetegemine loogilise ja loodusliku superventsuse vahel ning tähelepanuta jääks ka asjaolu, et mingis väga reaalses mõttes tulenevad meie maailma bioloogilised faktid loogiliselt füüsikalistest faktidest. Teised autorid (näiteks Teller (1989)) on teinud südame kõvaks ning otsustanud, et mittefüüsilise lisaainesega maailmad ei ole ei loogiliselt ega metafüüsiliselt võimalikud, kuigi pealtnäha tundub vastupidi; kuid sellisel juhul hakkavad loogiline ja metafüüsiline võimalikkus tunduma kaunis meelevaldsetena. Õnneks ei ole sääraste sammude järele vajadust. Selgub, et seda kasulikku mõistet — loogilist superventsust, mis on ühitatav ülalkirjeldatud maailmade võimalikkusega — on täiesti võimalik alal hoida, kui tema definitsiooni sobival viisil täpsustada.[[52]](#footnote-53) [[53]](#footnote-54)

Lahenduse võti on võtta superventsust kui teesi ainult meie maailma kohta (või üldisemalt, mingite kindlate maailmade kohta). See on kooskõlas intuitsiooniga, mis ütleb, et bioloogilised faktid determineeruvad loogiliselt meie maailma füüsikalistest faktidest, olenemata mingite veidrate maailmade olemasolust, kus säärast determinatsiooni ei ole. Täpsustatud definitsioon ütleb, et B-omadused on A-omaduste suhtes loogiliselt superventsed, kui B-omadused tulenevad meie maailmas loogiliselt A-omadustest selles mõttes, et igas võimalikus samade A-faktidega maailmas oleksid ka B-faktid samad.[[54]](#footnote-55) Nii ei satu võimalike maailmade olemasolu, milledes kehtiksid veel mingid muud B-faktid, vastuollu meie maailma loogilise superventsusega, juhul kui vähemalt need B-faktid, mis kehtivad meie maailmas, peavad paika kõigis temaga füüsiliselt identsetes maailmades. Nii see üldjuhul ongi (ühe erandiga, millest tuleb juttu allpool​[[55]](#footnote-56)). Kui meie maailmas on koaala, kes istub eukalüpti otsas ja sööb puulehti, siis leidub igas meie maailmaga füüsiliselt identses maailmas, olgu selles ingleid või mitte, temaga identne koaala, kes istub samuti eukalüpti otsas ja sööb puulehti.

​**[lk.40]**​ Siin on üks väike komplikatsioon. Meie maailma bioloogiliste faktide seas on niisuguseid, mis inglimaailmas ei kehti: näiteks ei ole meie maailmas elusat ektoplasmat, ja kõigi elusolendite aluseks on DNA. Pole võimatu seegi, et inglimaailma ektoplasma on põhjuslikus sõltuvuses füüsikalistest protsessidest, nii et vombatite kopulatsioon füüsilises sfääris võib põhjustada ektoplasmaatilise vombatipoja sündimise mittefüüsilises sfääris. Kui nii, siis võib meie maailmas leiduda (mingis mõttes)​[[56]](#footnote-57) järglasteta vombat, kelle vaste inglitega maailmas ei ole järglasteta. Sellest järeldub, et omadus olla järglasteta ei ole meie definitsiooni järgi superventne, nagu seda ei ole ka maailma mastaapi omadused, näiteks elusa ektoplasma puudumine. Mitte kõik meie maailma faktid ei tulene ainuüksi füüsikalistest faktidest.

Probleemis selgusele jõudmiseks paneme tähele, et kõik need faktid sisaldavad eitavaid eksistentsiaallauseid ja ei olene seega mitte ainult sellest, mis meie maailmas toimub, vaid ka sellest, mida siin ei toimu. Me ei saa eeldada, et need faktid tulenevad deterministlikult mingitest lokaalsetest faktidest, sest nad ei sõltu mitte ainult lokaalsetest sündmustest maailmas, vaid ka maailma kitsendustest. Kõik laused superventsuse kohta peaksid käima ainult positiivsete faktide ja omaduste kohta — niisuguste, mille eitamiseks ei piisa maailma lihtsast laiendamisest. Maailma ​*W*​ *positiivset fakti*​ võib defineerida kui fakti, mis peab paika igas maailmas, mis sisaldab maailma *W* oma osahulgana[[57]](#footnote-58) [[58]](#footnote-59); positiivne omadus on selline, mida, kui ta realiseerub maailmas ​*W*​, realiseerivad ka vastavad indiviidid kõigis maailmades, mis sisaldavad maailma ​*W* oma osahulgana.[[59]](#footnote-60) Enamik argielu fakte ja omadusi on positiivsed — mõelgem näiteks omadusele olla känguru, omadusele olla kahe meetri pikkune või omadusele olla järglaste vanem. Igas negatiivses faktis ja omaduses peitub alati ühel või teisel kujul mingi negatiivne eksistentsiaallause. Siia kuuluvad otseselt negatiivsed eksistentsiaalfaktid nagu näiteks see, et ektoplasmat ei ole olemas; universaalfaktid, näiteks et kõigi elusolendite aluseks on DNA; negatiivsed relatsioonilised omadused, näiteks omadus olla järglasteta; ja superlatiivsed omadused, näiteks omadus olla kõigist kõige suurema järglaste arvuga organism.

Kõik superventsussuhted, mida me allpool käsitlema hakkame, piirduvad positiivsete faktide ja omadustega. Öeldes, et bioloogilised omadused on füüsikaliste omaduste suhtes superventsed, peame edaspidi silmas ainult positiivseid bioloogilisi omadusi. Kuna kõik meid huvitavad omadused, näiteks lokaalsed füüsikalised ja fenomenaalsed omadused, on nagunii positiivsed, siis ei ole see kuigi tõsine kitsendus.

Seega võime B-omaduste globaalset loogilist superventsust A-omaduste suhtes defineerida järgmiselt: kõik B-faktid, mis kehtivad meie maailmas, kehtivad ka kõigis meie maailmast A-eristamatutes loogiliselt võimalikes maailmades[[60]](#footnote-61). Me ei tarvitse lisada siia klauslit positiivsuse kohta, kuid peame B-faktide ja -omaduste all edaspidi tavaliselt silmas positiivseid fakte ja omadusi. Analoogiliselt on B-omadused A-omaduste suhtes lokaalselt ja loogiliselt superventsed, kui kõik igas aktuaalses indiviidis​ ​*x* realiseeruvad B-omadused realiseeruvad ka igas temast A-eristamatus loogiliselt võimalikus indiviidis. Lühemalt ja üldisemalt: B-omadused on A-omaduste suhtes loogiliselt superventsed, kui aktuaalseid olukordi puudutavad B-faktid tulenevad A-faktidest, kui mõista olukordade all globaalsel juhul maailmu ja lokaalsel juhul indiviide. See definitsioon võtab arvesse tõsiasja, et väited superventsuse kohta käivad tavaliselt meie maailma kohta, ning jätab otsustava rolli loogilisele paratamatusele.[[61]](#footnote-62)

#### Superventsus ja materialism

Loogiline superventsus ja looduslik superventsus viivad ontoloogias välja üpris erinevate järeldusteni — erinevate vastusteni küsimusele, mis on maailmas olemas. Kui B-omadused on A-omaduste suhtes loogiliselt superventsed, siis tulenevad B-omadused A-omadustest täiesti tasuta. Niipea kui Jumal seadis (hüpoteetiliselt) paika kõik meie maailma füüsikalised faktid, ilmusid bioloogilised faktid automaatselt ja tasuta. B-faktid lihtsalt kirjeldavad veel kord seda, mida A-faktid juba kirjeldasid. Nad võivad olla teistsugused faktid (fakt elevantide kohta ei ole mikrofüüsikaline fakt), kuid nad ei ole uued faktid.

Pelgalt loodusliku superventsuse juhul ei ole ontoloogia nii sirgjooneline. Maailma mitmesuguseid fenomene ühendavad kontingentsed seaduspärased seosed. Kui B-tunnused on meie maailmas A-tunnuste suhtes superventsed vaid looduslikult, siis võiks olla olemas maailm, milles meie A-faktid kehtiksid, B-faktid aga mitte. Nagu [ülal](#PARA.34-16) nägime, tuli Jumalal pärast kõigi A-faktide fikseerimist teha veel tööd, et fikseerida ka B-faktid. B-faktid on midagi enamat kui A-faktid ja nende olemasolu tähendab, et maailma on ilmunud midagi uut.

Kõike seda silmas pidades võime anda täpse definitsiooni laialt levinud *materialismi* (ehk *füsikalismi*​) doktriinile, mille mõte on üldlevinud arvamuse järgi selles, et maailmas on kõik füüsikaline, või et ei ole olemas midagi, mis jääks väljapoole füüsilise maailma piire, või et maailma faktide hulk sisaldub mingis mõttes ammendavalt füüsikaliste faktide hulgas. Meie keelepruugis: materialism on tõene, kui kõik maailma positiivsed faktid on füüsikaliste faktide suhtes globaalselt-loogiliselt superventsed. See sõnastus kajastab intuitiivset arusaama, et kui materialism on tõene, siis olid kõik faktid fikseeritud, niipea kui Jumal fikseeris maailma füüsikalised faktid.

(Või vähemalt niipea kui ta fikseeris kõik positiivsed faktid. Piirdumine positiivsete faktidega on tarvilik selleks, et ekstoplasmaatiliste lisafaktidega maailmad ei takistaks materialismi kehtimist meie maailmas. Negatiivsed eksistentsiaalsed faktid, nagu näiteks “ingleid ei ole olemas”, ei ole füüsika suhtes rangelt võttes loogiliselt superventsed, kuid nende mittesuperventsus on materialismiga üsna täielikult ühitatav. Mingis mõttes pidi Jumal negatiivsete faktide fikseerimiseks tegema veel midagi lisaks füüsikaliste faktide fikseerimisele: ta pidi lõpuks teatama: “Ja see on kõik”. Soovi korral võiksime lisada materialismi definitsiooni superventsusbaasile teist järku fakti “ja see on kõik”; säärasel juhul saaks positiivsete faktidega piirdumist nõudva kitsenduse välja jätta.)

​**[lk.42]**​ Selle definitsiooni kohaselt on materialism tõene, kui kõik meie maailma positiivsed faktid on tuletatavad füüsikalistest faktidest.[[62]](#footnote-63) Teisiti öeldes, materialism on tõene, kui kõik meie maailma positiivsed faktid kehtivad ka igas loogiliselt võimalikus ja meie maailmast füüsikaliselt eristamatus maailmas[[63]](#footnote-64). See on omakorda samaväärne teesiga, et iga meie maailmast füüsikaliselt eristamatu maailm sisaldab meie maailma koopiat oma osamaailmana või pärisosamaailmana; see tundub olevat intuitiivselt õige definitsioon.[[64]](#footnote-65) (Sellega sarnane on *füsikalismi* definitsioon, mille esitas Jackson (1994) ja kus kriteeriumiks on, et iga meie maailma minimaalne füüsiline duplikaat peab olema meie maailma lihtne [​*simpliciter*​] duplikaat.[[65]](#footnote-66))

Need küsimused tulevad palju pikemalt arutusele [4. peatükis](#ORG.4), kus ma esitan ka uusi argumente materialismi siinse definitsiooni kasuks. Kui lugejale ei meeldi, et selles definitsioonis figureerib loogiline võimalikkus, aga mitte *tout court* võimalikkus või “metafüüsiline võimalikkus”, siis võib ta asendada loogilise võimalikkuse metafüüsilisega. Püüan allpool näidata, et tulemus sellest ei muutu.

### 2. Reduktiivsed seletused

Teaduse hoogne areng mõnel viimasel sajandil on andnud meile põhjust uskuda, et maailmas on väga vähe sügavaid saladusi. Tundub, et peaaegu igale looduse fenomenile mikrofüüsikast kõrgemal tasemel on põhimõtteliselt võimalik anda *reduktiivne seletus*​, s.t. seletada teda ainuüksi lihtsamate entiteetide kaudu. Neil juhtudel tuleneb madalamal tasemel kulgevate protsesside rahuldavast kirjeldusest automaatselt ka kõrgema taseme fenomeni seletus.

Heaks illustratsiooniks on bioloogilised fenomenid. Paljunemist saab seletada, kirjeldades geneetilisi mehhanisme ja rakuprotsesse, mille varal organism võib luua uusi organisme. Kohastumist saab seletada, kirjeldades mehhanisme, mis viivad kohaste muutusteni välises talitluses vastuseks keskkonnast saabuvatele stiimulitele. Bioloogilise elu enda seletamiseks tuleb kirjeldada mitmesuguseid mehhanisme, mille varal toimub paljunemine, kohastumine jne. Niipea kui meil on madalama taseme lugu küllaldastes üksikasjades käes, kaob igasugune fundamentaalse salapära tunne: fenomenid, mis vajasid seletust, on selle saanud.

Midagi sarnast võib öelda ka enamiku teiste looduslike fenomenide kohta. Füüsikas jutustatakse soojuse seletamiseks lugu energiast ja molekulide ergastumisest. Astronoomias käsitletakse Kuu faaside seletamiseks üksikasjalikult orbitaalliikumist ja valguse peegeldumist. Geofüüsikas kirjeldatakse maavärinate seletamiseks maa-aluste ainemasside vastasmõjusid. Kognitiivteadustes tuleb meil (näiteks) õppimise seletamiseks esimeses lähenduses  — kui jätta kõrvale kõik raskused, mis tekivad seoses õppimisega kaasnevate elamustega — vaid kirjeldada mitmesuguseid funktsionaalseid mehhanisme, mis töötlevad keskkonnast saabuvaid stiimuleid ​**[lk.43]**​ ning kutsuvad esile vajalikke muutusi käitumises. Nende seletuste paljud üksikasjad käivad meile esialgu veel üle jõu ning osutuvad tõenäoliselt väga keerukateks, ent me teame, et kui madalama tasandi lugu on küllalt selgeks saanud, siis tuleb pikapeale järele ka kõrgema taseme lugu.

Ma ei tahaks reduktiivse seletuse mõistele esialgu veel täpset definitsiooni anda. Mingi ettekujutuse annavad ülaltoodud näited. Kuid ma võin veidi selgitada, mida reduktiivne seletus ei ole. Mingi fenomeni reduktiivseks seletamiseks pole sugugi tingimata tarvis teda redutseerida, vähemalt mitte selle mitmemõttelise sõna igas tähenduses. Fenomenid, mis võivad realiseeruda paljudel erinevatel füüsilistel kandjatel, näiteks õppimine, ei tarvitse redutseeritavad ollagi, selles mõttes et õppimist ei tarvitse olla võimalik samastada ühegi konkreetse madalama taseme fenomeniga. Kuid see mitmene realiseeritavus [​*multiple realizability*​] ei takista ühegi konkreetse õppimistoimingu reduktiivset seletamist madalama taseme fenomenide kaudu.[[66]](#footnote-67) Ka tuleb meeles pidada, et fenomeni reduktiivne seletamine ja kõrgtaseme teooria redutseerimine on kaks ise asja. Mõnikord toob mingi fenomeni reduktiivne seletamine kaasa ka mõne senise kõrgtaseme teooria reduktsiooni, mõnel teisel juhtul aga selgub, et need teooriad olid ekslikud. Paljudel juhtudel aga ei leidugi kõrgtaseme teooriat, mida saaks redutseerida.

Seletamine ei piirdu ega ammendu reduktiivse seletamisega. On ka palju muusuguseid seletusi, millest mõni võib heita mõnele fenomenile rohkem valgust kui reduktiivne seletus. On fenomene — näiteks elu tekkimine —, millele leidub ajalooline seletus; siin annab reduktiivne seletus ainult sünkroonse ülevaate elussüsteemide talitlemisest. On ka mitmesuguseid kõrgtaseme seletusi: näiteks käitumise mitmesuguseid aspekte võidakse seletada uskumuste ja soovide kaudu. Kuigi samasugusele käitumisele saaks põhimõtteliselt anda ka reduktiivse seletuse, on kõrgtaseme seletused sageli arusaadavamad ja selgemad. Ei tule arvata, nagu peaksid reduktiivsed seletused tõrjuma välja kõik muud liiki seletused. Igal on oma koht.

#### Reduktiivne seletamine funktsionaalse analüüsi varal

Mis võimaldab reduktiivselt seletada nii erinevaid fenomene nagu paljunemine, õppimine ja soojus? Kõigil neil juhtudel on otsustav tähtsus vastava fenomeni iseloomustamiseks nõutavate mõistete olemusel. Me saaksime tigedaks, kui keegi vaidleks paljunemise rakuteoreetilisele seletusele vastu sõnadega: “See seletab, kuidas protsessid rakkudes võivad viia esialgse olendiga sarnase keeruka füüsilise süsteemi tekkimiseni, aga jätab seletamata paljunemise” — sest ainult seda paljunemine ju tähendabki. Üldiselt kaasneb fenomeni reduktiivse seletamisega vaikivalt või selgesõnaliselt kõnealuse fenomeni mõningane skemaatiline analüüs. Paljunemise mõistet võib laias laastus analüüsida kui ​**[lk.44]**​ organismi võimet luua mingil kindlal viisil teisi organisme. Sellest järeldub, et kui me oleme ära seletanud protsessid, mille varal mingi organism loob teisi organisme, siis oleme ära seletanud ka selle paljunemisviisi.

See jutt võib tunduda triviaalsena, kuid just säärase analüüsi võimalikkus on aluseks reduktiivsete seletuste võimalikkusele üldse. Ilma sellise analüüsita puuduks meil seletuslik sild, mis viiks madalama taseme füüsikalistelt faktidelt kõnealuse fenomenini. Kui selline analüüs on aga käes, jääb vaid näidata, kuidas nii- või teistsugused madaltasemelised füüsikalised mehhanismid sellesse analüüsi sobituvad, ja seletus ongi käes.

Kõige huvitavamate seletust nõudvate fenomenide, sealhulgas paljunemise ja õppimise puhul on tavaliselt võimalik analüüsida asjaomaseid mõisteid *funktsionaalselt*​, kirjeldades nende mõistete põhituuma mingi funktsiooni või mingite funktsioonide täitmise kaudu (kus sõna “funktsioon” tuleb mõista pigem põhjuslikus kui teleoloogilises mõttes) või võime kaudu neid funktsioone täita. See tähendab, et niipea kui me oleme seletanud, kuidas need funktsioonid talitlevad, on seletuse leidnud ka meid huvitav fenomen. Kui me oleme seletanud, kuidas organism täidab teise organismi loomise funktsiooni, siis oleme ära seletanud ka paljunemise, sest paljunemine tähendabki vaid selle funktsiooni täitmist. Samamoodi on õppimisega. Kui me ütleme mingi organismi kohta, et ta õpib, siis tähendab see üldjoontes, et selle organismi käitumisvõimed kohastuvad vastama keskkonnnast saabuvatele stiimulitele sobival viisil. Niipea kui me oleme seletanud, kuidas organism suudab sellekohaseid funktsioone täita, ongi ära seletatud ka õppimine.

(Halvemal juhul on seletamata jäänud kõik õppimise fenomenaalsed aspektid; need jätan ma siinkohal arusaadavatel põhjustel vaatluse alt välja. Kui õppimise mõistes sisaldub fenomenaalne element, siis võib õppimisprotsessi vastav osa jääda seletuseta, kuid mind huvitavad siin esialgu õppimise psühholoogilised aspektid, mis moodustavad nähtavasti õppimise mõiste tuuma.)

Põhimõtteliselt käib nende funktsioonide seletamine üsna sirgjooneliselt. Kuivõrd sääraste funktsioonide täitmise tulemusi on endid võimalik füüsikaliselt kirjeldada, ja kuivõrd kõigil füüsikalistel sündmustel on füüsikalised põhjused, peaks iga sellise funktsiooni talitlusele leiduma füüsikaline seletus. Tuleb vaid näidata, mil viisil nii- või teistsugused seisundid kutsuvad mingi põhjusliku protsessi varal ja kooskõlas loodusseadustega esile vastavad resultantseisundid. Muidugi ei tarvitse seesuguste füüsikaliste seletuste üksikasjad olla triviaalsed. Tegelikult moodustavadki lõviosa kõigist reduktiivsetest seletustest üksikasjad, analüütiline komponent on sageli triviaalne. Ent niipea kui vajalikud üksikasjad on käes, annab füüsikalise põhjuslikkuse madaltasemeline kirjeldus seletuse, kuidas asjaomased funktsioonid toimivad, ja seega ka seletuse meid huvitavale fenomenile.

Isegi füüsika mõisteid, näiteks soojust, saab tõlgendada funktsionaalselt: soojus on üldjoontes miski, mille toimel metallid paisuvad; mida põhjustab tuli; mis kutsub esile ​**[lk.45]**​ teatavat liiki aistingu; jms. Niipea kui me oleme ära seletanud, kuidas kõik need põhjuslikud suhted toimivad, on meil käes ka soojuse seletus. Soojuse mõiste tuleneb tema põhjuslikust rollist: teda määratlevad nähtused, mis tüüpiliselt põhjustavad soojust, ja nähtused, mida soojus sobivates tingimustes ise tüüpiliselt põhjustab. Niipea kui me oleme empiirilisel teel kindlaks teinud, kuidas toimivad soojuse põhjuslikud rollid, on ära seletatud ka soojuse fenomen ise.

Siin on tehnilisi keerukusi, kuid need ei muuda asja olemust. Näiteks Kripke (1980) on osutanud erinevusele termini, näiteks “soojus”, ja talle vastava põhjusliku rolli kirjelduse vahel: kui eeldada, et soojus realiseerub molekulide liikumise kaudu, siis oleks mõnes kontrafaktilises maailmas võimalik pidada molekulide liikumist soojuseks olenemata sellest, kas need molekulid mängivad vastavat põhjuslikku rolli või mitte. Kuid sellegipoolest ei seisne soojuse seletamine mitte molekulide liikumise, vaid soojuse põhjusliku rolli seletamises. Mõistmaks, et see on tõepoolest nii, paneme tähele, et meie teadmine soojuse võrdväärsusest molekulide liikumisega on aposterioorne: see teadmine on alles soojuse seletamise tagajärg. Aprioorset arusaama soojuse mõistest, mis meil oli enne sellele fenomenile seletuse leidmist, saab väljendada umbes niisuguste sõnadega: “soojus on see, mis mängib aktuaalses maailmas nii- ja niisugust põhjuslikku rolli”. Niipea kui me oleme välja uurinud, kuidas see põhjuslik roll toimib, on meil käes ka fenomeni seletus. Lisaboonusena oleme saanud teada, mida soojus endast üldse kujutab: ta on molekulide liikumine, sest just molekulide liikumine mängib vastavat põhjuslikku rolli aktuaalses maailmas.

Teine väike keerukus seisneb selles, et paljude põhjuslikule rollile toetuvate mõistete puhul ei ole päris selge, kas nad viitavad mingile üht või teist põhjuslikku rolli täitvale seisundile või selle rolli aktuaalsele talitlemisele. Soojuse all võidakse mõelda niihästi põhjuslikult toimivaid molekule kui ka põhjuslikku protsessi (soojenemist) ennast. Analoogiliselt võidakse tajumise all mõelda nii tajumisakti kui ka selle tagajärjel tekkivat seesmist seisundit.[[67]](#footnote-68) See kahemõttelisus ei muuda aga oluliselt midagi. Põhjusliku rolli täitmise seletus annab seletuse soojusele või tajumisele mõlemas tähenduses.

Kolmas komplikatsioon tuleb sellest, et paljusid põhjuslikust rollist tulenevaid mõisteid iseloomustab osalt ka nende mõju meie elamustele. Näiteks soojust on loomulik mõista kui soojusaistingute põhjust. Kas see tähendab, et soojust ei olegi võimalik seletama asuda, senikaua kui meil puudub seletus soojusaistingutele? Loomulikult ei ole meil neile aistingutele (nagu ka elamustele üldse) head seletust, nii et praktikas jääb see fenomeni külg seletuseta. Kui me suudame seletada, kuidas mingitel tingimustel tekib molekulaarliikumine ja kuidas ta põhjustab metallide paisumist ning stimuleerib mingil kindlal viisil meie nahka, siis piisab tähelepanekust, et see liikumine on soojatundega korrelatsioonis. Sellest korrelatsioonist me järeldame, et siin on peaaegu kindlasti põhjuslik seos. Kindlasti ei ole ükski soojuse seletus täielik, kuni me pole näidanud, kuidas see põhjuslik seos toimib. Kuid enamikuks eesmärkideks piisab ka ebatäielikust seletusest. On mõnevõrra paradoksaalne, et me suudame seletada fenomeni kohta peaaegu kõike, ainult mitte ​**[lk.46]**​ tema mõju meie fenomenoloogiale. Kuid praktikas ei ole sellest midagi katki. Hoopis halvem oleks, kui me ei saakski ülejäänud teadusega edasi liikuda, enne kui meil on olemas mingisugune teadvuseteooria.

#### Reduktiivsed seletused kognitiivteaduses

Paradigma, kus reduktiivne seletamine toimub funktsionaalse analüüsi teel, töötab enamikus kognitiivteaduste valdkondades suurepäraselt, vähemasti põhimõtteliselt. Nagu eelmises peatükis nägime​, saab funktsionaalset analüüsi kohaldada enamikule mittefenomenaalsetest vaimsetest mõistetest. Psühholoogilisi seisundeid saab kirjeldada nende põhjusliku rolli kaudu. Nende seisundite seletamiseks tuleb seletada vastavaid põhjuslikke mehhanisme.

Põhimõtteliselt saab seda teha nende seisundite aluseks olevate neurofüsioloogiliste protsesside tasemel. Niipea kui me oleme näidanud, kuidas mingid neurofüsioloogilised seisundid põhjustavad kõnealuste funktsioonide talitlemise, oleme andnud seletuse ka vastavale psühholoogilisele seisundile. Kuid iga kord polegi tarvis laskuda neurofüsioloogia tasemele. Sageli on mõnd vaimu aspekti võimalik seletada, esitades sobiva *kognitiivse mudeli*​, s.t. kirjeldades detailselt vastavate funktsioonide talitlemiseks piisavaid mehhanisme sisaldava süsteemi abstraktset põhjuslikku organisatsiooni, laskumata seejuures seda põhjuslikku organisatsiooni implementeeriva füüsikalis-keemilise kandja tasemele. Sel viisil oleme andnud seletuse, kuidas psühholoogia vastav aspekt on üldse võimalik, näidates, mil viisil võiksid ühed või teised põhjuslikud mehhanismid olla kõnealuste vaimsete protsesside aluseks. Ja kui me soovime seletada mingeid ühe või teise aktuaalse organismi või organismitüübi vaimseid seisundeid (näiteks õppimist inimestel, erinevalt õppimisest üldse), siis tuleb sellele seletusele lisaks näidata, et mudeli põhjuslik organisatsioon kajastab kõnealuse organismi põhjuslikku organisatsiooni.

Õppimise võimalikkuse seletamiseks võime pakkuda välja mudeli, mille mehhanismid põhjustavad mitmesugustele keskkonnast saabuvatele stiimulitele vastavaid muutusi käitumisvõimes — näiteks konnektsionistliku mudeli. Seletamaks õppimist inimesel, tuleb lisaks sellele näidata, et see mudel vastab põhjuslikule organisatsioonile, mis seisab nende funktsioonide taga inimesel. See teine samm on tavaliselt raske: neurofüsioloogiliste teadmiste lünklikkuse tõttu ei oska me säärasele vastavusele otseselt osutada ning peame seepärast otsima kaudseid tõendeid, näiteks kvalitatiivseid sarnasusi reaktsioonimustrites, ajalisi mõõteandmeid jms. See ongi üks kognitiivteaduste praeguse ebaküpsuse põhjusi. Kuid siingi tuleneb sääraste seletuste põhimõtteline võimalikkus, nagu ikka, otseselt psühholoogiliste mõistete funktsionaalsest iseloomust.

Ent kuigi funktsionaalsed seletused töötavad hästi psühholoogiliste seisundite vallas, ei ole tundu neist paraku olevat abi fenomenaalsete seisundite seletamisel. ​**[lk.47]**​ Põhjus on ilmne. Ükskõik, missuguse funktsionaalse seletuse me inimese kognitsioonile ka annaksime, ikka kerkib veel üks küsimus: miks kaasneb seesuguse funktsioneerimisega teadvus? Psühholoogiliste seisunditega säärast uut küsimust ei teki. Kui küsida mingi funktsionaalse õppimismudeli kohta: “Miks kaasneb nende funktsioonidega õppimine?”, siis siin on õige vastus semantiline: “Sellepärast, et õppimine ei tähendagi midagi muud kui neid funktsioone.” Teadvuse mõistet sarnasel viisil analüüsida ei saa. Erinevalt psühholoogilistest seisunditest ei määratle fenomenaalseid seisundeid nende põhjuslikud rollid. Järelikult ei piisa teadvuse seletamiseks ühe või teise põhjusliku rolli seletamisest. Ka siis, kui me oleme mingi funktsiooni talitluse ära seletanud, on asjaolu, et selle funktsiooniga kaasneb (kui ikka kaasneb) teadvus, jäänud endiselt täiesti seletuseta.

Seda mõtet võib sõnastada ka järgmiselt. Kui meil on olemas õppimise toimiv funktsionaalne seletus, siis on süsteem, milles see seletus realiseerub, kuid mis ei õpi, loogiliselt täiesti võimatu (vähemalt kuivõrd õppimine ei nõua teadvust). Ent olenemata sellest, kuidas me kognitsiooni funktsionaalselt seletaksime, tundub olend, kelles see seletus realiseeruks, kuid kellel ei kaasneks selle realiseeringuga teadvust, olevat loogiliselt võimalik. Ta võib olla looduslikult võimatu — kuna on võimalik, et teadvus kasvabki aktuaalses maailmas faktiliselt välja just sellest funktsionaalsest organisatsioonist —, kuid oluline on, et selles ettekujutuses ei ole loogilist vastuolu.

Kui see on tõepoolest loogiliselt võimalik, siis on vaimsete fenomenide iga funktsionaalne ja isegi iga füüsikaline seletus põhimõtteliselt ebatäielik. Või Levine’i (1983) sõnastuses: niisuguste seletuste ja teadvuse enda vahel on *seletuslünk*​ [​*explanatory gap*​]. Isegi kui sobiva funktsionaalse organisatsiooni pinnalt kasvabki praktikas alati välja teadvus, jääb ikkagi vastuseta küsimus, miks see nii on. Allpool arutleme selle punkti üle pikemalt.

Kui see on nii, siis kaasneb iga fenomenaalset elementi sisaldava vaimse mõistega osaline seletuslünk. Näiteks kui teadvuselamused on tarvilikud uskumiseks või õppimiseks, siis ei ole meil võimalik anda uskumisele või õppimisele täielikku reduktiivset seletust. Kuid meil on vähemalt põhjust uskuda, et põhimõtteliselt on võimalik anda reduktiivne seletus nende vaimsete võimete psühholoogilistele aspektidele, mis peaksid moodustama vastavate mõistete tuuma. Tundub, et kui raskused fenomenoloogiaga kõrvale jätta, siis on kognitiivteadusel olemas vahendid, et vaimsete nähtuste seletamisega edukalt toime tulla.

### 3. Loogiline superventsus ja reduktiivsed seletused

Reduktiivsete seletuste epistemoloogial on superventsuse metafüüsikaga täiesti otsene seos. Igale looduslikule fenomenile saab anda reduktiivse seletuse mingite madaltaseme omaduste kaudu parajasti siis, kui ta on nende omaduste suhtes ​**[lk.48]**​ loogiliselt superventne. Ja talle saab anda reduktiivse seletuse füüsikaliste omaduste kaudu — ehk lihtsalt “reduktiivse seletuse” —, kui ta on loogiliselt superventne füüsika suhtes.

Täpsemalt sõnastades: looduslikku fenomeni saab seletada reduktiivselt mingite madalama taseme omaduste kaudu, kui omadus olla selle fenomeni realiseering on nende madaltaseme omaduste suhtes globaalselt-loogiliselt superventne. Fenomen on reduktiivselt seletatav *simpliciter*​, kui omadus olla selle fenomeni realiseering on globaalselt-loogiliselt superventne füüsikaliste omaduste suhtes.

Seda võib, võib-olla mõneti tinglikult, vaadelda kui reduktiivse seletuse mõiste selgitust. See, et meie varasemast arusaamast reduktiivse seletamise kohta tuleneb (globaalne) loogiline superventsus, peaks olema selge eelnenud arutlustest​. Kui omadus olla mingi fenomeni realiseering ei ole mingite madalama taseme omaduste suhtes loogiliselt superventne, siis jätab nende omaduste iga madaltasemeline seletus vastamata küsimuse: miks kaasneb selle fenomeniga see madaltaseme protsess? Reduktiivne seletamine eeldab meid huvitava fenomeni niisugust analüüsi, mille realiseering on tuletatav madaltaseme faktidest. Seega eeldab reduktiivne seletamine loogilist superventsussuhet. Näiteks paljunemist saab seletada reduktiivselt madalama taseme faktide kaudu just sellepärast, et ta on nende faktide suhtes loogiliselt superventne.

See, et loogiline superventsus on reduktiivseks seletatavuseks ka piisav, on veidi vähem ilmne. Kui mingi fenomen​ ​*P* on mingite madalama taseme omaduste suhtes loogiliselt superventne ja meil on olemas selle fenomeni mingi realiseeringuga seostuvate madalama taseme faktide kirjeldus, siis on​ ​*P* realiseering selle kirjelduse loogiline järelm. Seega annab madalama tasandi faktide kirjeldus meile automaatselt fenomeni​ ​*P* seletuse. Kuid sellegipoolest võib säärane seletus tunduda mõnikord mitterahuldavana. Selleks on kaks põhjust. Esiteks võivad madalama taseme faktid osutuda pealtnäha juhuslike üksikasjade tohutuks segadikuks, milles puudub selge seletuslik ühtsus. Nii võib see olla näiteks kõigi molekulaarsete liikumistega, mis on aluseks mingile õppimisprotsessile. Teiseks võivad fenomeni​ ​*P* erinevate ilmingutega kaasneda täiesti erinevad madaltaseme faktistikud, mistõttu konkreetsete näidete seletused ei anna seletust kogu fenomenile kui tüübile.

Üks võimalus on öelda, et loogiline superventsus on reduktiivseks seletamiseks vaid tarvilik, kuid mitte piisav. Sellest piisab minu argumentideks teadvuse kohta [järgmises peatükis](#ORG.3). Kuid kasulikum on panna tähele, et siiski on olemas ka niisugune praktiline arusaam reduktiivsest seletamisest, mille puhul loogiline superventsus on niihästi tarvilik kui ka piisav. Selle asemel et järeldada ülalmainitud raskustest, et kõnealused kirjeldused ei ole seletused, võib võtta neid kui näiteid selle kohta, et reduktiivne seletus ei tarvitse tingimata olla selgusttoov [​*illuminating*​] seletus, vaid on pigem demüstifitseeriv seletus.

​**[lk.49]**​ Nagu ma [ülal](#PARA.37-5) juba märkisin, ei piirdu ega ammendu seletamine reduktiivse seletamisega. Reduktiivse seletamise peamine ülesanne on kaotada kõrgtaseme fenomeni ümbert sügava müsteeriumi oreool. Ta saavutab seda kõnealuse fenomeni toore meelevaldsuse taandamisega madalama taseme protsesside toorele meelevaldsusele. Kuivõrd madaltaseme protsessid võivad olla ise üsna toored ja meelevaldsed, ei pruugi reduktiivne seletus anda meile fenomenist väga sügavat arusaamist, kuid ta lubab meil vähemalt saada lahti tundest, et mängus on “veel midagi”.

Reduktiivsete seletuste ja selgusttoovate seletuste vahelist lünka saab aga üldjuhul täita palju täielikumalt tänu meie maailma füüsika kahele alusfaktile: *autonoomiale* ja *lihtsusele*​. Mikrofüüsikaline põhjuslikkus ja mikrofüüsikalised seletused on nähtavasti autonoomsed — selles mõttes et igal füüsikalisel sündmusel on füüsikaline seletus —, mis tähendab, et füüsika seadused suudavad seletada füüsikalisi sündmusi omaenda vahenditega. Lisaks on need seadused kaunis lihtsad, mistõttu ka seletused on küllalt kompaktsed. Kumbki neist punktidest ei tarvitseks kehtida. Maailmas, kus me elame, võiksid valitseda läbinisti emergentsed fundamentaalseadused, mis reguleeriksid kõrgtaseme konfiguratsioonide, näiteks organismide käitumist koostoimes ülalt alla kulgeva põhjuslikkusega, mis on mikrofüüsika seaduste suhtes ülimuslikud. (Millekski selletaoliseks pidasid meie maailma briti emergentistid, näiteks Alexander (1920) ja Broad (1925).) Mõeldav oleks ka, et seadusi, mis reguleerivad meie maailmas mikrofüüsikaliste entiteetide käitumist, oleks tohutu hulk ja nad oleksid äärmiselt käänulised, või et mikrofüüsikalised protsessid oleksid täiesti kaootilised ja seadusi ei olekski. Säärastes maailmades poleks kuigi palju lootust jõuda selgusttoovate reduktiivsete seletusteni, sest me ei suudaks iial lihtsustada madaltaseme kirjelduste meelevaldsust.

Kuid tundub, et aktuaalse maailma madatasemeline autonoomia ja lihtsus võimaldab tuua selgust ka väga keerukatesse protsessidesse. Madaltaseme faktidele, milledele kõrgtaseme fenomen toetub, on sageli omane mingi sügav ühetaolisus, mis võimaldab anda arusaadavaid seletusi. Mingist kindlast kõrgtasemelise põhjuslikkuse juhtumist, näiteks triklile vajutamisest, mille tagajärjel püssist tuleb lask, võime kõigepealt eraldada välja kimbu madaltaseme fakte, mis selle põhjuslikkuse tagavad, ning seejärel korrastada need faktid mingite lihtsate printsiipide alusel, ja nii õnnestubki meil kaunis lihtsalt selgitada, kuidas see põhjuslikkus teoks saab. See ei tarvitse alati õnnestuda. Mõnest valdkonnast, näiteks sotsioloogiast ja majandusteadusest, võib madaltaseme protsesside lihtsus olla nii kaugel, et selgusttoovate reduktiivsete seletusteni ei olegi võimalik jõuda, isegi kui fenomenid on ise loogiliselt superventsed. Seal, kus see nii on, tuleb sellega leppida: me võime piirduda nende valdkondade seletustega kõrgtasemel ja märkida, et tänu loogilisele superventsusele on reduktiivset seletust põhimõtteliselt võimalik leida, ehkki seda suudaks siis mõista ehk ainult mingi ülivõimekas olend.

​**[lk.50]**​ Paneme tähele ka, et ülaltoodud arutluskäigu põhjal osutuvad reduktiivsed seletused põhimõtteliselt *partikulaarseteks* — nad seletavad mingi fenomeni üksikuid juhtumeid, kuid mitte tingimata kõiki juhtumeid üheskoos. Seda oligi oodata. Kui mingi omadus võib realiseeruda mitmel eri viisil, siis ei olegi meil ju alust loota, et neid kõiki saaks katta üheainsa seletusega. Näiteks temperatuur realiseerub erinevates keskkondades üpris erinevalt, ja iga erijuhu kohta leidub erinev seletus. Ning — et tuua näide ka palju kõrgemalt tasemelt — on äärmiselt ebatõenäoline, et mõrvadele saaks anda üheainsa seletuse, mis kataks kõik mõrvajuhtumid. Siiski on üksikjuhtude seletustes mingi ühtsus sageli siiski olemas, selles mõttes et ühe juhu head seletust saab sageli kohaldada paljudele teistele juhtudele. Ka see on taas meie maailma fundamentaalse lihtsuse järelm, mitte seletuste paratamatu omadus. Paljud lihtsad ühtsustavad lood, mida madaltaseme protsesside kohta jutustada saab, on meie maailmas rakendatavad kõikjal või vähemalt paljudel väga erinevatel üksikjuhtudel. Samuti on ükskjuhud, eriti bioloogilistes teadustes, sageli mingis osas ühist päritolu, mistõttu ka nendega seostuvates madaltaseme protsessides on midagi sarnast. Seega ei ole teine mainitud raskustest — ühe või teise fenomeni konkreetsete juhtude seletuste ühtsustamine — nii tõsine, kui ta olla võiks. Igatahes on kesksel kohal just partikulaarsete juhtude seletamisel.

Reduktiivsete seletuste ja selgusttoovate seletuste vahelise lünga täitmise üle võiks arutleda veel pikalt, kuid olgugi et see teema vääriks omaette põhjalikku käsitlust, ei ole ta praegu minu jaoks nii oluline. Peaasi on, et loogilise superventsuse puudumisel (ja nagu ma allpool​ näidata püüan, teadvuse puhul see nii ongi) on igasugune reduktiivne seletamine võimatu, isegi kui olla kriteeriumides väga leplik. Oluline on ka see, et loogiline superventsus kaotab kõik jäljed kõrgtaseme fenomene ümbritsevast metafüüsilisest salapärast, taandades igasuguse neis fenomenides leiduva tooruse madalama taseme faktide toorusele​. Tähtis, kuigi mitte nii tähtis, on ka see, et kui loogiline superventsus kehtib, siis leidub ka nii- või teistsugune reduktiivne seletus. Kuigi need seletused ei pruugi olla selgusttoovad või kasulikud, ei ole see kaugeltki nii fundamentaalne puudus kui seletuste võimatus valdkondades, kus loogiline superventsus ei kehti.

#### Veel märkusi reduktiivse seletamise kohta

Teeme veel mõned märkused. Esiteks ei toimu laskumine fenomenide reduktiivsel seletamisel praktikas enamasti mikrofüüsika tasemeni välja. See oleks pööraselt raske ning tooks kaasa kõik raskused toore elementaarsusega​, millest äsja juttu oli. Selle asemel seletatakse kõrgtaseme fenomene veidi madalama taseme omaduste kaudu, näiteks paljunemist rakumehhanismide kaudu ja Kuu faase orbitaalliikumise kaudu. Seejuures loodetakse, et neile fundamentaalsematele fenomenidele ​**[lk.51]**​ saab omakorda anda reduktiivse seletuse mingite veelgi fundamentaalsemate omaduste kaudu. Kui kõik läheb hästi, saab bioloogilisi fenomene seletada rakufenomenide kaudu, need omakorda biokeemiliste fenomenide, neid edasi keemiliste fenomenide ja viimaseid füüsikaliste fenomenide kaudu. Füüsikalisi fenomene püütakse unifitseerida niipalju kui võimalik, ent mingil tasemel tuleb füüsikat võtta lõplikult elementaarsena: küsimusele, miks fundamentaalsed seadused või rajatingimused on just niisugused, nagu nad on, ei pruugi enam vastust olla. Esialgu ei ole see seletuste astmik palju enamat kui õhuloss, ent edu on siiski olnud märkimisväärne. Tänu loogilisele superventsusele ühes lihtsuse ja autonoomiaga kõige madalamal tasemel peaks niisugune teaduste seletuslik ühendus olema põhimõtteliselt võimalik. Kuid esialgu me veel ei tea, kas tegelikkuse keerukus ei takista selle unelma praktilist teostamist.

Teiseks on vähemasti mõeldav, et mingi fenomen võib olla reduktiivselt seletatav madaltaseme omaduste kaudu, kuid mitte reduktiivselt seletatav *simpliciter*​. Seda võib juhtuda olukorras, kus C-omadused on B-omaduste suhtes loogiliselt superventsed ja seega seletatavad B-omaduste kaudu, B-omadused ise aga ei ole füüsika suhtes loogiliselt superventsed. On selge, et niisugune seletus on mingis mõttes reduktiivne, mingis teises mõttes aga mitte. Mind huvitab siin peamiselt reduktiivne seletamine füüsikaliste omaduste kaudu, või omaduste kaudu, mis on seletatavad füüsikaliste omaduste kaudu jne. Isegi kui need C-omadused on mingis suhtelises mõttes reduktiivselt seletatavad, kõneleb juba nende olemasolu sellest, et reduktiivne seletatavus ei ole kõikehõlmav.

Kolmandaks, lokaalse loogilise superventsuse nõue on liiga range, et võimaldada reduktiivseid seletusi. Reduktiivse seletuse saab anda isegi nendele indiviidi omadustele, mis sõltuvad kontekstist. Selleks tuleb kirjeldada, kuidas realiseeruvad nende omaduste seosed keskkonnaga. Senikaua kui fenomen on globaalselt superventne, on teda võimalik reduktiivselt seletada mingite madaltaseme faktide kaudu, isegi kui nende faktide omavaheline ajaline ja ruumiline kaugus on väga suur.

Neljandaks jaguneb elu, õppimise, soojuse jts. fenomenide reduktiivne seletamine põhimõtteliselt kaheks eraldi projektiks. Esimene neist on *eksplitseerimine*​, kus me selgitame analüüsi teel, mida me õieti seletada kavatseme. Näiteks õppimist saab analüüsida kui teatavat liiki kohanemisprotsessi. Teine projekt on *seletamine* — siin me näitame, kuidas madaltaseme faktid sellesse analüüsi sobituvad. Esimene projekt on kontseptuaalne, teine on empiiriline. Paljude või enamiku fenomenide puhul on kontseptuaalne etapp üsna triviaalne. Kuid mõne fenomeni, näiteks uskumuste puhul, võib juba eksplitseerimine valmistada suuri raskusi. Muidugi ei ole praktikas nende kahe projekti vahel kunagi selget vahet, sest eksplitseerimine ja seletamine kulgevad paralleelselt.

### ​**[lk.52]**​ 4. Kontseptuaalne tõesus ja paratamatu tõesus\*

Minu arutlused superventsuse ja seletamise kohta toetusid suuresti loogilise võimalikkuse ja paratamatuse mõistele. Nüüd on aeg kõnelda sellest põhjalikumalt. Väite loogilise paratamatuse all võib mõelda eelkõige tema tõesust kõigis loogiliselt võimalikes maailmades. Niihästi asjaomase maailmade klassi mõtestamine kui ka väidete tõesuse hindamine neis maailmades nõuab mõningat ettevaatlikkust; sellest tuleb veidi pikemalt juttu käesolevas jaos allpool​. Väite loogilist paratamatust võib sõnastada ka tähendusliku tõesusena: väide on loogiliselt paratamatu, kui tema tõesuse tagavad temas figureerivate mõistete tähendused. Ent ka tähenduste tõlgendamine nõuab mõningat ettevaatust. Mõlemad kirjeldatud mõistmisviisid ja nendevaheline seos tulevad jutuks [siinsamas jaos allpool](#ORG.2.4.4).

(Nagu ülal, nii ka siin ei tule loogilist paratamatust mõista kitsas mõttes — tuletatavusena esimest järku loogikas või mõnes muus süntaktilis-formaalses süsteemis. On isegi põhjust oletada, et nende süsteemide aksioomide ja tuletusreeglite õigsus tulenebki nende loogilisest paratamatusest avaramas, algsemas mõttes.)

Kõik see eeldab, et me võtame vähemalt mingil määral tõsiselt *kontseptuaalset tõesust* — arusaama, et on väiteid, mille tõesus või väärus tuleneb ainuüksi neis figureerivate terminite tähendustest. Minu seniste arutluste põhipunktid on toetunud mitmesuguste mõistete kirjeldustele. Näiteks paljunemise reduktiivsest seletamisest kõneldes väitsin ma, et teatud funktsioonide talitlus tuleneb madaltaseme seikadest ning et nende funktsioonide talitlemine ammendabki paljunemise mõiste.

Kontseptuaalse tõe mõistel on mõnedes ringkondades halb maine sestsaadik, kui seda kritiseeris Quine (1951), kes väitis, et kontseptuaalsete tõdede ja empiiriliste tõdede vahel puudub igasugune kasulik erinevus. Argumendid kontseptuaalse tõe​[[68]](#footnote-69) vastu rühmituvad enamasti ühe või teise ümber järgnevatest:

1. Enamikul mõistetest puudub definitsioon, millest nähtuksid tarvilikud ja piisavad tingimused. (Seda on kirjanduses mainitud palju kordi, kuid tavaliselt omistatakse seda tähelepanekut Wittgensteinile (1953).)
2. Enamikku väiteid, mis on pealtnäha kontseptuaalsed tõed, on tegelikult võimalik revideerida ja piisava empiirilise tõendusmaterjali alusel kummutada. (Selle tõdemuse autor on Quine.)
3. Kripke (1972) esitatud argumentidest aposterioorse paratamatuse kohta nähtub, et paljude terminite puhul ei ole võimalik​ ​*a priori* teada, missugustel tingimustel nad on kõigis võimalikes maailmades rakendatavad.

Neil kolmel punktil on jõudu liigselt lihtsustava arusaamise vastu kontseptuaalsest tõest, kuid mitte selle vastu, kuidas mina neid mõisteid kasutan. Muu seas ​**[lk.53]**​ ei mõjuta need kaalutlused *superventsuskonditsionaale* — lauseid kujul “kui mingi olukorra A-faktid on ​*X*​, siis B-faktid on​ ​*Y*​”, kus A-faktid annavad olukorra täieliku kirjelduse fundamentaalsel tasemel. Mingeid muid kontseptuaalseid tõdesid minu argumendid ei vaja ja me näeme peagi, et ükski ülaltoodud punktidest neid ei riiva. Analüüsin allpool põhjalikumalt ka kontseptuaalse tõe ja paratamatu tõe omavahelist suhet ning täpsustan nende rolli loogilise superventsuse mõtestamisel.

#### Mõisted

Üheste ja rangete definitsioonide puudumine on kõigist raskustest, mida kontseptuaalsed tõede valmistavad, kõige vähem tõsine. Mitte ükski minu argumentidest ei sõltu sääraste definitsioonide olemasolust. Aeg-ajalt ma küll viitan ühe või teise mõiste analüüsile, kuid mulle piisab sellest, kui need analüüsid asja sisu umbkaudseltki tabavad, ega oota, et nad annaksid mulle kätte täpsed tarvilikud ja piisavad tingimused. Enamik mõisteid (näiteks “elu”) on tegelikus kasutuses nagunii mõnevõrra ebamäärased, ja pole erilist mõtet püüda ebamäärasusest meelevaldsete täpsustustega lahti saada. Selle asemel, et öelda: “Süsteem on elus parajasti siis, kui ta paljuneb, tema kohastumistegur on 800​[[69]](#footnote-70) või enam ning tema ainevahetuse efektiivsus on vähemalt 75 %, või ilmutab neid omadusi sobivas ühenduses nii- ja niisuguste omadustega”, aitab lihtsast märkusest, et kui mingi süsteem ilmutab neid omadusi küllalt suurel määral, siis vastab ta sõna “elus” tähendusele. Kui asjakohaste madaltaseme faktide kirjeldus sedastab need faktid süsteemi paljunemise, kohastumise, ainevahetuse jne. kohta, siis on ühtlasi sedastatud ka faktid selle kohta, kas see süsteem on elus — niivõrd kui see üldse on faktiküsimus.

Seda võib näitlikult kokku võtta skemaatilise diagrammi abil (joonis 2.1), mis kujutab kaht eri viisi, kuidas kõrgtaseme omadus P võib sõltuda kahest madaltaseme parameetrist A ja B, millest kumbki võib omandada väärtusi mingis vahemikus. Selge definitsioon, mis näitaks ära tarvilikud ja piisavad tingimused, vastaks vasakpoolsele diagrammile, kus must ristkülik kujutab omaduse P realiseerumispiirkonda. Tegelikus elus on sõltuvused aga alatasa niisugused nagu parempoolsel diagrammil, kus P-suspiirkonna kontuurid on laialivalguvad: on olemas suur piirkond, kus ei ole võimalik kindlalt öelda, kas omadus P kehtib või mitte, kuid leidub ka piirkond, kus asi on selge. (Ei ole võib-olla kindel, kas bakter ja arvutiviirus on elusolendid, kuid koer on seda kahtlemata.) Objekti puhul, mis jääb diagrammi mustale alale, on A ja B määr P realiseerumiseks piisavad ja konditsionaal “kui objektil​ ​*x* on niisugused A ja B väärtused, siis *x* on P” on kontseptuaalne tõde hoolimata sellest, et omadusel P puudub ühene definitsioon. Igasugune ebaselgus sellistes konditsionaalides — hallidel aladel — kajastab ebaselgust tegelikes faktides, nagu peabki. Seda pilti saab triviaalselt üldistada suvalisest arvust teguritest ​**[lk.54]**​ sõltuvatele omadustele ja superventsuskonditsionaalidele üldiselt.

**Joonis 2.1.** Kaks viisi, kuidas omadus P võib sõltuda omadustest A ja B.

Oleme jõudnud olulise tulemuseni: üks faktistik võib tuleneda teisest ka siis, kui tema mõisted ei ole puhtalt defineeritavad teise faktistiku kaudu. Äsjakirjeldatud juhtum ongi näide just selle kohta: omadusel P puudub lihtne definitsioon parameetrite A ja B kaudu, kuid mõnel juhul saab faktidest parameetrite A ja B kohta tuletada fakte P kohta. Teiseks näiteks võime tuua kinniste kõverate ümmargususe kahemõõtmelises ruumis (joonis 2.2). Vististi ei ole võimalik anda ümmargususe täiuslikku definitsiooni lihtsamate matemaatiliste mõistete kaudu.[[70]](#footnote-71) Kuid vaatame vasakpoolset kõverat, mida kirjeldab võrrand 2x² + 3y² = 1. Kui ümmargususe puhul saab üldse rääkida faktidest, siis on see kõver faktiliselt ümmargune (eriti võrreldes parempoolse kõveraga, mis kindlasti ei ole ümmargune.) Lisaks tuleneb see fakt kujundi matemaatilisest kirjeldusest — kirjeldus ja ümmarguguse mõiste määratlevad ühiselt kõvera faktilise ümmargususe. Sellest, et B-faktid võivad tuleneda A-faktidest, ilma et B-faktid oleksid defineeritavad A-faktide kaudu, järeldub, et definitsioonide puudumine ei mõjuta loogilise superventsuse mõistet. (Allpool, kui tuleb juttu keerukamatest küsimustest ja vastuväidetest loogilise superventsuse teemal, tasub meenutada seda näidet.)

​**[lk.55]**​ **Joonis 2.2.** Ümar kõver 2x² + 3y² = 1 ja tema mitteümar sõber.

​**[lk.54]**​ Seda tulemust võib sõnastada ka teisiti, öeldes, et mõiste “tähenduseks”, mida meil enamasti tarvis on, ei ole mitte tema definitsioon, vaid *intensioon* — funktsioon, mis kirjeldab, kuidas mõistet ühes või teises olukorras kohaldatakse. Mõnikord on intensiooni võimalik formuleerida definitsioonina, kuid, nagu toodud näidetest selgub, ei ole see alati nii. Ent niipea kui me teame, kuidas mõiste ühes või teises olukorras rakendub, on meil tema intensioon teada; ja, nagu ma peagi näitan, mingeid muid “tähendusi” minu argumendid üldiselt ei vajagi.

#### Revideeritavus

Teine Quine’i (1951) vastuväide ütleb, et kontseptuaalseid tõdesid on alati võimalik revideerida, kui selleks on piisavalt empiirilist tõendusmaterjali. Näiteks kui tõendid sunnivad meid mingi teooria üht või teist aluslauset​[[71]](#footnote-72) revideerima, siis võib juhtuda, et väide, mida seni peeti kontseptuaalselt tõeseks, osutub vääraks.

Nii on see tõepoolest paljude kontseptuaalseteks peetavate tõdede puhul, kuid superventsuskonditsionaalidega, mida meie siin vaatleme, kujul “kui madaltaseme faktid osutuvad nii-  ja niisugusteks, siis on kõrgtaseme faktid nii-  ja niisugused”, on asi teisiti. Selle tingimuse eesliikmes​[[72]](#footnote-73) mainitud faktide hulka kuuluvad õigupoolest kõik relevantsed empiirilised tegurid. Empiirilised tõendid võivad meile näidata konditsionaali eesliikme, kuid mitte konditsionaali enda väärust. Me võime minna äärmuseni ja sõnastada eesliikme nii, et see sisaldaks maailma kõigi madaltaseme faktide täieliku kirjelduse. Just eesliikme täielikkus tagab konditsionaali tõeväärtuse sõltumatuse empiirilistest andmetest. (Pilti teeb veidi keerukamaks aposterioorse paratamatuste olemasolu, millest [peagi](#ORG.2.4.3) juttu tuleb. Siin huvitavad mind ainult episteemilised konditsionaalid aktuaalses maailmas võimalike asjaseisude kohta.)

Kuigi revideeritavuskaalutlused lubaksid veenvalt väita, et lihtsaid ja lühikesi kontseptuaalseid tõdesid ei ole palju, ei ole neis kaalutlustes midagi, mis räägiksid vastu seesugustele paljude kitsendustega ja keerukatele kontseptuaalse tõdedele, mida mina siin vaatlen. Nende tähelepanekute kokkuvõte on, et kõrgtaseme lausete tõesustingimused ei tarvitse olla hõlpsasti lokaliseeritavad, kuna neid võivad ühel või teisel viisil kaudselt mõjutada väga mitmesugused tegurid; kuid see oht ei ähvarda globaalseid tõesustingimusi, mida väljendavad superventsuskonditsionaalid. Tõepoolest, kui tähendus määratleb funktsiooni võimalike maailmade hulgast osutusklasside (s.t. intensioonide) hulka, ja kui võimalikele maailmadele saab anda lõplikke kirjeldusi (näiteks ​**[lk.56]**​ nende maailmade aluskvaliteetide jaotuse kaudu), siis saame automaatselt tulemuseks tohutu hulga kontseptuaalselt tõeseid konditsionaale.

#### Aposterioorne paratamatus.

Traditsiooniliselt on arvatud, et kõiki kontseptuaalseid tõdesid on, nagu kõiki paratamatuid tõdesidki, võimalik teada ​*a priori*​, ning et aprioorsete tõdede, paratamatute tõdede ja kontseptuaalsete tõdede klass on omavahel lähedases suguluses või koguni mahult kattuvad. Saul Kripke raamat “Nimetamine ja paratamatus” [“Naming and Necessity”] (1972) lõi selle pildi segi, näidates, et on olemas suur klass paratamatult tõeseid lauseid, mille tõesust ei ole võimalik ​*a priori* teada. Niisugune on näiteks lause “Vesi on H₂O”. Meil ei ole võimalik ​*a priori* teada, et see lause on tõene; meie poolest (teadmiste põhjal, mis meil olid enne seda, kui hakkasime asja uurima) võiks vesi koosneda ka millestki muust, võib-olla näiteks XYZ-st. Kripke väidab, et kuna aktuaalses maailmas on vesi H₂O, siis on ta ka kõigis võimalikes maailmades H₂O. Sellest järeldub, et “vesi on H₂O” on oma aposterioorsusest hoolimata paratamatu tõde.

See asjaolu toob minu mudelisse paar raskust. Näiteks mõne tõlgenduse järgi on need paratamatud tõed kontseptuaalsed tõed, millest järeldub, et mitte kõiki kontseptuaalseid tõdesid ei ole võimalik teada ​*a priori*​. Teiste tõlgenduste järgi ei ole säärased laused kontseptuaalsed tõed, kuid sellisel juhul katkeb seos kontseptuaalse tõe ja paratamatuse vahel. Käesolevas raamatus rakendan ma paratamatu tõesuse väljaselgitamiseks mitmes kohas aprioorseid meetodeid — teguviis, mille Kripke käsitlus paljude arvates kahtluse alla seab.

Analüüs näitab minu arvates, et need raskused ei muuda minu argumentides põhimõtteliselt midagi. Kuid tasub näha vaeva ja teha endale selgeks, milles siin asi on. Kulutan nüüd pisut aega, et luua süstemaatiline mudel nende küsimuste uurimiseks, mis hakkavad edaspidi korduma. Muu hulgas kirjeldan, kuidas kujutada Kripke tõdemusi kahemõõtmelise pildina, mille koordinaatideks on tähendus ja paratamatus. See mudel sünteesib omavahel paljusid ideid, mida on välja öelnud Kripke, Putnam, Kaplan, Stalnaker, Lewis, Evans, Davies ja Humberstone ning teised, kes on neid kahemõõtmelisi fenomene uurinud.

Traditsioonilise arusaama kohaselt osutusest​, mis ulatub tagasi Fregeni, ent mida mina väljendan siin nüüdisaegses terminoloogias, defineerib iga mõiste funktsiooni $f:W\rightarrow R$, mille argumentideks on võimalikud maailmad ja väärtusteks vastava mõiste referendid. Niisugust funktsiooni nimetatakse sageli *intensiooniks*​; ühes maailma täpse kirjeldusega $w$ defineerib ta *ekstensiooni*​ $f\left(w\right)$. Frege enda järgi on igal mõistel *mõte* [​*sense*​], mis määratleb mõiste osutuse [​*reference*​] olenevalt maailma seisundist; seega vastavad need mõistete mõtted küllalt täpselt intensioonidele. Seda mõtet on sageli tõlgendatud kui vastava mõiste *tähendust*​.

Uuemates töödes on hakatud tunnistama, et ükski intensioon ei suuda üksinda teha kõike, mida peab tegema tähendus. Pilt, mille töötas välja Kripke, teeb olukorra keerukamaks, näidates, et osutusmehhanismid ​**[lk.57]**​ on aktuaalses maailmas hoopis teistsugused kui võimalikes kontrafaktilistes maailmades. Mingis mõttes võib öelda, et Kripke pilt lahutab Frege pildi kaheks eraldi tasemeks.

Kripke arusaama väljendamiseks võib öelda, et iga mõistega kaasneb tegelikult kaks intensiooni. Teisiti öeldes, mõiste referent sõltub maailma seisundist kahel täiesti erineval viisil. Üks sõltuvus fikseerib referendid aktuaalses maailmas, olenevalt sellest, missugune see maailm tegelikult on: ühel juhul eraldab mõiste välja ühe asja, teisel teise. Teine sõltuvus määratleb referendid kontrafaktilistes maailmades, kui referendid aktuaalses maailmas on juba fikseeritud. Kummalegi sõltuvusele vastab omaette intensioon, mida ma hakkan edaspidi nimetama vastavalt *primaarseks* ​ja *sekundaarseks* intensiooniks.

Mõiste *primaarne intensioon* on funktsioon, mille argumendiks on maailm ja väärtuseks ekstensioon, mis kajastab referendi fikseeringut aktuaalses maailmas. Ta eraldab maailmast välja referendi, millele mõiste viitaks, kui see maailm osutuks aktuaalseks maailmaks. Võtame näiteks mõiste “vesi”. Kui aktuaalses maailmas täidaks ookeane ja järvi XYZ, siis oleks mõiste “vesi” referendiks XYZ,[[73]](#footnote-74) ent kuna faktiliselt täidab neid H₂O, siis on selleks H₂O. Seega on mõiste “vesi” primaarseks intensiooniks funktsioon, mis seab​ XYZ-maailma vastavusse​[[74]](#footnote-75) XYZ-ga ja H₂O-maailma H₂O-ga. Umbkaudse lähendusena võiks öelda, et primaarne intensioon osutab läbipaistvale joogikõlblikule vedelikule, mis moodustab suurema osa ookeanide ja järvede ainest või, lühidalt öeldes, *vesijale ollusele* [​*watery stuff*​].

Ent Kripke (nagu ka Putnam 1975) tähendab, et kuna aktuaalses maailmas osutub sõna “vesi” referendiks H₂O, siis on mõistlik öelda, et vesi on kõigis kontrafaktilistes maailmades H₂O. Mõiste “vesi” sekundaarne intensioon osutab igas kontrafaktilises maailmas veele, nii et kui Kripkel ja Putnamil on õigus, siis osutab sekundaarne intensioon H₂O-le kõigis maailmades.[[75]](#footnote-76)

Minu jaoks on mõiste juures oluline eelkõige tema primaarne intensioon: just primaarne intensioon hõlmab kõik selle, mis looduslike fenomenide mõistetes seletust vajab. Kui keegi nõuab meilt seletust vee kohta siis, kui me veel ei tea, et vesi on faktiliselt H₂O, siis soovib ta õigupoolest enam-vähem seletust tema keskkonnas leiduva läbipaistva joogikõlbliku vedeliku kohta. Seda, et vesi on H₂O, teame me alles siis, kui meil on täielik seletus käes. Erinevalt sekundaarsest intensioonist ei sõltu mõiste primaarne intensioon empiirilistest teguritest: ta annab meile täpse kirjelduse, kuidas osutus sõltub sellest, missugune on väline maailm, ega sõltu seega ise sellest, missuguseks väline maailm osutub.

Mõiste primaarse intensiooni lühike kirjeldus *à la* “kõige levinum läbipaistev joogikõlblik vedelik keskkonnas” on loomulikult lihtsustus. Tegeliku intensiooni kindlaksmääramiseks tuleb üksikasjalikult vaadelda konkreetseid stsenaariume: mida me ütleksime, kui selguks, et ​**[lk.58]**​ maailm on niisugune? mida me ütleksime, kui selguks, et maailm on naasugune? Näiteks kui selguks, et vedelik järvedes on H₂O ja vedelik ookeanides on XYZ, ütleksime arvatavasti, et nii üks kui teine on vesi; kui ookeanide ja järvede ollus oleks ainete segu, milles oleks 95 % ainet A ja 5 % ainet B, ütleksime arvatavasti, et A on vesi, aga B midagi muud; kui selguks, et maailmas leidub aine, mis pole ei läbipaistev ega joogikõlblik, kuid millel on sobiv mikrofüüsikaline seos meie keskkonnas leiduva läbipaistva, joogikõlbliku vedelikuga, nimetaksime arvatavasti ka seda ainet veeks (nagu me teemegi jääga ja reoveega). Täielik tingimuste kogum, mida aine peab rahuldama, et me nimetaksime teda veeks, on äärealadel kaunis laialivalguv ja ei tarvitse refleksioonile otsekohe ilmne olla; kuid kõik need asjaolud ei muuda pilti oluliselt. Tehisliku terminiga “vesijas ollus” püüan ma väljendada primaarset intensiooni, olgu see milline tahes.[[76]](#footnote-77)

Otsustamaks, mis on mingi mõiste referent aktuaalses maailmas, on mõnel juhul tarvis üpris palju mõtelda: näiteks mõiste “mass” puhul, kui selgub, et aktuaalses maailmas kehtib relatiivsusteooria[[77]](#footnote-78) [[78]](#footnote-79) [[79]](#footnote-80), ja võib-olla ka küsimuses, mida tuleks aktuaalses maailmas pidada uskumuseks. Niisiis võib otsustamine, millisele mitmest võimalikust kandidaadist primaarne intensioon aktuaalses maailmas osutab, nõuda küllalt palju mõttetööd. Kuid see ei tähenda, et me ei teaks seda ​*a priori*​: me saame selle üle arutleda olenemata sellest, missuguseks maailm tegelikult osutub. Võib näiteks juhtuda, et keegi seab relatiivsusteooriat kinnitavad katsed kahtluse alla, mistõttu me ei ole kindlad, kas aktuaalne maailm on relativistlik; ent nii või teisiti on meil võimalik arutleda sõna “mass” osutuse üle, olenemata sellest, kumb võimalus osutub aktuaalseks.

(Keelekogukonnas käibivate mõistete primaarsete intensioonide analüüsimisel tekib mitmesuguseid keerukusi. Nendega võib toime tulla, kui panna tähele, et indiviidi jaoks võib mõiste primaarset intensiooni mõjutada tema respekt vastava mõiste vastu niisugusena, nagu see on käibel kogukonnas: nii võib minu mõiste “jalakas” osutada sellele, mida nimetavad “jalakaks” teised. Kuid nii või teisiti ei mängi seesugused probleemid rolli küsimustes, mis mind siin huvitavad — nende vaatlemisel võib väga hästi eeldada, nagu koosneks kogukond ühestainsast isikust või nagu oleks kõigil indiviididel ühevõrra täielik informatsioon, või isegi nagu oleks kogukond üksainus hiiglasuur indiviid. Üldise semantikateooria konstrueerimisel primaarsete intensioonide alusel võib kerkida ka tehnilisi raskusi. Näiteks: kas mingi mõiste osutus on sellele mõistele olemuslikult omane? Või: kas eri kõnelejad võivad seostada ühe ja sama sõnaga erinevaid primaarseid intensioone? Ent kuna ma ei püüa siin luua täielikku semantilist teooriat, siis võib need peensused vaatluse alt välja jätta.

Filosoofid vaatavad primaarsetele intensioonidele mõnikord kahtlustava pilguga, sest need meenutavad neile nn. deskriptiivset osutusteooriat[[80]](#footnote-81). Kuid kirjeldused ei mängi minu mudelis olulist rolli; ma kasutan ​**[lk.59]**​ neid üksnes selleks, et paremini iseloomustada mõningaid funktsioone võimalike maailmade hulgast ekstensioonide hulka​[[81]](#footnote-82). Tõeliselt olulised ei ole ekstensioonide kirjeldused, vaid funktsioon ise. See pilt on täiesti ühitatav nn. põhjusliku osutusteooriaga[[82]](#footnote-83): tuleb lihtsalt tähele panna, et näiteks vee mõiste primaarne intensioon võib eeldada sobivat põhjuslikku seost referendi ja subjekti vahel. Tõepoolest, meie usk põhjuslikku osutusteooriasse saab alguse eelkõige mitmesugusete võimaluste vaatlusest, missuguseks aktuaalne maailm osutuda võiks, ja sedastamisest, mis oleks igal neist juhtudest mõiste referendiks — see tähendab, mõiste primaarse intensiooni kindlakstegemisest neis maailmades.).

Kuna aktuaalses maailmas osutab mõiste “vesi” vesijale ollusele, siis võiks arvata, et vesi on vesijas ollus kõigis võimalikes maailmades. Kripke ja Putnam näitasid, et see ei ole nii: kui vesi on aktuaalses maailmas H₂O, siis on vesi kõigis võimalikes maailmades küll H₂O, kuid maailmas, kus kõige levinum läbipaistev ja joogikõlblik vedelik ei ole H₂O, vaid XYZ (Putnami “Teisikmaal”), ei ole see vedelik vesi, vaid ainult vesijas ollus. Kõik seda väljendab mõiste “vesi” sekundaarne intensioon, mis osutab kõigis maailmades veele: see tähendab, ta osutab kõigis maailmades H₂O-le.

Mõiste (näiteks vee mõiste) sekundaarset intensiooni ei saa kindlaks teha ​*a priori*​, sest sekundaarne intensioon sõltub asjaseisust aktuaalses maailmas. Kuid siiski on sekundaarsel intensioonil lähedane seos primaarse intensiooniga. Mõiste sekundaarse intensiooni kindlakstegemiseks tuleb kõigepealt leida tema primaarne intensioon aktuaalses maailmas ja seejärel rigidiseerida tulemus nii, et mõiste hakkaks osutama samaliiki asjale kõigis võimalikes maailmades. Kuna mõiste “vesi” primaarne intensioon — “vesijas ollus” — osutab aktuaalses maailmas H₂O-le, siis osutab sekundaarne intensioon pärast rigidiseerimist H₂O-le kõigis võimalikes maailmades.

Kokkuvõtlikult võib öelda, et “vee” kontseptuaalseks ekvivalendiks on “​*dthat*​(vesijas ollus)”, kus *dthat* on variant Kaplani rigidiseerimisoperaatorist​, mis teeb intensioonist tema väärtuse põhjal aktuaalses maailmas rigiidse designaatori (Kaplan 1979). Ühtne fregeaanlik intensioon laguneb kaheks: primaarseks intensiooniks (“vesijas ollus”), mis fikseerib referendi aktuaalses maailmas, ja sekundaarseks intensiooniks (“H₂O”), mis osutab referendile kontrafaktilistes võimalikes maailmades ning sõltub asjaseisust aktuaalses maailmas.

(Mõnikord kaldutakse arvama, et aposterioorne paratamatus teeb aprioorse kontseptuaalse analüüsi irrelevantseks, kuid see on alusetu oletus. Juba enne seda, kui jäikadest designatsioonidest ja muust sarnasest on mõtet kõnelema hakatagi, oskame me juba öelda, mis nimelt teeb aktuaalse maailma objektist​ ​*X* midagi niisugust, mida on võimalik pidada mõiste ​“​*X*​” referendiks. Kuid selleks tuleb analüüsida primaarset intensiooni. Ja see analüüs on aprioorne, kuna ta eeldab vastust küsimusele, millele viitaks mõiste *X*​, kui asjaseis aktuaalses maailmas oleks nii- või naasugune. Kui meil on võimalik aktuaalse maailma asjaseisu tundes teada, millele meie mõisted viitavad, siis on meil võimalik teada, millele meie mõisted ​**[lk.60]**​ viitaksid, kui asjaseis aktuaalses maailmas oleks nii- või naasugune. See, kas asjaseis aktuaalses maailmas on tegelikult nii- või naasugune, ei mõjuta kuigivõrd vastust sellele küsimusele, kuid aitab tähelepanu koondada.)

Nii primaarseid kui ka sekundaarseid intensioone võib vaadelda kui funktsioone võimalike maailmade hulgast ekstensioonide hulka ​*f* :​ ​*W* → *R*​, kusjuures siin tuleb võimalikke maailmu mõista veidi teisiti. Võiksime öelda, et mõiste primaarne intensioon osutab mõiste referendile mingis maailmas, kui seda maailma vaadeldakse aktuaalse maailmana — s.t. aktuaalse maailma kandidaadina mõtleja jaoks —, sekundaarne intensioon aga osutab mõiste referendile mingis maailmas, kui seda maailma vaadeldakse kontrafaktilise maailmana eeldusel, et mõtleja aktuaalne maailm on juba fikseeritud. Kui vaadelda XYZ-maailma aktuaalse maailmana, siis osutab minu sõna “vesi” XYZ-le, kui aga kontrafaktilisena, siis H₂O-le.

Vahetegemine nende kahe viisi vahel, kuidas me maailmu vaatleme, on väga sarnane Kaplani (1989) eristusega, kelle järgi ütlustel on *lausungikontekst*​ ja *otsustussituatsioon*​. Kui me vaatleme maailma *w* kontrafaktilisena, siis on meie lausungi kontekstiks aktuaalne maailm ja otsustussituatsiooniks maailm ​*w*​. Näiteks kui ma ütlen siinses maailmas “maailmameri koosneb veest” ja püüan leida selle tõeväärtust XYZ­-maailmas, siis on “vee” referendiks H₂O ja minu lause on väär. Kui me aga vaatleme maailma *w* aktuaalse maailmana, siis mõtleme temast kui potentsiaalsest lausungikontekstist ja uurime, missugune oleks asjaseis, kui ütluse kontekstiks osutuks maailm *w*​. Kui minu lause “maailmameri koosneb veest” kontekstiks osutuks XYZ-maailm, siis oleks lause tõene, kui hinnata tema tõeväärtust XYZ-maailmas. Primaarne intensioon on seega väga sarnane sellega, mida Kaplan nimetab termini *karakteriks*​, kuigi mõnede erinevustega[[83]](#footnote-84) [[84]](#footnote-85), sekundaarne intensioon vastab aga sellele, mida ta nimetab termini *sisuks*​.

Nende kahe vahel on mõningane asümmeetria, sest Quine (1969) nimetab *keskpunktiga*​ võimalikuks maailmaks lausungikonteksti, aga mitte otsustussituatsiooni. Keskpunktiga võimalik maailm on järjestusega paar, mis koosneb maailmast ja *keskpunktist*​, milleks on kõnealust sõna lausuva agendi vaatepunkt. Keskpunkti koosseisu kuuluvad (peale võib-olla veel millegi) “väljavalitud” indiviid ja ajahetk. (Idee autoriks on Lewis (1979); Quine arvab, et keskpunktiks võiks olla mingi punkt aegruumis.) Keskpunkti ülesanne on kajastada asjaolu, et sõnad (näiteks “vesi”) osutavad minu jaoks teistsugustele ekstensioonidele kui minu teisikule Teisikmaal, kuigi me elame samas universumis.[[85]](#footnote-86) Ainsaks erinevuseks meie vahel on meie asukoht universumis, ja just asukohast tulenevad olulised erinevused referendi fikseerimise protsessis.

See fenomen on eriti selgesti jälgitav indeksikaalsete sõnade puhul nagu näiteks “mina”, mille referent ei sõltu ainult maailma üldisest seisundist, vaid ka lausujast: sõna “mina” primaarne intensioon osutab indiviidile, kes on maailma keskpunktis. (Minu “mina”-mõiste sekundaarne intensioon osutab David Chalmersile kõigis võimalikes maailmades). ​**[lk.61]**​ Ent ka sellistel mõistetel nagu “vesi” on, kuigi vähem ilmsena, olemas indeksikaalne element, mida võib laias laastus sõnastada kui “​*dthat*​(üldlevinud läbipaistev, joogikõlblik vedelik meie keskkonnas)”.[[86]](#footnote-87) Just selle indeksikaalse elemendi tõttu tuleb primaarsed intensioonid panna sõltuma keskpunktiga maailmadest. Ent kui referent aktuaalses maailmas on fikseeritud, ei ole referentide kindlaksmääramiseks kontrafaktilises maailmas keskpunkti enam tarvis. Otsustussituatsiooniks saab seega võtta lihtsa, keskpunktita võimaliku maailma.

Kõike seda saab formaliseerida, kui panna tähele, et kontrafaktilistes maailmades ei piisa referentide täielikuks aprioorseks kindlaksmääramiseks ühekohalisest funktsioonist $f:W\rightarrow R$. Kontrafaktilises maailmas sõltub osutus nii sellest maailmast endast kui ka asjaseisust aktuaalses maailmas. Teisiti öeldes, mõiste defineerib kahekohalise funktsiooni

kus $W^{}$ on võimalike keskpunktiga maailmade ruum ja $W$ on tavaliste võimalike maailmade ruum. Funktsiooni esimeseks argumendiks[[87]](#footnote-88) on lausungikontekst ehk aktuaalses maailmas võimalik asjaseis, teiseks argumendiks aga otsustussituatsioon ehk kontrafaktiline võimalik maailm. Võib ka öelda, et mõiste defineerib funktsioonide pere

iga aktuaalses maailmas võimaliku asjaseisu $v\in W^{}$ puhul, kus $F\_{v}\left(w\right)=F\left(v,w\right)$. Kui näiteks mõisteks on “vesi” ja $a$ on maailm, kus vesijas ollus on H₂O, siis osutab $F\_{z}$ kõigis võimalikes maailmades H₂O-le. Arvestades, et meie maailmas on vesi tõepoolest H₂O, määrab see funktsioon $F\_{a}$ sõna “vesi” õiged rakendustingimused kontrafaktilistes maailmades. Kui meie maailm osutuks teistsuguseks maailmaks $b$, kus vesijas ollus oleks XYZ, siis määratleks vastavad rakendustingimused teistsugune intensioon $F\_{b}$, mis osutaks igas võimalikus maailmas XYZ-le.

Funktsiooni $F$ saab kindlaks määrata ​*a priori*​, kuna kõik aposterioorsed tegurid sisalduvad tema argumentides. Funktsioonist $F$ on võimalik taastuletada meie mõlemad ühekohalised intensioonid. Primaarne intensioon on funktsioon $f:W^{}\rightarrow R$, mille määrab “diagonaalkujutus” $f:w↦F\left(w,w’\right)$, kus maailm $w’$ erineb maailmast $w$ ainult keskpunkti puudumise poolest. See funktsioon fikseerib mõiste referendi aktuaalses maailmas. Sekundaarne intensioon on kujutus $F\_{a}:w↦F\left(a,w\right)$, kus $a$ on meie aktuaalne maailm. See intensioon osutab referentidele kontrafaktilistes maailmades. Siit järeldub otseselt, et aktuaalsele maailmale rakendamisel primaarne ja sekundaarne intensioon ühtivad: $f\left(a\right)=F\_{a}\left(a’\right)=F\left(a,a’\right)$.

Tavaliselt saab toimida ka vastupidises suunas: kahekohalise funktsiooni $F$ ja seega ka sekundaarse intensiooni $F\_{a}$ saab enamasti tuletada primaarsest intensioonist $f$​**[lk.62]**​ “reegli” abil, mis sätestab, kuidas sekundaarne intensioon sõltub primaarsest intensioonist ja aktuaalsest maailmast $a$. Reegel oleneb mõiste tüübist. Mõiste puhul, mis on jäik designaator, seisneb reegel selles, et sekundaarne intensioon osutab maailmas $w$ sellele, millele primaarne intensioon osutab maailmas $a$ (looduslikke liike tähistavate sõnade puhul aga võib-olla millelegi, millel on samasugue alusstruktuur nagu sellel, millele primaarne intensioon osutab maailmas $a$). Formaalsemalt: olgu $D:R×W\rightarrow R$ “projektsioonioperaator”, mis teeb mingis maailmas välja valitud klassi liikmest “sama” klassi liikme mingis teises võimalikus maailmas. Siis on sekundaarseks intensiooniks $F\_{a}$ lihtsalt funktsioon $D\left(f\left(a\right),−\right)$, mida võib vaadelda kui *dthat*​-operaatori rakendust intensioonile, mille defineerib $f$.

Teistel (mitteindeksikaalsetel)[[88]](#footnote-89) mõistetel on sekundaarset intensiooni lihtsam tuletada kui primaarset. “Kirjeldavate” väljendite puhul, nagu “arst”, “ruut” või “vesijas ollus” ei mängi rigiidsed designaatorid mingit erilist rolli: nende mõistete kohaldamine kontrafaktilistele maailmadele ei sõltu sellest, missuguseks osutub aktuaalne maailm. Nende mõistete sekundaarsed intensioonid on primaarsete intensioonide lihtsad koopiad (kui mitte arvestada erinevusi, mis tulenevad keskpunkti olemasolust). Mudel, mida ma siin kirjeldasin, võimaldab käsitleda mõlemat liiki mõisteid.

Omadusi tähistavaid sõnu, nagu näiteks “kuum”, saab intensionaalsuse seisukohalt kujutada kahel viisil. Me võime vaadelda omaduste intensioone funktsioonidena maailmade hulgast indiviidiklasside hulka (mille elementideks on seda omadust realiseerivad indiviidid), või funktsioonidena maailmade hulgast omaduste endi hulka. Minu mudelisse sobivad mõlemad võimalused: mõlemal juhul on hõlpus määrata kindlaks primaarne ja sekundaarne intensioon, ja lihtne on ka mõlema viisi vahel edasi-tagasi liikuda. Allpool toimin ma tavaliselt siiski esimesel viisil, nii et sõna “kuum” primaarne intensioon osutab entiteetidele, mida saab aktuaalses maailmas — olenevalt sellest, missugune see maailm osutub — pidada kuumadeks, ja sekundaarne intensioon osutab asjadele, mis on kontrafaktilises maailmas kuumad eeldusel, et aktuaalne maailm osutub niisuguseks, nagu ta on.

.

Mõiste tähenduse kandidaatideks võib pidada niihästi primaarseid kui ka sekundaarseid intensioone. Ma arvan, et ei ole mõtet valida välja üks neist, et pidada seda siis tõeliseks tähenduseks; “tähendus” on siin pigem autiitel​. Me võime väga hästi mõelda primaarsest ja sekundaarsest intensioonist ka kui tähenduse aprioorsest ja aposterioorsest aspektist.

Niisugusel korral on kumbagi liiki intensioon aluseks kindlat liiki kontseptuaalsetele tõdedele ehk tõdedele tähenduse alusel​[[89]](#footnote-90). Primaarne intensioon on aluseks aprioorsetele tõdedele, nagu “Vesi on vesijas ollus”. Sellised laused on tõesed olenemata sellest, missuguseks osutub aktuaalne maailm, kuigi nad ei tarvitse kehtida kõigis mitteaktuaalsetes võimalikes maailmades. Sekundaarne intensioon ei ole aluseks aprioorsetele tõdedele, kuid on aluseks tõdedele, mis kehtivad kõigis kontrafaktilises võimalikes maailmades — näiteks “Vesi on H₂O”. Mõlemat erikuju võib pidada tõeks tähenduse alusel; lihtsalt tähenduse aspektid, mille alusel nad tõesed on, on erinevad.

​**[lk.63]**​ Mõlemat võib vaadelda ka kui paratamatute tõdede erikujusid. Teine neist vastab paratamatu tõe mõiste tavapärasemale tõlgendusele. Kuid esimest võib tõlgendada ka kui kõigis võimalikes maailmades kehtivat tõde, kui mõista neid võimalikke maailmu lausungikontekstidena või aktuaalse maailma võimalike variantidena. Sellise veidi teistsuguse tõlgenduse korral on lause paratamatult tõene, kui ta osutub tõeseks olenemata sellest, missuguseks osutub aktuaalne maailm[[90]](#footnote-91). Kui aktuaalne maailm osutub maailmaks, kus vesijas ollus on XYZ, siis osutub minu väide “XYZ on vesi” tõeseks. Niisiis, selles tõlgenduses, mis vaatleb võimalikke maailmu kui aktuaalseid maailmu, on “vesi on vesijas ollus” paratamatu tõde.

Evans (1979) nimetab seesugust paratamatust *süvaparatamatuseks* [​*deep necessity*​], vastandades seda *pindparatamatustele* [​*“superficial” necessity*​] nagu “vesi on H₂O”. Davies ja Humberstone (1980) analüüsivad seda paratamatuse liiki üksikasjalikult modaaloperaatori abil, mida nad nimetavad *fikseeritud aktuaalsuseks*​. Erinevalt pindparatamatusest ei mõjuta süvaparatamatust aposterioorsed kaalutlused. Need kaks võimalikkuse ja paratamatuse liiki rakenduvad alati lausungitele. Seevastu maailmade võimalikkus selles mõttes kahte liiki ei jagune; kirjeldatud kaks paratamatuse liiki erinevad ainult lausete tõeväärtuse hindamise viisilt maailmas.

Me võime öelda seda ka teisiti, pannes tähele, et igale väitele vastab kaks tõesustingimuste kogumit. Kui me vaatleme mingis lauses sõnade tähendusi vastavalt nende primaarsetele intensioonidele, saame lause *primaarsed* tõesustingimused — see tähendab, võimalike keskpunktiga maailmade hulga, kus see lause osutub tõeseks, kui hinnata tema tõeväärtust vastavalt temas sisalduvate mõistete primaarsetele intensioonidele. Primaarsed tõesustingimused ütlevad meile, missugune peab olema aktuaalne maailm, et lause oleks selles maailmas tõene, või teisiti öeldes, nad määravad kindlaks kontekstid, milles ta osutub tõeseks. Näiteks lause “vesi on märg” primaarsed tõesustingimused ütlevad ligikaudu seda, et see lause on tõene niisuguste maailmade hulgas, kus vesijas ollus on märg.

Kui me aga vaatleme lauses sisalduvate sõnade tähendusi nende sekundaarsete intensioonide alusel, rakenduvad meile tuttavamad *sekundaarsed tõesustingimused*​. Need määratlevad ühe või väite tõeväärtuse kontrafaktilistes maailmades eeldusel, et aktuaalne maailm on osutunud niisuguseks, nagu ta on. Näiteks (siinses maailmas öeldud) lause “Vesi on märg” sekundaarsed tõesustingimused ütlevad, et lause kehtib maailmades, kus vesi on märg: seega kui eeldada, et vesi on H₂O, määratlevad nad nende maailmade hulga, kus H₂O on märg. Paneme tähele, et aktuaalse maailma tõe asjus siin kahemõttelisuse ohtu ei ole: aktuaalses maailmas saame lausele alati sama tõeväärtuse, ükskõik kas me hindame lauset primaarsete või sekundaarsete tõesustingimuste põhjal.

Kui me vaatleme propositsioone kui funktsioone võimalike maailmade hulgast tõeväärtuste hulka, siis annavad need kaks tõetingimuste kogumit iga lause kohta kaks propositsiooni. Lauses sisalduvate terminite primaarsete intensioonide kombinatsioon annab *primaarpropositsiooni*​, mis on tõene parajasti nendes lausungikontekstides, ​**[lk.64]**​ kus lause osutuks väljendavat mingit tõde. (See on Stalnakeri (1978) “diagonaalpropositsioon”. Rangelt võttes on see keskpunktiga propositsioon, s.t. funktsioon keskpunktiga maailmade hulgast tõeväärtuste hulka). Sekundaarsed intensioonid annavad *sekundaarpropositsiooni*​, mis on tõene niisugustes kontrafaktilistes olukordades, kus lause, lausutuna aktuaalses maailmas, on tõene. Sekundaarpropositsioon on lausungi “sisu” Kaplani mõttes — ja sagedamini peetakse just sekundaarpropositsiooni propositsiooniks, mida üks või teine lause väljendab, kuigi ka primaarpropositsioon on oluline.

Need kaks lausete paratamatu tõesuse liiki on täpses vastavuses kahe neile vastavate propositsioonide paratamatuse liigiga. Esimeses (aprioorses) mõttes on lause paratamatult tõene, kui vastav primaarpropositsioon kehtib kõigis võimalikes keskpunktiga maailmades (see tähendab, kui lause osutuks kõigis lausungikontekstides tõeseks). Lause on paratamatult tõene aposterioorses mõttes, kui vastav sekundaarpropositsioon kehtib kõigis võimalikes maailmades (see tähendab, kui see lause, lausutuna aktuaalses maailmas, on tõene kõigis kontrafaktilistes maailmades). Esimene liik vastab Evansi süvaparatamatusele ja teine meile tuttavamale pindparatamatusele.

Võtame näiteks lause “Vesi on H₂O”. Mõistete “vesi” ja “H₂O” primaarsed intensioonid on erinevad, mistõttu meil ei ole võimalik ​*a priori* teada, et vesi on H₂O; lausega seonduv primaarpropositsioon ei ole paratamatu (kuna ta kehtib ainult niisugustes keskpunktiga maailmades, kus vesijal ollusel on teatud kindel molekulaarne struktuur). Nende mõistete sekundaarsed intensioonid aga ühtivad, nii et lause “vesi on H₂O” on tõene kõigis võimalikes maailmades, kui hinnata tema tõeväärtust sekundaarse intensioonide alusel — see tähendab, temaga seonduv sekundaarpropositsioon on paratamatu. Kripke aposterioorne paratamatus tuleb mängu just siis, kui lause sekundaarsed intensioonid on aluseks mingile paratamatule propositsioonile, primaarsed intensioonid aga mitte.

Nüüd aga vaatleme lauset “Vesi on vesijas ollus”. Siin sõnade “vesi” ja “vesijas ollus” primaarsed intensioonid ühtivad, mistõttu see lause on ​*a priori* tõene, senikaua kui meil on need mõisted olemas​. Meie lausele vastav primaarpropositsioon on paratamatu, ja seega on ta paratamatult tõene Evansi “süvaparatamatuse” mõttes. Samade sõnade sekundaarsed intensioonid on aga erinevad, sest mõistet “vesi” tuleb võtta rigiidses mõttes, mõistet “vesijas ollus” aga mitte: maailmas, kus läbipaistev joogikõlblik vedelik on XYZ, osutab väljendi “vesijas ollus” sekundaarne intensioon XYZ-le, sõna “vesi” oma aga mitte. Lausele vastav sekundaarpropositsioon ei ole seega paratamatu ja lause ei ole paratamatult tõene harjumuspärasemas mõttes; see on näide Kripke “kontingentsest aprioorsusest”.

Üldiselt on paljude neist Kripke ideedest tulenevate näiliste raskuste põhjus selles, et ettekujutust osutusest, mille annab ülalkirjeldatud kahekohaline funktsioon, püütakse suruda ühteainsasse tähenduse või paratamatuse mõistesse. Neist raskustest ülesaamiseks piisab tavaliselt osutuse kahemõõtmelisusega arvestamisest ​**[lk.65]**​ ning kõnealuse tähenduse või paratamatuse liigi hoolikast ja selgesõnalisest määratlemisest.[[91]](#footnote-92)

Selle kahemõõtmelise mudeli abi on võimalik kirjeldada nii mõtlemise kui ka keele semantikat. Ma arutlen selle üle palju põhjalikumalt teisal (Chalmeres 1994c). Mudeli see aspekt ei ole siinkohal oluline, kuid seda tasub mainida, sest paar korda tuleb meil sellega siiski riivamisi kokku puutuda. Peamine idee on väga sarnane: mõistega, mida indiviid oma mõtlemises kasutab, saab seostada primaarse intensiooni vastavalt sellele, millele see mõiste osutab olenevalt sellest, milliseks osutub aktuaalne maailm, ja sekundaarse intensiooni vastavalt sellele, millele ta osutab kontrafaktilistes maailmades eeldusel, et aktuaalne maailm osutub niisuguseks, nagu ta on. Sarnasel viisil saab omistada primaar- ja sekundaarpropositsiooni (mida ma nimetan mujal vastavalt *mõisteliseks sisuks* ja *relatsiooniliseks sisuks*​) ka uskumustele.

Näiteks mõistete “Hesperos” ja “Phosphoros” primaarsed intensioonid on erinevad (üks osutab igas keskpunktiga maailmas Ehatähele, teine Koidutähele), kuid nende sekundaarsed intensioonid ühtivad (mõlemad osutavad kõigis maailmades Veenusele). Mõtte “Hesperos on Phosphoros” primaarpropositsioon on tõene kõigis keskpunktiga maailmades, kus Ehatäht on Koidutäht: asjaolu, et see mõte ei ole triviaalne, vaid informatiivne, vastab asjaolule, et primaarpropositsioon on nende kahe mõiste primaarsete intensioonide erinevuse tõttu kontingentne.

Primaarpropositsioon kajastab asjade seisu subjekti seisukohalt vaadates paremini kui sekundaarpropositsioon: ta paneb paika keskpunktiga maailmade kogumi, mida subjekt seda propositsiooni uskudes tunnistab oma potentsiaalseteks elukeskkondadeks. (Uskudes, et Hesperos on Phosphoros, tunnistan ma sellisteks kõik keskpunktiga maailmad, kus Ehatäht ühtib keskpunkti ümbruses Koidutähega.) Samuti on üsna lihtne näidata, et mõtete vahelisi kognitiivseid ja ratsionaalseid seoseid suunab just primaar-, mitte sekundaarpropositsioon. Seepärast on loomulik vaadelda primaarpropositsiooni kui mõtte kognitiivset sisu.[[92]](#footnote-93)

#### Loogiline paratamatus, kontseptuaalne tõde ja mõeldavus

Sellele mudelile toetudes võime nüüd formuleerida suhted loogilise paratamatuse, kontseptuaalse tõe ja mõeldavuse vahel. Alustame loogilisest paratamatusest. See on lihtsalt paratamatus ülalselgitatud mõttes: lause on loogiliselt paratamatu parajasti siis, kui ta on kõigis loogiliselt võimalikes maailmades tõene. Muidugi on lausete loogilisel paratamatusel kaks kuju olenevalt sellest, kas me hindame lause tõeväärtust võimalikes maailmas primaarsete või sekundaarsete intensioonide alusel. Neid erikujusid võiks nimetada vastavalt *1-paratamatuseks* ja *2-paratamatuseks*​.

​**[lk.66]**​ See analüüs selgitab lause loogilist paratamatust ja võimalikkust a) maailmade loogilise võimalikkuse kaudu ja b) lauses figureerivate terminite intensioonide kaudu. Intensioone ma juba käsitlesin. Mis puutub loogiliselt võimaliku maailma mõistesse, siis see on enam-vähem elementaarmõiste: nagu ennegi, võime intuitiivselt mõelda loogiliselt võimaliku maailma all maailma, mille Jumal oleks võinud luua. (Küsimused Jumala enda kohta jätame siinkohal kõrvale.) Ma ei hakka puudutama hella küsimust nende maailmade ontoloogilisest staatusest, vaid võtan neid lihtsalt kui endastmõistetavat töövahendit, samasugust nagu matemaatika.[[93]](#footnote-94) [[94]](#footnote-95) [[95]](#footnote-96) Selle klassi mahu kohta võib öelda eeskätt seda, et iga maailm, mis on mõeldav, on ka loogiliselt võimalik. (Sellest tõsiasjast tuleb peagi pikemalt juttu.)

Mis puutub kontseptuaalsesse tõesse, siis on kerge luua seos tähendusliku tõe ja loogilise paratamatuse vahel, kui pidada tähendust samaseks intensiooniga (primaarse või sekundaarsega). Kui lause on loogiliselt paratamatu, siis tuleneb tema tõesus automaatselt terminite intensioonidest, millest ta koosneb (ja lause kompositsioonilisest struktuurist). Me ei pea maailma sisse tooma ühtki muud rolli, sest kõnealused intensioonid peavad paika igas võimalikus maailmas. Ja ümberpöördult: kui lause tõesus tuleneb tema intensioonidest, siis on ta igas võimalikus maailmas tõene.

Nagu ennegi, võib kontseptuaalsel tõel eristada kaht kuju olenevalt sellest, kas me samastame “tähendust” primaarse või sekundaarse intensiooniga. Need kaks kuju on paralleelsed paratamatu tõe kahe kujuga. Senikaua kui me langetame otsuseid mõlemal juhul paralleelselt, on väide kontseptuaalselt tõene parajasti siis, kui ta on paratamatult tõene. Lause “vesi on vesijas ollus” on kontseptuaalselt tõene ja paratamatult tõene primaarsete intensioonide alusel, ja lause “vesi on H₂O” on kontseptuaalselt tõene ja paratamatult tõene sekundaarsete intensioonide alusel.[[96]](#footnote-97) *A priori* on üldiselt võimalik jõuda ainult esimest liiki kontseptuaalste tõdedeni. Teise liigi alla kuuluvad paljud aposterioorsed tõed, sest sekundaarne intensioon sõltub sellest, missuguseks osutub aktuaalne maailm.

(Ma ei ütle, et ainuõige viis tähendustest mõtlemiseks käib intensioonide kaudu. Tähendus on paljutahuline mõiste, mille mõningaid tahke intensioonid ei tarvitse täiuslikult kajastada, mistõttu tähenduse ja intensiooni samastamisele võidakse ka vastu vaielda, vähemalt mõnel juhul.[[97]](#footnote-98) Pigem tuleks tähenduse ja intensiooni võrdsustamist siinkohal võtta kui kokkulepet: postuleerides, et nad omavahel ühtivad, võib leida mitmesuguseid kasulikke seoseid. Sõna “tähendus” tarvitusviisist ei olene palju. Igatahes on intensioonipõhine tõesus ainus tõesusele toetuva tähenduse liik, mida minul siin tarvis läheb.)

Küllalt ettevaatlikult talitades võime leida ka seose lausete loogilise võimalikkuse ja nende mõeldavuse vahel. Võime öelda, et lause on *mõeldav*​ (või *mõeldavalt tõene*​), kui leidub mõeldav maailm, milles ta on tõene. Seda ütlemisviisi ei tohiks segi ajada “mõeldavuse” teiste tähendustega. Näiteks on olemas ka tähendus, mille järgi lause on mõeldav, kui ta on meie teada tõene või ei ole meie teada võimatu. Selles mõttes on ​**[lk.67]**​ nii Goldbachi konjektuur kui ka selle eitus mõeldavad. Kuid üks neist lausetest, see, kumb on faktiliselt väär, ei ole mõeldav minu siinses mõttes, sest ei ole mõeldavat maailma, kus ta oleks tõene, või teisiti öeldes, ta on kõigis maailmades väär.

Mõeldavuse niisuguses käsitluses tähendab lause mõeldavus kaht asja: esiteks mingi vastava maailma mõeldavust, teiseks lause tõesust selles maailmas.[[98]](#footnote-99) See tähendab, et mõeldavusotsustusi tehes tuleb olla kindel, et me oleme maailma, mida me mõeldavaks peame, õigesti kirjeldanud; selleks tuleb õigesti hinnata väite tõeväärtust selles maailmas. Esimesel pilgul võiks tunduda, et Goldbachi konjektuuri väärus on mõeldav, kui pidada mõeldavaks maailma, kus matemaatikud kuulutavad ta vääraks; aga kui Goldbachi konjektuur on tegelikult tõene, siis on selle maailma kirjeldus ebaõige: tegelikult on Goldbachi konjektuur ka selles maailmas tõene ja matemaatikud eksivad.

Praktikas tarvitseb mõeldavusotsustuse tegemiseks vaid vaadelda mingit mõeldavat olukorda — väikest osa mingist maailmast — ning seejärel hoolitseda selle eest, et meil oleks selle olukorra õige kirjeldus. Kui leidub mõeldav olukord, kus mingi lause on tõene, siis leidub ilmselt ka mõeldav maailm, kus see lause on tõene; seega annab see meetod mõttekaid tulemusi, ilma et meil tuleks oma kognitiivseid ressursse üle koormata vajadusega kujutada ette tervet maailma!

Mõnikord öeldakse, et laused “vesi on XYZ” ja teised sarnased näited tõestavad, et mõeldavusest ei saa järeldada võimalikkust; kuid minule tundub, et asi pole nii lihtne. Tegelikult on mõeldavusel kaks erikuju, mida võiks nimetada *1-mõeldavuseks*​ ja *2-mõeldavuseks*​, olenevalt sellest, kas me hindame lause tõeväärtust mingis mõeldavas maailmas asjaomaste terminite primaarsete või sekundaarsete intensioonide alusel. Lause “vesi on XYZ” on 1-mõeldav, sest leidub mõeldav maailm, kus see lause (kui hinnata tema tõeväärtust primaarsete intensioonide alusel) on tõene; kuid ta ei ole 2-mõeldav, sest ta (kui hinnata tema tõeväärtust sekundaarsete intensioonide alusel) ei ole tõene üheski mõeldavas maailmas. Need kaks mõeldavuse liiki on kahe ülalkirjeldatud loogilise võimalikkuse liigi täpsed paarikud.

Sageli samastatakse lausete mõeldavust 1-mõeldavusega (s.t. mõeldavusega selles mõttes, nagu lause “vesi on XYZ” on mõeldav), sest just seesuguse mõeldavuseni on võimalik jõuda ​*a priori*​. Seevastu lause võimalikkust samastatakse enamasti 2-võimalikkusega (s.t. mõeldavusega selles mõttes, nagu lause “vesi on XYZ” on võimatu). Kui mõista mõeldavust ja võimalikkust nii, siis ei saa mõeldavusest järeldada võimalikkust. Kuid sellegipoolest tuleneb 1-mõeldavusest 1-võimalikkus ja 2-mõeldavusest 2-võimalikkus. Tuleb lihtsalt olla ettevaatlik ja mitte kõnelda 1-mõeldavusest, kui jutt on 2-võimalikkusest. Teisiti öeldes, tuleb olla ettevaatlik ja mitte kirjeldada maailma, mida vaadeldakse mõeldavana (näiteks XYZ maailma), primaarsete intensioonide alusel seal, kus sekundaarsed intensioonid oleksid kohasemad.[[99]](#footnote-100)

Kõigest sellest järeldub, et sagedane vahetegemine “loogilise” ja “metafüüsilise” võimalikkuse vahel, mis ulatub tagasi Kripke näideteni — millede järgi ​**[lk.68]**​ on loogiliselt, kuid mitte metafüüsiliselt võimalik, et vesi oleks XYZ, — ei ole eristus maailmade, vaid parimal juhul lausete tasemel. Selles mõttes on lause loogiliselt võimalik, kui leidub maailm, milles ta on tõene primaarsete intensioonide alusel, ja metafüüsiliselt võimalik, kui leidub maailm, milles ta on tõene sekundaarsetete intensioonide alusel; asjaomane maailmade ruum on mõlemal juhul sama.[[100]](#footnote-101)

Peaasi on siin see, et ükski vaadeldud juhtumitest ei anna alust pidada võimatuks ühtki mõeldavat maailma. Kõik meie mured mõeldavuse ja võimalikkuse vahelise lünga pärast puudutavad lauseid, mitte maailmu: lause mõeldava maailma kohta kas kirjeldab seda maailma valesti (nagu Kripke juhul ja teisel Goldbachi juhul), või siis me väidame, et lause on mõeldav, ilma et me ühelegi maailmale mõtleksimegi (nagu esimesel Goldbachi juhul). Seepärast pole nähtavasti põhjust eitada, et maailma mõeldavusest tuleneb tema võimalikkus. Siitpeale lähtun ma eeldusest, et see väide on loogiliselt võimalik; iga võimalikkuse erikuju, mille puhul mõeldavusest ei saa järeldada võimalikkust, on sellisel juhul väiksema mahuga klass. Lugeja otsustab võib-olla, et on olemas “metafüüsiliselt võimalike maailmade” klass väiksema mahuga; kuid põhjused, miks ta niisuguse klassi olemasolusse usub, peaksid siis olema täiesti sõltumatud neist tavalistest põhjustest, mida ma siin vaatlesin. Igal juhul on seletuslikes küsimustes kesksel kohal just loogiline võimalikkus. (Tugevam “metafüüsiline” modaalsus sobib parimal juhul ehk ontoloogia, materialismi jms. üle arutlemiseks; sellest tuleb juttu [sobival kohal 4. peatükis](#ORG.4.2.2).)

Vastupidises suunas — loogiliselt võimalikkuselt mõeldavusele — pole järelduskäik nii lihtne, sest kuna meie kognitiivsel võimekusel on piirid, siis võib juhtuda, et mõned olukorrad, mis on võimalikud, ei ole meie jaoks mõeldavad oma keerukuse tõttu. Kui aga mõista mõeldavust kui *põhimõttelist mõeldavust* — mõeldavust võib-olla mingi ülivõimeka olendi jaoks —, siis on usutav, et ühe või teise maailma loogilisest võimalikkusest saab järeldada selle maailma mõeldavuse ning et lause loogilisest võimalikkusest tuleneb seega tema (1- või 2-)[[101]](#footnote-102)mõeldavus​. Igal juhul huvitab neist kahest järelduskäigust mind allpool rohkem teine.

Kui lause on loogiliselt võimalik või paratamatu oma primaarsete intensioonide alusel, siis on seda võimalikkust või paratamatust võimalik ​*a priori*​ teada, vähemasti põhimõtteliselt. Modaalsus ei ole episteemiliselt kättesaamatu: lause võimalikkus on tema koosseisu kuuluvate sõnade intensioonide ja võimalike maailmade ruumi funktsioon; põhimõtteliselt on nad mõlemad meile episteemiliselt kättesaadavad, ja kumbki neist ei sõltu[[102]](#footnote-103) aposterioorsetest faktidest. Seega on 1-võimalikkuse ja 1-mõeldavuse üle põhimõtteliselt võimalik otsustada tugitoolist tõusmata. Seevastu 2-võimalikkust ja 2-mõeldavust on paljudel juhtudel võimalik sedastada ainult *a posteriori*​, kuna sekundaarsete intensioonide tuvastamisel võivad mängida rolli välise maailma faktid.

​**[lk.69]**​ 1-paratamatute tõdede klass kattub täpselt aprioorsete tõdede klassiga. Kui lause on ​*a priori*​ tõene, siis on ta tõene olenemata sellest, missuguseks osutub aktuaalne maailm; see tähendab, ta on tõene kõigis maailmades, mida me käsitleme aktuaalsetena, ning on seega 1-paratamatu. Ja ümberpöördult, kui lause on 1-paratamatu, siis on ta tõene, ükskõik missuguseks osutub aktuaalne maailm, ja seega on ta tõene *a priori*​. Enamikul sellistest juhtudest on meil võimalik teada väite tõeväärtust ​*a priori*​; eranditeks võivad olla mõned matemaatilised väited, mille tõeväärtust me ei oska hinnata, ja mõned laused, mis on nii keerulised, et me ei suuda neid mõista. Kuid nendelgi juhtudel on vististi mõistlik öelda, et neid on vähemalt põhimõtteliselt võimalik teada *a priori*​, kuigi nad ületavad meie kognitiivse võimekuse piirid. (Selle küsimuse juurde tulen ma *sobivas kohas allpool* veel tagasi.)

#### Loogiline paratamatus ja loogiline superventsus

Meie käsutuses on kaks pisut erinevat loogilise superventsuse mõistet olenevalt sellest, kas me toetume loogilise paratamatuse primaarsele või sekundaarsele kujule. Kui sõnaga “sodi” seondub nii primaarne kui ka sekundaarne intensioon, siis võib sodisus olla füüsikaliste omaduste suhtes loogiliselt superventne sõna “sodi” primaarse või sekundaarse intensiooni alusel. Superventsus sekundaarse intensiooni alusel — s.t. superventsus, millele vastavaks modaalsuseks on aposterioorne paratamatus — vastab sellele, mida mõned on nimetanud metafüüsiliseks superventsuseks, ent mida võib, [nagu juba nägime](#PARA.46-11)​, pidada üheks loogilise superventsuse erikujuks.

(Tegelikult on omaduste loogilisel superventsusel ainult üks liik nagu ka propositsioonide loogilisel paratamatusel on ainult üks liik. Kuid [me oleme näinud](#PARA.45-26)​, et iga sõna või mõiste määratleb tegelikult kaks omadust — ühe omaduse tema primaarse ja ühe tema sekundaarse intensiooni kaudu (näiteks vastavalt “vesijas ollus” ja “H₂O”). Seega võivad iga mõistega (näiteks mõistega “vesi”) seonduvad omadused olla superventsed kahel viisil. Mõnikord kõnelen ma ka mitterangelt ka omadusega seonduvast primaarsest ja sekundaarsest intensioonist ning omaduste kahest võimalikust superventsuse liigist.)

Allpool vaatlen ma loogilise superventsuse mõlemat erikuju, nii primaarset kui ka sekundaarset, konkreetsetel juhtudel. Kuid esiplaanil hakkab olema esimene kuju. Primaarsed intensioonid on olulisemad kui sekundaarsed, eriti kui jutt on seletamisest. Nagu juba tähendatud​, on meil uurimist alustades käepärast ainult primaarne intensioon, ja just see intensioon määrab, kas seletus on rahuldav või mitte. Näiteks selleks, et seletada vett, tuleb seletada tema läbipaistvust, vedelust jne. Sekundaarne intensioon (“H₂O”) tuleb mängu alles siis, kui seletamine on lõpule jõudnud, ja seetõttu ei saa ta üksinda anda meile seletuse korrektsuse kriteeriumi. Just loogilisest superventsusest primaarse intensiooni alusel oleneb, kas reduktiivne seletamine on võimalik. Allpool ​**[lk.70]**​ kõnelen ma, kui pole öeldud teisiti, loogilisest superventsusest kõikjal primaarse intensiooni alusel.

Valides ühe intensiooni liigi, näiteks primaarse, ja jäädes selle juurde pidama, näeme, et loogilise superventsuse mitmesugused võimalikud formuleeringud on samaväärsed. Vastavalt [käesoleva peatüki alguses](#PARA.34-2) esitatud definitsioonile on B-omadused A-omaduste suhtes loogiliselt superventsed, kui iga loogiliselt võimaliku olukorra *Y* puhul, mis on A-eristamatu mingist aktuaalsest olukorrast ​*X*​, kehtivad kõik B-faktid, mis kehtivad olukorras ​*X*​, ka olukorras ​*Y*​. Või lihtsamalt öeldes, B-omadused on A-omaduste suhtes loogiliselt superventsed, kui iga aktuaalse olukorra *X* A-faktidest tulenevad selle olukorra B-faktid. (“Q tuleneb P-st” tähendab siin “on loogiliselt võimatu, et P ja mitte Q”.)

Kui jääda pidama globaalse superventsuse juurde, siis tähendab see, et B-omadused on A-faktide suhtes loogiliselt superventsed, kui aktuaalse maailma B-faktid tulenevad tema A-faktidest. Analoogiliselt on B-omadused loogiliselt superventsed A-omaduste suhtes, kui ei leidu mõeldavat maailma, mille A-omadused ühtiksid meie maailma omadega, B-omadused aga mitte. Me võime ka öelda, et loogiline superventsus kehtib, kui kõigi meie maailma *W* A-faktide ühendi ​*A\** ja suvalise meie maailma B-fakti *B* puhul tuleneb seose ​*A\**​(​*W*​) → ​*B*​(​*W*​) tõesus A-terminite ja B-terminite tähendustest (millede all tuleb siin mõista intensioone).

Ning kui B-omadused on A-omaduste suhtes loogiliselt superventsed primaarsete intensioonide alusel, siis on B-omadused tuletatavad A-omadustest ​*a priori*​. See tähendab, et kui keegi teab mingi aktuaalse olukorra kõiki A-fakte, siis on tal ainuüksi nende põhjal võimalik teha kindlaks selle olukorra B-faktid, kui ta valdab kõnealuseid B-mõisteid. Praktikas võib see tuletamine vaadeldavate olukordade keerukuse tõttu olla keeruline või võimatu, kuid vähemalt põhimõtteliselt on see võimalik. Kui aga võtta loogilise superventsuse aluseks sekundaarsed intensioonid, siis on olukorra B-fakte samuti põhimõtteliselt võimalik tuletada A-faktidest, kuid alles *a posteriori*​. A-faktidele tuleb lisada kontingentsed faktid aktuaalse maailma kohta, sest viimaseid on tarvis vastavate B-intensioonide määratlemiseks.

Seega saab ütlusi loogilise superventsuse kohta põhjendada vähemalt kolmel viisil — mõeldavuse kaudu, epistemoloogiliselt ja analüüsi teel. See tähendab, et kinnitamaks B-omaduste loogilist superventsust A-omaduste suhtes, tuleb kas (1) näidata, et A-omaduste realiseerumine on mõeldamatu, ilma et realiseeruksid ka B-omadused; (2) näidata, et see, kes valdab A-fakte, võib saada teada ka B-faktid (vähemasti kui jutt on superventsusest primaarsete intensioonide alusel); või (3) B-omaduste intensioonide piisavalt üksikasjaliku analüüsi teel näidata, et B-laused tulenevad A-lausetest ainuüksi oma intensioonide varal. Needsamad viisid sobivad ka loogilise superventsuse puudumise tuvastamiseks. Põhjendades allpool oma väiteid loogilise superventsuse kohta, rakendan ma kõiki kolme meetodit.

​**[lk.71]**​ Kuna mitte igaüks ei pruugi loogilise superventsuse erinevate formuleeringute samaväärsuses nii kindel olla, siis esitan ma allpool loogilist superventsust puudutavaid olulisi järeldusi põhjendades oma argumendid mitmel kujul, lastes käiku kõik need formuleeringud. Nii võib lugeja näha, et minu argumentidel on jõudu ning et neis pole midagi, mis sõltuks peentest erinevustest superventsuse mitmesuguste käsituste vahel.

### 5. Peaaegu kõik on füüsika suhtes loogiliselt superventne\*

Järgnevas peatükis püüan näidata, et teadvuselamus ei ole füüsika suhtes loogiliselt superventne, mistõttu teda ei ole võimalik reduktiivselt seletada. Sageli vastatakse säärastele väidetele, et teadvuselamus ei ole selles suhtes üksi ning et omadusi, mis ei ole füüsikaliste omaduste suhtes loogiliselt superventsed, on väga mitmesuguseid. Näiteks ei olevat säärastel omavahel nii erinevatel omadustel nagu laud-olemine[[103]](#footnote-104), elusus ja majanduslik heaolu mingit loogilist seost faktidega, mis puudutavad aatomeid, elektromagnetvälju jne. Ei saavat ju need kõrgtaseme faktid loogiliselt tuleneda mikrofüüsikalistest faktidest?

Minu arvates selgub hoolikal analüüsil õige pea, et see on ei ole nii ja et nimetatud kõrgtaseme faktid on füüsika suhtes (globaalselt) loogiliselt superventsed, niivõrd kui nad üldse on faktid.[[104]](#footnote-105) Seega on teadvuselamus loogilise superventsuse puudumise poolest peaaegu unikaalne nähtus. Seos teadvuse ja füüsikaliste faktide vahel on oma olemuselt teistsugune kui tavalised seosed kõrgtaseme ja madaltaseme faktide vahel.

Enamiku omaduste loogilist superventsust füüsikaliste omaduste suhtes on võimalik tuvastada mitmel viisil. Siinkohal huvitavad mind ainult omadused, mis iseloomustavad *looduslikke fenomene* — s.t. maailma kontingentseid aspekte, mis nõuavad seletust. Omadus olla ingel ei pruugi olla loogiliselt superventne füüsika suhtes, ent kuna meil pole põhjust inglitesse uskuda, siis ei tohiks superventsuse puudumine meile siin muret teha. Samuti ei huvita mind faktid abstraktsete entiteetide kohta, nagu seda on matemaatilised objektid või propositsioonid, sest neid tuleb käsitleda eraldi.[[105]](#footnote-106)

Tuleb märkida, et kui ma ütlen, et enamik kõrgtaseme omadusi on füüsikaliste omaduste suhtes superventsed, ei taha ma väita, nagu tuleneksid kõrgtaseme faktid ja seadused mikrofüüsika seadustest, olgu üksi või ühenduses mikrofüüsikaliste rajatingimustega. See oleks liiga tugev väide, ja kuigi tal võib sobivas tõlgenduses mõningast usutavustki olla, pole tema kasuks esialgu tõendeid. Minu väide on palju nõrgem: ma ütlen ainult, et kõik kõrgtaseme faktid tulenevad mikrofüüsikalistest faktidest (võib-olla ühenduses mikrofüüsika seadustega). Sellesse tohutult mahukasse hulka kuuluvad faktid ​**[lk.72]**​ iga viimase kui osakese ja välja jaotuse kohta igas viimaseski aegruumi nurgakeses — aatomitest Napoleoni kübaras elektromagnetväljadeni Saturni välimises rõngas. Niipea kui see faktide hulk on fikseeritud, ei jää kõigele muule, nagu allpool näeme, varieerumiseks enam kuigi palju ruumi.

Enne argumentide juurde asumist mainiksin mõningaid põhjusi, miks loogiline superventsus füüsika suhtes mõnel juhul ei kehti. Esiteks, mõned kõrgtaseme omadused ei ole loogiliselt superventsed sellepärast, et nad sõltuvad teadvuselamusest. Võib-olla konstitueerib teadvuselamus mingis osas näiteks armastust. Fenomenaalsetest kvaliteetidest võivad, nagu allpool näeme, sõltuda ka mõnede väliste omadustega, näiteks värvuse või kuumusega kaasnevad primaarsed (kuigi mitte sekundaarsed) intensioonid. Sel juhul ei ole armastus ja võib-olla ka kuumus füüsika suhtes loogiliselt superventsed. Neid ei tuleks aga pidada minu väidet kummutavateks vastunäideteks, sest nad ei too kaasa ühtki uut lünka loogilises superventsuses. Minu väidet ennast on vahest kõige parem sõnastada nii: kõik faktid on loogiliselt superventsed füüsikaliste faktide ja fenomenaalsete faktide ühendi suhtes, või: kõik faktid on loogiliselt superventsed füüsikaliste faktide suhtes, kui teadvuselamus arvestusest välja jätta.[[106]](#footnote-107) Analoogiliselt võib sõltuvus teadvuselamusest olla takistuseks mõnede kõrgtaseme fenomenide reduktiivsele seletatavusele, kuid sellegipoolest võib öelda, et need fenomenid on reduktiivselt seletatavad, kui mitte arvestada teadvuselamust.

Teiseks kuulub mõnede primaarsete — kuigi mitte sekundaarsete — intensioonide juurde, nagu [ülal](#PARA.45-21) nägime, vältimatult indeksikaalne element. Näiteks sõna “vesi” primaarne intensioon on umbkaudu “läbipaistev, joogikõlblik vedelik meie keskkonnas”, nii et kui aktuaalses universumis on olemas vesijas H₂O ja vesijas XYZ, siis oleneb vastus küsimusele, kumba neist tuleb nimetada “veeks”, sellest, missugune on seda sõna lausuva agendi keskkond. Seepärast tuleb superventsusbaasile mõnel juhul lisada agendi asukohale vastav keskpunkt. See annab meile loogilise superventsuse ja reduktiivse seletuse, kui mitte arvestada teadvuselamuse ja indeksikaalsuse panust.

Ja viimaseks: loogilisele superventsusele ei räägi vastu juhud, kus kõrgtaseme faktid on määramatud. Ma väidan ainult niipalju, et kuivõrd kõrgtaseme faktid on deterministlikud, determineerivad neid füüsikalised faktid. Kui kogu maailmastki ei piisa kõrgtaseme faktide fikseerimiseks, siis ei maksa oodata, et selleks piisab füüsikalistest faktidest. Võidakse öelda, et loogiline superventsus ei kehtiks, kui leiduks kaks võrdselt head kõrgtaseme teooriat maailma kohta, mis erineksid kõrgtaseme faktide kirjelduselt. Näiteks võivad viirused olla ühe teooria järgi elusolendid, teise järgi aga mitte, ja seepärast ei determineeri füüsikalised faktid bioloogilise elu fakte. See ei ole aga vastunäide, vaid olukord, kus elu puudutavad faktid on määramatud. Sellise määramatuse korral võib sõnatarvituse fikseerida nii või teisiti, nagu parajasti mugavam on. Kui faktid on selgepiirilised — näiteks kui väide, et viirused on elusolendid, on tõene —, siis on üks neist kirjeldustest lihtsalt väär. Nii või teisiti ​**[lk.73]**​ tulenevad olukorra faktid, kuivõrd nad üldse kindlapiirilised on, füüsikalistest faktidest.

Järgnevas hakkan ma tooma loogilise superventsuse kõikehõlmavuse põhjendamiseks argumente, mis apelleerivad mõeldavusele, epistemoloogilistele kaalutlustele ja asjaomaste mõistete analüüsile.

##### Mõeldavus.

Enamiku kõrgtaseme faktide loogilist superventsust on kõige lihtsam näha, kui võtta loogilise võimalikkuse testiks mõeldavus. Missugune peaks olema maailm, et olla iga viimsegi mikrofüüsikalise fakti poolest meie omaga identne, kuid bioloogiliselt sellest erinev? Ütleme näiteks, et meie maailmas on vombat, kellel on kaks poega. Meie maailma füüsikaliste faktide hulka kuuluvad faktid kõigi osakeste jaotuse kohta sellele vombatile vastavas aegruumipiirkonnas, samuti faktid tema poegade ning nende keskkonna ja evolutsiooniloo kohta. Milles võiks seisneda meie maailma erinevus maailmast, mille füüsikalised faktid ühtiksid meie maailma omadega, kuid sellel vombatil ei oleks kaht poega? Selline maailm on täiesti mõeldamatu. Fikseerides mingi võimaliku maailma niisugusena, et kõik selle füüsikalised faktid oleksid samad, oleme automaatselt fikseerinud ka faktid vombatsuse ja järgluse kohta. Need bioloogilised faktid ei ole midagi niisugust, mida saaks lahutada oma füüsilisest alusest — see pole võimalik isegi kontseptuaalselt.

Sama kehtib ka arhitektuuriliste faktide, astronoomiliste faktide, käitumisfaktide, keemiliste faktide, majanduslike faktide, meteoroloogiliste faktide, sotsioloogiliste faktide jne. kohta. Maailm, mis oleks meie maailmaga füüsikaliselt identne, ent kus need faktid oleksid erinevad, on mõeldamatu. Kujutleda mikrofüüsikaliselt identset maailma tähendab kujutleda maailma, kus iga viimsegi osakese asukoht aegruumis on sama. Sellest järeldub, et see maailm ühtib meie maailmaga ka makroskoopiliselt struktuurilt ja makroskoopiliselt dünaamikalt. Niipea kui see kõik on paigas, ei jää sääraste faktide varieerumiseks lihtsalt enam ruumi (kui vahest mitte arvestada erinevusi, mis tulenevad teadvuselamuste erinevustest).

Ning see mõeldamatus ei tulene nähtavasti mitte meie kognitiivse võimekuse kontingentsetest kitsendustest, vaid säärane maailm on mõeldamatu põhimõtteliselt. Isegi ülivõimekas olend, isegi Jumal ei suudaks sellist maailma endale ette kujutada. Selles lihtsalt poleks midagi ette kujutada. Niipea kui nad kujutlevad maailma koos kõigi füüsikaliste faktidega, kujutlevad nad automaatselt maailma, kus kehtivad ka kõik kõrgtaseme faktid. Füüsikaliselt identne maailm, kus kõrgtaseme faktid ei kehtiks, on seega loogiliselt võimatu, ja kõnealused kõrgtaseme omadused on füüsika suhtes loogiliselt superventsed.

##### Epistemoloogia.

Mõeldavust puudutavate intuitiivsete arusaamade juurest edasi liikudes võime tähendada, et võimalik maailm, mis oleks meie omaga füüsikaliselt identne, kuid erineks sellest bioloogiliselt, seaks meie ette radikaalseid epistemoloogilisi probleeme. Kuidas ​**[lk.74]**​ teha kindlaks, et maailm, milles me viibime, on meie maailm, aga mitte too teine? Kuidas teha kindlaks, et bioloogilised faktid meie maailmas on niisugused, nagu nad on? Et asja selgemini mõista, paneme tähele, et kui ma oleksin tolles alternatiivses maailmas, näeksin ma seda kindlasti täpselt samasugusena nagu meie maailm. Temas realiseeruksid samasugused osakeste jaotused nagu need, mida võib täheldada meie maailma taimedes ja loomades; täpselt samasugused oleksid neilt entiteetidelt peegelduvate footonite konfiguratsioonid; kõige hoolikamalgi uurimisel ei ilmneks ainsatki erinevust. Sellest järeldub, et kõigist välistest tõenditest, mis meil võtta on, ei piisa nende maailmade eristamiseks. Kuivõrd meie maailma bioloogilised faktid ei ole loogiliselt superventsed, ei ole meil mitte mingil viisil võimalik neist välise tõendusmaterjali põhjal teadmisi saada.

Tegelikkuses aga ei seondu bioloogiaga ühtki sügavat epistemoloogilist probleemi. Oma maailma bioloogiliste faktide kohta saame me teadmisi väliste tõendite põhjal kogu aeg, ilma et kerkiks erilisi skeptilisi probleeme. Sellest järeldub, et bioloogilised faktid on füüsika suhtes loogiliselt superventsed. Sama kehtib ka arhitektuuri, majanduse ja meteoroloogia faktide kohta: nende teadasaamisest välise tõendusmaterjali põhjal ei tulene mingeid erilist skeptilisi probleeme, ja seega peavad need faktid olema füüsika suhtes loogiliselt superventsed.

Selle teesi kinnituseks võib tähendada, et aladel, kus epistemoloogilisi raskusi tõepoolest ilmneb, kaasneb nendega loogilise superventsuse puudumine, ja ümberpöördult: aladel, kus loogilist superventsust ei ole, toob selle puudumine kaasa epistemoloogilisi raskusi.

Kõige ilmsemaks näiteks on üks teadvust puudutav epistemoloogiline probleem — *teiste vaimude probleem*​, mis seisneb selles, et nii teadvuse olemasolu meid ümbritsevatel olenditel kui ka selle puudumine tunduvad olevat ühtviisi loogiliselt ühitatavad kõigi väliste tõenditega. Meil ei ole võimalik piiluda näiteks koera ajju ja nentida teadvuselamuse olemasolu või puudumist. Selle probleemi staatus on vaieldav, kuid juba ainuüksi tema *prima facie*​ olemasolust piisab, et kummutada ülaltoodud argumentidega paralleelne epistemoloogiline argument teadvuse loogilise superventsuse kasuks. Seevastu “teiste bioloogiate probleemi” või “teiste majanduste probleemi” ei ole olemas isegi *prima facie*​. Need faktid on kõigile otseselt kättesaadavad just sellepärast, et nad on fikseeritud, niipea kui fikseerida füüsikalised faktid.

(Küsimus: miks ei sunni analoogiline argument meid järeldama, et kui teadvuselamus ei ole loogiliselt superventne, siis ei ole meil võimalik midagi teada omaendagi teadvusest? Vastus: sellepärast, et teadvuselamus on otse meie episteemilise universumi keskmes. Skeptilised probleemid seoses mittesuperventsete bioloogiliste faktidega tulevad sellest, et me pääseme bioloogilistele faktidele ligi ainult füüsilisel teel, väliste tõendite kaudu; välised mittesuperventsed faktid jääksid teispoole meie otsese episteemilise vaatevälja piire. Omaenda teadvusega see nii ei ole.)

Teine kuulus epistemoloogiline probleem puudutab põhjuslikkuse fakte. Nagu ütles Hume, annavad välised tõendid meile ainult reeglipärasid ​**[lk.75]**​ sündmuste järgnevuses, aga mitte mingeid põhjuslikkuse fakte veel nendele lisaks. See tähendab, et kui mõista põhjuslikkust millegi enamana kui reeglipärasus (nagu minu arvates ongi õige), siis ei ole selge, kas meil on võimalik teada, et põhjuslikkus üldse eksisteerib. Ka see skeptiline probleem käib käsikäes loogilise superventsuse puudumisega: kõnealusel juhul ei ole põhjuslikkuse faktid konkreetsete füüsikaliste faktide suhtes loogiliselt superventsed. Isegi kui meie valduses oleksid kõik faktid füüsikaliste entiteetide jaotuse kohta aegruumis, oleks ikkagi loogiliselt võimalik, et kõik need reeglipärad on hiiglaslik kosmiline juhus, mille taga pole mingit reaalset põhjuslikkust. Lähemas plaanis näeme, et iga juhtumi puhul, mille konkreetsetes faktides tundub avalduvat põhjuslikkus, on loogiliselt võimalik pidada seda pelgaks järgnevuseks. Me järeldame põhjusliku seose olemasolu abduktsiooni teel — kuna vastasel korral peaksime uskuma üüratutesse seletamatutesse kokkusattumustesse —, kuid usk põhjuslikkusesse ei sunni end meile peale nii otseselt nagu usk bioloogiasse.

Et põhjuslikkuse superventsusega seonduvatest probleemidest ümbert ringi minna, olen ma postuleerinud, et siinses käsitluses hõlmab superventsusbaas lisaks konkreetsetele füüsikalistele faktidele ka kõiki füüsikaseadusi. On mõistlik oletada, et seaduste lisamine fikseerib põhjuslikkuse faktid. Kuid nagu varem põhjuslikkusega, kerkib nüüd seoses seadustega jällegi skeptiline probleem — meenutagem Hume’i induktsiooniprobleemi ning loogilist võimalust, et kõik, mis näib seadusena, võib olla vaid juhuslik reeglipära.

Niipalju kui mina öelda oskan, piirduvad epistemoloogilised probleemid, mis tulenevad loogilise superventsuse puudumisest füüsika suhtes, nende kahega. Veel paar epistemoloogilist probleemi käivad mingis mõttes nende eel, kuna nad puudutavad füüsikaliste faktide eneste olemasolu. Esimene neist on välismaailma olemasolu kartesiaanlik probleem. Meie elamuslike andmetega ei ole vastuolus võimalus, et maailma, mida me arvame nägevat, ei ole olemas; võib-olla on see meie hallutsinatsioon, võib-olla me oleme vaadiajud. Selle probleemi allikat võib näha just nimelt selles, et välismaailma faktid ei ole loogiliselt superventsed meie elamuse faktide suhtes. (Idealistid, positivistid ja teised on püüdnud vaieldavate argumentidega põhjendada, et selline superventsus on olemas. Paneme tähele, et kui nende vaadet aktsepteerida, langeb see skeptiline probleem ära.) Teine epistemoloogiline probleem puudutab teaduses postuleeritud teoreetilisi entiteete — elektrone, kvarke jms. Nende mitteolemasolu ei oleks loogiliselt vastuolus otseselt vaadeldavate faktidega objektide kohta meie ümbruses, ja seepärast on nende olemasolu kohta väljendatud skeptilisi kahtlusi. Nagu näitab analüüs, tuleb see probleem sellest, et teoreetilised faktid ei ole vaatlusfaktide suhtes loogiliselt superventsed. Nii ühel kui teisel juhul ei ole skeptiliste kahtluste hajutamiseks arvatavasti paremat võimalust kui abduktsioon, just nagu põhjuslikkuse puhulgi. Kuid põhimõtteline võimalus, et me eksime, sellest ei kao.

Igatahes lähen ma sedasorti skeptilistest probleemidest ümbert ringi eeldusega, et füüsiline maailm on lihtviisiliselt olemas, ning fikseerin superventsusbaasis kõik maailma ​**[lk.76]**​ füüsikalised faktid (eeldades seega, et väline maailm on olemas ning et olemas on ka elektronid jne.). Kui need faktid on teada, ei jää enam ruumi skeptilisteks kahtlusteks enamikus kõrgtaseme faktides, ja seda just sellepärast, et nad on loogiliselt superventsed. Või mõttekäiku teisipidi pöörates: kuna kõik meie väliste andmete allikad on mikrofüüsikaliste faktide suhtes loogiliselt superventsed, siis ei saa välised tõendid anda meile põhjust uskuda fenomenidesse, mis nende faktide suhtes superventsed ei ole. Võiks küsida, kas mikrofüüsikaliste faktide seletamiseks võiks ehk abduktsiooni teel postuleerida veel mingeid teisi fenomene nagu ülal​[[107]](#footnote-108). Tõepoolest, see protsess kulgeb konkreetsetelt faktidelt lihtsate seadusteni, mis on nende aluseks (näidates seega kätte põhjuslikud seosed), kuid seejärel nähtavasti peatub. Fundamentaalsete seaduste olemusse kuulubki see, et nad on seletusahela viimane lüli. (Edasi võivad tulla veel ehk teoloogilised spekulatsioonid.) Järele jäävad fenomenid, millede kohta meil on seesmised tõendid — nimelt teadvuselamus —, ja see ongi kõik. Kui teadvuselamust mitte arvestada, on kõik fenomenid füüsika suhtes loogiliselt superventsed.

Epistemoloogilist argumenti loogilise superventsuse kasuks saab esitada ka sirgjoonelisemalt, öeldes, et see, kes valdaks kõiki füüsikalisi fakte, võiks põhimõtteliselt saada teada kõik kõrgtaseme faktid, kui ta valdab asjaomaseid kõrgtaseme mõisteid. Tõsi küll, praktikas ei suudaks ta mikrofüüsikaliste faktide kogumist lähtudes iialgi kõrgtaseme faktideni jõuda. Viimaste kogumi pöörasest suurusest piisab, et seda võimalust välistada. (Ammugi ei soovi ma väita, et neid saaks tuletada formaalsel teel; ülalselgitatud põhjustel​[[108]](#footnote-109) ei ole formaalsed süsteemid siin relevantsed.) Kuid põhimõtteliselt on võimalik väga mitmel viisil näidata, et keegi (mingi ülivõimekas olend?), kellel oleksid käepärast ainult mikrofüüsikalised faktid ja nende juurde kuuluvad mõisted, suudaks tuletada neist kõrgtaseme faktid.

Kõige lihtsam viis selleks on panna tähele, et põhimõtteliselt oleks võimalik konstrueerida maailma suur mõtteline simulatsioon ja jälgida seda nii-öelda vaimusilmas. Võtame näiteks mingi vihmavarjuga mehe. Olukorraga seonduvatest mikrofüüsikalistest faktidest saaks otseteed tuletada faktid ainemasside jaotumise ja keemilise koostise kohta mehe ümbruses; tulemuseks oleks selle piirkonna struktuurne kirjeldus kõrgtasemel. Lihast ja verest kahejalgse meessoost olendi olemasolu saaks kindlaks teha üsna sirgjooneliselt. Näiteks saaks olemasolevast struktuuriinformatsioonist järeldada, et selles piirkonnas asub organism, mis toetub kahele pikale jalale ning et need on tema kulgemisorganid, et selle olendi anatoomia vastab mehe anatoomiale jne. Selguks, et ta kannab mingit seadeldist, mis ei luba veepiiskadel, mida tema ümbruses leidub väga palju, talle peale sadada. Kahtlusi selles, et ese on tõepoolest vihmavari, saab hajutada, kui panna tähele, et tema füüsiline struktuur võimaldab teda kokku panna ja avada; et ta rippus kunagi varem ühel hommikul nagis, enne seda aga on ta algselt valmistatud tehases koos teiste sarnaste seadeldistega jne. Kahtlusi selles, et see lihast ja verest kahejalgne olend on tõepoolest inimene, võib hajutada, jälgides tema DNA koostist, tema evolutsioonilugu, sugulust teiste olenditega jne. Selleks tarvitseb ​**[lk.77]**​ vaid eeldada, et simulatsiooni jälgija valdab vajalikku mõistet piisavas mahus, et seda konkreetsetel juhtudel õigesti rakendada (või teisiti öeldes, ta valdab vastavat intensiooni). Kui nii, siis sisaldab mikrofüüsikaliste faktide kogum kõiki andmeid, mida on tarvis, et kohaldada neid mõisteid ning tuvastada, et siin on tõepoolest vihmavarjuga mees.

Sama mõttekäiku saab kohaldada peaaegu kõigile kõrgtaseme fenomenidele: laudade ja toolidele, bioloogilisele elule, majanduslikule õitsengule. Teades kõiki madaltaseme fakte, on simulatsiooni jälgijal põhimõtteliselt võimalik tuletada kõik faktid, mida ta vajab otsustamaks, kas kõnealune omadus ühes või teises olukorras realiseerub või mitte. Selleks tuleb konstrueerida mikrofüüsikaliste faktidega ühilduv võimalik maailm ning seejärel kõrgtaseme faktid sellest maailmast sobivat intensiooni rakendades lihtsalt maha lugeda (kuna relevantsed faktid on füüsiliselt identsetes võimalikes maailmades ühesugused). Seega on kõrgtaseme faktid füüsika suhtes loogiliselt superventsed.

##### Analüüsitavus.

Olen siiani öelnud, et mikrofüüsikaliste faktidega on kõrgtaseme faktid fikseeritud, ent pole kõrgtaseme mõistete kohta, mida selleks tarvis läheb, veel kuigi palju öelnud. Igal konkreetsel juhul aga on selle tulenevussuhte aluseks mõistete intensioonid. Kui mingi kõrgtaseme fakt tuleneb mikrofüüsikalistest faktidest, siis on see nii sellepärast, et mikrofüüsikalistest faktidest piisab maailma nende omaduste fikseerimiseks, mis kõrgtaseme intensiooni võimalikuks teevad. Teisiti öeldes, me peame oskama analüüsida, mis tingimustel mingi entiteet kõrgtaseme mõiste intensioonile vastab — vähemalt piisaval määral selleks, et mõista, miks saab neid vastavustingimusi rahuldada füüsikaliste faktide fikseerimise teel. Seepärast on kasulik vaadelda lähemalt kõrgtaseme mõistete intensioone ning uurida maailma omadusi, mis neid rakendada lubavad.

Nende intensioonide selgitamise ja sõnades väljendamise teel on mõningaid takistusi. Nagu [ülal](#PARA.43-2) nägime, on mõistete rakendustingimused sageli mingis osas ebamäärased. Kas kehakudedest​[[109]](#footnote-110) koosnev tassikujuline ese on tass? Kas arvutiviirus on elus? Kas aineosakestest juhuslikult moodustunud raamatusarnane objekt on raamat? Meie tavalised mõisted nendele küsimustele sirgjoonelisi vastuseid ei anna, ja mingis mõttes on need lihtsalt kokkuleppe küsimused. Seega ei ole meil selgepiirilisi rakendustingimusi, mida tuletamisprotsessis kohaldada. Kuid nagu me oleme [juba](#PARA.43-2) näinud, on see määramatus faktide endi määramatuse täpne peegelpilt. Kuivõrd sõna “tass” intensioon on kokkuleppe küsimus, on ka tasse puudutavad faktid kokkuleppe küsimus. Meie jaoks on oluline, et intensioon determineerib koos mikrofüüsikaliste faktidega kõrgtaseme faktid, kuivõrd need on tõepoolest faktilised. Ebaselgus ja ebamäärasus võivad jutu kohmakaks teha, kuid küsimustele endile neil mingit olulist mõju ei ole.

Lisaks on siin veel teine eelmisega sarnane raskus, mis seisneb selles, et mõistete lühidad analüüsid on alati ebatäpsed. Nagu nägime, ei ole mõistete definitsioonid tavaliselt kuigi selgepiirilised. Esimeses lähenduses võime öelda, et üks või teine objekt on laud, kui tal ​**[lk.78]**​ on tasane horisontaalne pind, mis toetub jalgadele; kuid see määratlus on ühekorraga liiga avar (Frankensteini koletis kõmpidel?) ja liiga kitsas (seinale kinnituv jalgadeta laud?). Definitsiooni saab täpsustada lisatingimuste ja -klauslitega, kuid nendega tulevad kaasa uued määramatused, ja laitmatut tulemust me lõpuks ikkagi ei saa. Kuid iga erijuhtumi käsitlemiseks vajalikesse üksikasjadesse ei ole tarviski süüvida: mingist punktist edasi hakkavad üksikasjad juba korduma. Asja mõistmiseks piisab, kui me teame, mis liiki omaduste alusel me intensiooni rakendame.

Nagu ülal nägime, ei nõua B-faktide tuletatavus A-faktidest B-omaduste definitsioone A-omaduste kaudu. Viimselt ei representeeri tähendust definitsioon, vaid intensioon. Analüüsi roll seisneb siin vaid intensioonide kirjeldamises küllalt üksikasjalikult, et tulenevus oleks selgesti näha. Selleks ei pea analüüs olema kuigi detailne. Üldiselt on intensioonid võimaliku maailma indiviididele rakendatavad tänu oma mingitele ühtedele omadustele, teised aga pole nii vajalikud; analüüsi mõte on mõista, millised on need omadused, tänu millele intensioon on kohaldatav, ja näidata, et need omadused on tuletatavad füüsikalistest omadustest.

Kolmanda raskuse allikaks on asjaolu, et paljudel mõistetel ei ühti aprioorsed rakendatavustingimused aposterioorsetega. Senikaua kui me hoiame primaarsete ja sekundaarsete intensioonide vahel selget vahet, ei ole häda kuigi suur. On ilmne, et sõna “vesi” sekundaarne intensioon — umbes nagu “H₂O” — on füüsika suhtes loogiliselt superventne. Kuid primaarne intensioon — umbes nagu “läbipaistev, joogikõlblik vedelik meie keskkonnas” — on samuti loogiliselt superventne, kuna vee läbipaistvus, joogikõlblikkus ja vedelus tulenevad füüsikalistest faktidest.[[110]](#footnote-111) Me võime võtta asja ükskõik kumba pidi.​[[111]](#footnote-112). Esmajoones aga huvitab meid primaarne intensioon, sest, [nagu nägime](#PARA.47-3), toimub reduktiivne seletamine just primaarse intensiooni alusel. Kui primaarne intensioon *I* on füüsika suhtes loogiliselt superventne, siis on seda üldiselt ka rigidiseeritud sekundaarne intensioon *dthat*​(​*I*​), kuna selle aluseks on üldjuhul mingi seesmise füüsikalise struktuuri projektsioon kõigile maailmadele.

Mõned autorid leiavad, et kõrgtaseme faktid ei ole loogiliselt tuletatavad madaltaseme faktidest, viidates kinnituseks aposterioorsele paratamatusele. Tavaliselt kõneldakse umbes nii: “Vesi on paratamatult H₂O, kuid see ei ole tõde tähenduse alusel, mistõttu siin kontseptuaalne seos puudub”. Kuid see on tohutu lihtsustus. Esiteks võib sekundaarset intensiooni “H₂O” pidada mingis mõttes sõna “vesi” tähenduse osaks, ja see intensioon on kindlasti loogiliselt superventne. Kuid veelgi olulisem on see, et primaarne intensioon (“läbipaistev, joogikõlblik vedelik …”), mis fikseerib referendi, on samuti superventne, kuigi võib-olla mitte arvestades elamuse ja indeksikaalsuse panust. Just vastavuse tõttu sellele intensioonile me ju olemegi otsustanud pidada vett H₂O-ks. Primaarse intensiooni ​*I* abiga on kõrgema tasandi faktid raskusteta tuletatavad ​**[lk.79]**​ mikrofüüsikalistest faktidest (kui mitte arvestada elamuse ja indeksikaalsuse panust). Kripke tähelepanek, et *dthat*​(​*I*​) representeerib mõistet paremini, ei mõjuta seda tuletatavust sugugi. Semantilise rigidiseerimise fenomen üksi ei muuda ontoloogilises mõttes midagi.

Need takistused jalust ära, võime nüüd hakata uurima mitmesuguste kõrgtaseme mõistetega seonduvaid intensioone. Enamikul juhtudest saab neid kirjeldada funktsionaalses või struktuurses mõistetes või nende mõlematega. Näiteks selleks, et miski oleks laud, on nõutav, et: 1) tal oleks lame pind, mis toetub jalgadele, ja 2) ta oleks kasutatav tarbeesemena, mille peale saab asju panna. Esimene neist tingimustest on struktuuriline, puudutades objekti seesmist füüsilist struktuuri. Teine tingimus on funktsionaalne, puudutades entiteedi välist põhjuslikku rolli ja iseloomustades tema suhteid teiste entiteetidega. Struktuursed omadused on silmnähtavalt mikrofüüsikaliste faktide tulem. Üldiselt võib seda öelda ka funktsionaalsete omaduste kohta, kuigi siin ei ole tulenevus nii sirgjooneline. Funktsionaalsed omadused sõltuvad palju laiemast mikrofüüsikaliste faktide superventsusbaasist, mistõttu tuleb sageli arvesse võtta ka fakte objekti ümbruse kohta; ja niivõrd kui nende omaduste kirjeldus on soodumuslik (selleks, et teha kindlaks aine lahustuvus vees, tuleb viia ta kokkupuutesse veega), tuleb võtta appi kontrafaktuaalid. Ent kuna me lisasime oma superventsuskonditsionaalidesse eesliikmetena füüsikaseadused, mis nende kontrafaktuaalide tõeväärtused fikseerivad, siis sellest raskusi juurde ei tule.

Võtame teise näite. Tingimused, mille põhjal midagi elusaks pidada, võib umbkaudselt taandada mingile võimete kombinatsioonile, millede hulka kuuluvad (muu kõrval) reproduktsiooni-, kohanemis- ja ainevahetusvõime. (Nagu mujalgi, nii ka siin ei pruugi me sätestada täpseid proportsioone ja kõiki asjakohaseid tegureid). Kõiki neid omadusi saab kirjeldada funktsionaalselt — indiviidi suhete kaudu teiste indiviididega ning tema võime kaudu muundada väliseid ressursse energiaks ja reageerida kohasel viisil oma keskkonnale. Kõik need funktsionaalsed omadused on põhimõtteliselt tuletatavad füüsikalistest faktidest. Isegi kui bioloogilisele elule ei leidu täiuslikku funktsionaalset definitsiooni, näitab juba see kirjelduski, et elu on funktsionaalne omadus, mille realiseeringud on seega tuletatavad füüsikalistest faktidest.

Üks keerukus tuleneb asjaolust, et entiteedi funktsionaalseid omadusi kirjeldatakse sageli tema põhjusliku rolli kaudu tema suhetes teiste kõrgtaseme entiteetidega. Sellest järeldub, et nende omaduste loogiline superventsus sõltub teiste nendega suhestuvate kõrgtaseme mõistete loogilisest superventsusest, kusjuures neid mõisteid võidakse samuti kirjeldada funktsionaalselt. Senikaua kui põhjuslikke rolle saab viimselt taandada mittefunktsionaalsetele omadustele — tavaliselt kas struktuursetele või fenomenaalsetele —, ei too see kaasa mingeid raskusi. Erinevate funktsionaalsete omaduste definitsioonides üksteise kaudu võib tekkida mõningane tsirkulaarsus — näiteks klammerdaja mõistesse võib kuuluda see, et ta lööb klambreid, ja klambri mõistesse see, et teda lüüakse klammerdajaga. Tsirkulaarsusest ülesaamiseks võib kõigi nende omaduste juures ühekorraga jätta arvestamata​[[112]](#footnote-113) nende põhjuslikud rollid[[113]](#footnote-114) [[114]](#footnote-115); see on võimalik, senikaua kui neil analüüsidel on olemas mittetsirkulaarne komponent, mis ​**[lk.80]**​ toetub lõppkokkuvõttes struktuursetele või fenomenaalsetele omadustele. (Võib tunduda, nagu poleks fenomenaalsete omaduste appivõtmine ühitatav loogilise superventsusega füüsika suhtes; kuid allpool näeme, et see ei ole nii. Igal juhul on toetumine fenomenaalsetele omadustele kooskõlas loogilise superventsusega, kui mitte arvestada teadvuselamuse panust.)

Paljude omaduste kirjeldus on relatsiooniline, s.t. viitab entiteedi suhetele oma keskkonnaga. Tavaliselt on need suhted põhjuslikud ning kõnealused omadused seega funktsionaalsed, kuid mitte alati — näiteks ei ole seda omadus paikneda mõne kindla pardiga samal mandril. Mõned omadused sõltuvad jällegi ajaloost (kuigi neid saab enamasti tõlgendada põhjuslikult); selleks, et olend oleks känguru, peavad tal olema õiget liiki esivanemad. Igal juhul ei valmista säärased omadused loogilisele superventsusele raskusi, sest asjaomased ajaloo ja keskkonna faktid fikseeruvad ise ühes globaalsete füüsikaliste faktidega.

Isegi sääraseid keerukaid ühiskondlikke fakte nagu “1950-ndad aastad olid majandusliku õitsengu aeg”[[115]](#footnote-116) saab kirjeldada suuremas osas funktsionaalselt, mistõttu neid võib pidada tulenevateks füüsikalistest faktidest. Täielik analüüs oleks õitsengu mõiste ebamäärasuse tõttu väga keeruline, kuid et saada mingit ettekujutust selle analüüsi käigust, võime endalt küsida: miks me üldse ütleme, et 1950-ndad aastad olid majandusliku õitsengu aeg? Esimeses lähenduses sellepärast, et tööhõive oli kõrge ja inimeste ostujõud erakordselt suur, inflatsioon oli väike, ehitati palju eluasemeid jne. Edasi võib hakata andma umbkaudseid analüüse kasutatud mõistetele, näiteks eluasemele (koht, kus inimesed magavad ja söövad), tööhõivele (organiseeritud töö, mille eest makstakse tasu) ja monetaarsetele mõistetele (oletatavasti on raha umbkaudselt analüüsitav süstemaatilise võimalusena vahetada teda teiste esemete vastu ning raha väärtus selle kaudu, kui palju tema eest vastu saab). Me andsime siin kõik need analüüsid naeruväärsuseni lihtsustatud kujul, kuid mõte on küllalt selge. Üldiselt on need kõik funktsionaalsed omadused, mida saab tuletada füüsikalistest faktidest.

Paljud autorid on väljendanud kahtlusi kontseptuaalse analüüsi võimalikkuses. Sageli on seda tehtud põhjustel, mis ei mõjuta argumente, mida mina siin esitan — siia kuuluvad näiteks meie mõistete ebamäärasus ja kristallselgete definitsioonide puudumine. Kuid mõnel juhul on skepsisel ka sügavam alus. Ent kui see, mida ma olen käesolevas peatükis seni öelnud, on õige, ja kui võimaliku maailma füüsikaliste faktidega on fikseeritud ka kõrgtaseme faktid, siis võiks oodata, et ka vastavate kõrgtaseme mõistete intensioonid on analüüsitavad — vähemasti küllalt heas lähenduses, et näha, kuidas nende kohaldatavust saab kindaks määrata füüsikaliste faktide põhjal. Seda ma püüdsingi siinsetes näidetes teha. Sarnaselt võib käsitleda ka teisi näiteid.[[116]](#footnote-117)

Ma ei propageeri selliste analüüside sooritamise programmi üldiselt. Mõisted on liiga keerulised ja tõrksad asjad, et sellest kuigi palju kasu oleks, ning iga sõnadesse pandud analüüs on enamasti vaid asja enda kahvatu vari. Oluline on ​**[lk.81]**​ üldine mõte, et enamik kõrgtaseme mõisteid ei ole analüüsimatud elementaarmõisted. Üldiselt on nad analüüsitavad niivõrd, kui nende intensioonides võib näha funktsionaalsete või struktuursete omaduste kirjeldust. Just tänu oma analüüsitavusele on kõrgtaseme fakte põhimõtteliselt võimalik tuletada mikrofüüsikalistest faktidest ja reduktiivselt seletatada füüsikaliste faktide kaudu.

#### Mõned raskemad juhud

Mõnda tüüpi omaduste kohta võiks oodata, et nende loogiline superventsus ja seega ka reduktiivne seletamine valmistab erilisi raskusi. Allpool vaatlen mõningaid selliseid omadusi lähemalt, pöörates erilist tähelepanu küsimusele, kas raskused, mida valmistab vastavate fenomenide reduktiivne seletamine, on analoogilised raskustega, mida valmistavad reduktiivsele seletamisele teadvusega seonduvad probleemid. Mulle tundub, et kui paari võimalikku erandit mitte arvestada, siis siin mingeid olulisi uusi raskusi ei teki.

#####  Teadvusest sõltuvad omadused

Nagu juba mainitud, eeldavad mõnede mõistete primaarsed intensioonid mingit suhet teadvuselamustega. Ilmekas näide on punasus, kui mõista selle all väliste objektide omadust. Vähemalt mõne käsitluse järgi nõuab punasusega seonduv primaarne intensioon, et selleks, et mingi objekt oleks punane, peab tal olema soodumus kutsuda sobivates tingimustes esile punaseid elamusi.[[117]](#footnote-118) Seega ei ole punasus oma primaarses intensioonis füüsika suhtes loogiliselt superventne, ehkki ta on seda, kui mitte arvestada teadvuselamusi. Seevastu punasuse mõiste sekundaarne intensioon on peaaegu kindlasti superventne. Kui selgub, et aktuaalses maailmas kalduvad punast elamust põhjustama pindade omadus valgust kindlal viisil peegeldada, siis on sellise peegeldusvõimega objektid punased ka maailmades, kus pole teadvusega olendeid neid nägemas. Punasus on *a posteriori* samastatav nende pinnaomadustega, mis on loogiliselt superventsed ainult füüsika suhtes.

[Ülal](#PARA.47-3) nägime, et primaarsetel intensioonidel kaasneb loogilise superventsuse puudumisega reduktiivse seletamise võimatus. Kas see tähendab, et punasust on võimatu reduktiivselt seletada? Mingis nõrgas mõttes on vastus jaatav. Kui mõista punasuse all kalduvust põhjustada punaseid elamusi, siis kuivõrd elamus ei ole reduktiivselt seletatav, ei ole seda ka punasus. Kuid reduktiivsele seletusele võib lähedale jõuda. Võib öelda, et üks või teine füüsiline omadus põhjustab punaseid elamusi, ja võib isegi seletada põhjuslikku suhet selle omaduse ja punaste otsustuste vahel. Seletamata jääb vaid viimane samm, mis viib elamuse endani. Praktikas on meie nõuded seletustele küllalt nõrgad, et pidada ka niisugust seletust seletuseks​.[[118]](#footnote-119) Seletades fenomeni, mille referent fikseerub mingi elamuse kaudu, ei nõua me elamuse seletamist. Vastasel korral tuleks meil kaua oodata.

Sama võib öelda ka soojuse, valguse, helide ja teiste taoliste fenomenide kohta. Kuigi nende sekundaarsed intensioonid hõlmavad struktuurseid omadusi (molekulaarliikumine, ​**[lk.82]**​ footonite jaotus, õhulainetus), kuulub nende primaarsete intensioonide koosseisu seos teadvuselamusega: soojus on see, mis põhjustab soojaaistinguid, valgus põhjustab nägemiselamusi jne. Kuid nagu on tähendanud Nagel (1974) ja Searle (1992), ei ole soojuse seletamiseks tarvis seletada soojaaistinguid. Piisavalt hea on juba niisugunegi seletus, mis elamustega ei arvesta.

Teised omadused sõltuvad teadvuselamusest veelgi otsesemalt, sest lisaks sellele, et elamused mängivad rolli referentide fikseerimisel, konstitueerivad nad osalt ka aposterioorset mõistet. Ilmseks näiteks on omadus seista kõrvuti isikuga, kellel on teadvus. Mõne tõlgenduse kohaselt sõltuvad vaimsed omadused, näiteks armastus ja usk, mis ei ole ise fenomenaalsed omadused, kontseptuaalselt siiski teadvuselamuste olemasolust. Kui see on nii, siis ei oleks need omadused realiseeritavad teadvuseta maailmas. Seega ei ole sellised omadused loogiliselt superventsed isegi aposterioorselt, ja reduktiivne seletamine on võimatu veelgi tugevamas mõttes kui eelmistel juhtudel. Ent kui teadvuselamusi mitte arvestada, on nad loogiliselt superventsed ja reduktiivselt seletatavad, seega on see takistus reduktiivse seletamise teel meile juba tuttav ja mingeid uusi raskusi ei ilmne.

##### Intentsionaalsus.

Eraldi tasub käsitleda intentsionaalsuse staatust, kuna mõnikord arvatakse, et see seab meie ette raskusi, mis on analoogilised nendega, mida valmistab teadvus. Võib siiski arvata, et intentsionaalsetel omadustel tuleneb loogilise superventsuse puudumine alati teadvuse mittesuperventsusest. Nagu ma 1. peatükis mainisin​, ei ole vististi mõeldavat maailma, mis oleks meie omaga füüsiliselt ja fenomenaalselt identne, kuid erineks sellest oma intentsionaalselt sisult.[[119]](#footnote-120) Kui intentsionaalset sisu konstitueerib mingis osas fenomenoloogia, nagu mõned filosoofid arvavad, siis ei tarvitse intentsionaalsed omadused füüsika suhtes loogiliselt superventsed olla, ent nad on seda, kui teadvuselamused arvestusest välja jätta. Väide, et mingis osas konstitueerib intentsionaalset sisu teadvus, on vaieldav, kuid nii või teisiti ei ole kuigi palju põhjust uskuda, nagu oleks intentsionaalsusele omane mingi erilist liiki mittesuperventsus, mis ei tulene teadvuse mittesuperventsusest​.

Kõiki fenomenoloogilisi aspekte kõrvale jättes on inimeste käitumise seletamisel kõige parem vaadelda intentsionaalseid omadusi omamoodi kolmanda isiku konstruktina, ning sellistena peaks neid olema võimalik analüüsida käitumise ja keskkonna põhjuslike seoste kaudu. Sel juhul on intentsionaalsed omadused füüsika suhtes otsesel moel loogiliselt superventsed. Lewis (1974) on teinud tõsise katse näidata intentsionaalsete faktide tulenevust füüsikalistest faktidest sobiva funktsionaalse analüüsi varal. Hilisemaid intentsionaalsuse kontseptsioone, sealhulgas neid, mida on esitanud Dennett (1987), Dretske (1981) ja Fodor (1987), võib vaadelda kaastööna samale projektile. Ükski neist analüüsidest ei ole lõpuni veenev, kuid pole võimatu, et nende edasiarendamisel sünnib kunagi ka midagi paremat. Pole ühtki teadvuse superventsuse vastaste argumentidega analoogilist argumenti, mis näitaks, et intentsionaalsus ​**[lk.83]**​ ei saa olla füüsikaliste ja fenomenaalsete omaduste suhtes loogiliselt superventne.[[120]](#footnote-121) [[121]](#footnote-122) Tõepoolest, mõeldavusest lähtuvad argumendid näitavad, et kui intentsionaalsed omadused üldse realiseeruvad, siis peavad nad olema nende suhtes loogiliselt superventsed, ja ka epistemoloogilised argumendid viivad meid sarnasele järeldusele. Seega ei too intentsionaalsus kaasa mingeid uusi ontoloogilisi raskusi.

##### Moraalsed ja esteetilised omadused.

Sageli kinnitatakse, et pole olemas kontseptuaalset vahelüli, mis seostaks füüsikalisi omadusi moraalsete ja esteetiliste omadustega. Moore’i (1922) järgi ei ole näiteks headuse mõistes midagi, mis lubaks tuletada headust puudutavaid fakte füüsikalistest faktidest. Tegelikult väitis Moore, et pole olemas kontseptuaalset ühendusteed looduse faktidest moraalsete faktideni, kui arvata looduse faktide alla nii psüühilised kui füüsikalised faktid (mismõttu viitamine teadvuselamust mitte arvestavale superventsusele ei anna meile siin midagi). Kas see tähendab, et moraalsed omadused on sama problemaatilised nagu teadvuselamus?

On siiski kaks asjaolu, mis seda analoogiat rikuvad. Esiteks ei tundu maailm, mis oleks meie omaga looduslikult identne, kuid temast moraalselt erinev, mõeldaana, mistõttu on ebatõenäoline, et moraalsed faktid oleksid mingis vähegi tugevas mõttes lisafaktid. Teiseks ei ole moraalsed faktid fenomenid, mis endid meile peale sunniksid. Kui kitsas käes, võib moraalifaktide olemasolu isegi täielikult eitada. Just niisuguse strateegia ongi valinud mõned moraalsed antirealistid, näiteks Blackburn (1971) ja Hare (1984). Nemad väidavad, et kuna moraalsed faktid ei ole tuletatavad looduslikest faktidest ega ole ka piisavalt “veidrad”, et pidada neid lisafaktideks, siis ei ole neil mingit objektiivset olu, mistõttu moraal tuleks relativiseerida meie kognitiivse aparaadi konstruktiks või projektsiooniks. Seda strateegiat ei saa aga kohaldada fenomenaalsetele omadustele, mille olemasolus pole põhjust kahelda.

Moraalsete omaduste käsitlemiseks on vähemalt kaks mõttekat alternatiivi. Esimene on mingit liiki antirealism, mis näiteks relativiseeriks “objektiivsed moraalsed faktid” “subjektiivseteks moraalseteks faktideks”[[122]](#footnote-123), või ütleks, et moraalidiskursuses ei olegi juttu faktidest. Teine võimalus on postuleerida mingi aprioorne ühendustee, mis viiks looduslikelt faktidelt moraalifaktideni ja mille kehtivust saaks (vastukaaluks Moore’ile) näidata moraalimõistete analüüsi ja selgitamise teel. Kui mõni moraalimõiste, näiteks “hea”, defineerib mingi stabiilse mitteindeksikaalse primaarse intensiooni, siis ongi tulemuseks see viimatine positsioon: siis on meil käes aprioorne funktsioon maailma looduslikult kirjelduselt sama maailma moraalifaktidele. Ent kui see mõiste defineerib ainult indeksikaalse primaarse intensiooni, või kui eri subjektid võivad seostada temaga erinevaid, kuid ühevõrra mõttekaid primaarseid intensioone, või kui ta ei määratlegi ühtki primaarset intensiooni, siis jõuame mingil kujul esimesele seisukohale.

On pakutud ka muid lahendusi, kuid ükski neist ei tundu olevat vettpidav. Moore’i järgi ion looduslike ja moraalsete faktide vahel mittekontseptuaalne aprioorne seos, mille avab salapärane võime nimega “moraalne intuitsioon”. Kuid see arvamus ei ole kuigi palju poolehoidu leidnud, kuna pole selge, millega saaks põhjendada moraalsele intuitsioonile toetuvate väidete tõesust või väärust. Vaade, mille järgi moraalsed omadused on superventsed ​**[lk.84]**​ mingi fundamentaalse noomilise vahelüli kaudu, langeb vist kohe ära, kuna pole mõeldavat maailma, mille looduslikud faktid ühtiksid meie omadega, kuid moraalsed faktid oleksid teised. Meieaegsete moraalsete realistide seas (vt. nt. Boyd 1988; Brink 1989) on populaarne seisukoht, et moraalifaktid on loodusfaktide suhtes superventsed aposterioorse paratamatusega, s.t. superventsed mitte moraalimõistete primaarsete, vaid nende sekundaarsete intensioonide järgi. Sellele seisukohale on aga raske pidama jääda, kuna ka aposterioorsed ekvivalentsussuhted peavad toetuma viidete aprioorsele fiksatsioonile​: kuigi vesi on H₂O aposterioorselt, tulenevad vett puudutavad faktid mikrofüüsikalistest faktidest ​*a priori*​. Ja​[[123]](#footnote-124) kui moraalimõistetel on olemas primaarne intensioon ja kui looduslikult identsed keskpunktiga maailmad on ka moraalselt identsed, siis näib sellest tulenevat, et on olemas aprioorne vahelüli, mis ühendab looduslikke fakte moraalifaktidega. (Umbkaudu niisuguses vaimus kritiseerivad seda seisukohta Horgan ja Timmons (1992a; 1992b)).

Analoogiliselt saab käsitleda ka esteetilisi omadusi. Siin on antirealistlik käsitlusviis vähemalt sama veenev. Kõike kokku võttes võib öelda, et kuigi mõtiskluste käigus selle kohta, kuidas moraali ja esteetika valdkonda käsitleda tuleks, kerkib esile huvitavaid kontseptuaalseid küsimusi, ei too need kaasa metafüüsilisi ja seletuslikke probleeme, mis oleksid võrreldavad probleemidega, mida seab meie ette teadvuselamus.

##### Nimed.

Paljud autorid (vt. nt. Kaplan 1989) kinnitavad, et nimede, näiteks nime “Rolf Harris” taga ei ole mingit analüüsi, vaid nad lihtsalt osutavad otseselt oma referendile. Kas see tähendab, et omadus olla Rolf Harris ei ole füüsika suhtes loogiliselt superventne? Sekundaarse intensiooni superventsusega siin küsimust ei ole (näiteks võib Rolf kõigis võimalikes maailmades olla ühest kindlast seemne- ja ühest kindlast munarakust eostunud inimene), kuid on võimalik arvata, et primaarse intensiooni puudumine valmistab reduktiivsele seletamisele raskusi. Siiski on tõenäoline, et kuigi nimedel puudub primaarne intensioon, mis kehtiks kogu kogukonnas, kaasneb nime lausumisega iga kord siiski mingi primaarne intensioon. Kui ma lausun nime “Rolf Harris”, siis sõltub nime referent mingil süstemaatilisel viisil asjade seisust maailmas; minu jaoks võib primaarseks intensiooniks olla näiteks midagi niisugust nagu “mees nimega Rolf Harris, kes kolistab värvipurkidega ja kellel on minuga nii- ja niisugune põhjuslik suhe”.[[124]](#footnote-125) Selline intensioon oleks loogiliselt superventne. Kuid selle asemel, et hakata seda väidet üksikasjalikult põhjendama, piisab juba lihtsast nentimisest, et loogilise superventsuse puudumine ei too siin seletuslikus mõttes kaasa mingit müsteeriumi. Omadus olla Rolf Harris ei ole fenomen, mis nõuaks seletust (kusjuures seletamine ei tähenda siin sama mis selgitamine). Seletust nõuab vaid isiku olemasolu, kelle nimi on Rolf Harris, kes kolistab värvipurkidega jne. Need omadused on kindlalt superventsed ja neid on põhimõtteliselt võimalik seletada tavalisel viisil.

##### Indeksikaalid

Paljudel mõistetel, mõistest “vesi” alates ja mõistega “minu koer” lõpetades, sisaldab referendi fiksatsioon indeksikaalset elementi. Nende mõistete referent fikseerub ​**[lk.85]**​ nii füüsikaliste faktide kui ka subjektikohase *indeksikaalse fakti* alusel — viimane viitab vastavat mõistet kasutava subjekti asukohale. See fakt on iga subjekti puhul ühene, mistõttu ka referent fikseerub üheselt. Superventsus kehtib ja seletus peab paika, kui seda indeksikaalset fakti mitte arvestada.

Kas indeksikaalsus on reduktiivsele seletamisele takistuseks? Suvalise kõneleja puhul vahest mitte, sest indeksikaalsest “faktist” saab relativiseerimise teel vabaneda. Iseenda puhul ei ole see nii lihtne. Indeksikaalne fakt väljendab midagi väga olulist maailma kohta, nagu mina selle eest leian: asjaolu, et David Chalmers olen mina. Kuidas saaks seda pealtnäha elementaarset fakti seletada? Või kas see ongi fakt, mis nõuab seletust, aga mitte lihtsalt tautoloogia? Seda küsimust on erakordselt raske käsitada, kuid mulle tundub, et kuigi indeksikaal ei ole võib-olla maailma objektiivne fakt, on ta fakt maailma kohta niisugusena, nagu mina ta eest leian, ja just maailm, nagu mina ta eest leian, nõuabki seletust. Tooreste indeksikaalide olemus on aga üsna udune ja pole hoopiski selge, kuidas neid seletada.[[125]](#footnote-126) (Jah, me võime reduktiivselt seletada, miks lausung “Mina olen David Chalmers” on tõene, kui seda ütleb David Chalmers . Kuid see mitteindeksikaalne fakt tundub olevat hoopis midagi muud kui indeksikaalne fakt, et mina olen David Chalmers).

Selles seoses oleks ahvatlev pöörata pilk teadvuse poole. Ent kuigi teadvuse seletus võiks ehk anda seletuse “vaatepunktidele” üldiselt, on raske näha, kuidas ta võiks seletada, miks just see pealtnäha meelevaldne vaatepunkt on minu vaatepunkt — kui just mitte olla solipsist. Pole võimatu, et indeksikaalset fakti tuleb võtta kui midagi elementaarset. Sellisel juhul on reduktiivne seletamine selles küsimuses võimatu, nagu ta on võimatu ka teadvuse puhul, kusjuures need on kaks erinevat võimatust. Siiski teeb võimatus selles küsimuses vähem muret kui teadvuse puhul, sest fakt, mis siin seletamatuks jääb, on nii “kõhn” võrreldes teadvuse faktidega nende täies hiilguses. Selle elementaarse indeksikaalse fakti tunnistamine nõuaks meie materialistlikus maailmapildis palju vähem olulisi muudatusi kui mitteredutseeritavate teadvuselamuse faktide tunnistamine.

##### Negatiivsed faktid.

Nagu ülal nägime, ei tulene mõningad faktid, milles figureerivad eitavad eksistentsiaal- ja universaalkvantorid, loogiliselt füüsikalistest faktidest ega üldse ühestki lokaliseeritud​[[126]](#footnote-127) faktide kogumist. Vaatleme järgmisi meie maailma fakte: ingleid ei ole olemas; Don Bradman on suurim kriketimängija; kõigi elusolendite aluseks on DNA. Kõiki neid fakte oleks võimalik väga lihtsalt falsifitseerida, sattumata vastuollu ühegi meie maailma füüsikalise faktiga: selleks tuleb vaid lisada maailmale uusi mittefüüsilisi entiteete, näiteks kriketit mängivad ja ektoplasmast koosnevad inglid. Selle vastu ei aita isegi teadvuselamuse või indeksikaalsuse faktide lisamine.[[127]](#footnote-128) [[128]](#footnote-129)

Kas see tähendab, et neile faktidele ei olegi võimalik reduktiivset seletust leida? Nähtavasti küll, kuivõrd meil puudub füüsikaline seletus, miks ei ole meie maailmas mittefüüsilisi lisaentiteete. See on tõepoolest lisafakt. Parim viis selle olukorraga toimetulemiseks on lisada maailmale teist järku fakt, mis ütleb ​**[lk.86]**​ kõigi konkreetsete baasfaktide kogumi kohta, olgu need siis mikrofüüsikalised, fenomenaalsed, indeksikaalsed või veel ükskõik missugused faktid: “See on kõik.” See viimane fakt ütleb, et kõik maailma partikulaarsed faktid sisalduvad selles baasfaktide kogumis või tulenevad neist. Sellest teist järku faktist koos kõigi konkreetsete baasfaktidega saab tuletada kõik negatiivsed faktid.

Reduktiivsete seletuste võimatus ei ole siin väga tõsine raskus. Arvatavasti ongi säärane fakt “see on kõik” tõene suvalises maailmas, ja niisugust fakti ei ole kunagi võimalik tuletada konkreetsetest faktidest, vaid ta väljendab lihtsalt asjaolu, et meie maailmal ja üldse kõigil maailmadel on olemuslikud piirid. Nii saame suuri kulusid kandmata hõlmata kõik negatiivsed eksistentsiaal- ja universaalkvantoritega faktid.

##### Füüsikaseadused ja põhjuslikkus.

Kõige veenvamate kontseptsioonide järgi ei ole füüsikaseadused füüsikaliste faktide suhtes loogiliselt superventsed, kui vaadelda viimaseid maailma aegruumiloo konkreetsete faktide kogumina. Seda võib mõista, kui panna tähele, et kogu oma aegruumiloo ulatuses meie omast füüsikaliselt eristamatu, kuid teistsuguste seadustega maailm on loogiliselt võimalik. Näiteks võib tolles teises maailmas kehtida seadus, et kui koguda vaakumis üheks tükiks kokku kakssada tonni puhast kulda, muundub see pliiks. Kõiges muus on selle maailma seadused samad — väikeste muudatustega, kus tarvis. Meie maailma tegelikus aegruumiloos ei ole mitte kunagi olnud kahtsadat tonni kulda vaakumis ühes tükis. Järelikult on meie maailma ja teise maailma ajalugu identne, seadused aga sellegipoolest erinevad.

Taolised argumendid viivad mõttele, et loodusseadused ei ole konkreetsete füüsikaliste faktide kogumi suhtes loogiliselt superventsed.[[129]](#footnote-130) Sarnaste mõttekäikude varal saab näidata, et põhjuslik seos kahe sündmuse vahel on midagi enamat kui lihtne sündmuste reeglipära. Mitmesugused ​Hume’i vaadetest lähtuvad​[[130]](#footnote-131) mõtlejad püüavad neid järeldusi kahtluse alla seada, kuid mulle tundub, et siin tuleb neil argumentidest puudu.[[131]](#footnote-132) Seaduste olemasolus ja põhjuslikkuses on midagi taandamatut.

Teisal olen ma neist raskustest vabanemiseks arvanud superventsusbaa­si koosseisu füüsikaseadused, kuid sellega olen ma pigem metafüüsilisest mõistatusest mööda hiilinud kui sellele vastust otsinud. Tõsi, reduktiivsete seletuste võimatus on seaduste ja põhjuslikkuse puhul vähem tõsine puudujääk kui teadvuse puhul. Seadused ja põhjuslikud suhted on ise postuleeritud olemasolevate füüsikaliste fenomenide, nimelt hulgaliste looduses aset leidvate reeglipärade seletamiseks, teadvus on aga täiesti elementaarne *explanandum*​. Sellegipoolest esitab selliste redutseerimatute lisafaktide olemasolu juba iseenesest sügavaid küsimusi nende metafüüsilise olemuse kohta. Teadvuselamuste ja võib-olla ka indeksikaalsuse kõrval on need ainsad lisafaktid, millesse meil on vähegi põhjust uskuda. On üsna loomulik spekuleerida, et need kaks mittesuperventsete nähtuste eri liiki — teadvus ja põhjuslikkus — on ehk lähedases metafüüsilises suguluses.

#### ​**[lk.87]**​ Kokkuvõte

Meile on seega jäänud seisukoht, et peaaegu kõik faktid on füüsikaliste faktide (sealhulgas füüsikaseaduste) suhtes loogiliselt superventsed; võimalikeks eranditeks on teadvuselamused, indeksikaalsus ja negatiivsed eksistentsiaalfaktid. Teisiti sõnastades võime öelda, et maailma faktide ammendav loetelu on järgmine: 1) partikulaarsed füüsikalised faktid, 2) teadvuselamuse faktid, 3) loodusseadused, 4) teist järku fakt “ja see on kõik”, ja võib-olla ka 5) minu asukoha indeksikaalne fakt. (Kaks viimast on teistega võrreldes vähetähtsad ja viimase staatus on kahtlane, kuid täielikkuse huvides võtsin ma nad loetellu siiski sisse.) Näib, et kui mitte arvestada teadvuselamuse ja indeksikaalsuse panust, siis on kõik positiivsed faktid füüsika suhtes loogiliselt superventsed. Et seda lõplikult kinnitada, tuleks uurida kõikvõimalikke fenomene üksikasjalikumalt, ent kõige põhjal, mida seni näinud oleme, võib seda pidada arukaks järelduseks. Ontoloogilise ja epistemoloogilise olukorra võime kokku võtta paari väljamõeldud looga. Pole võimatu, et neis lugudes peitub tõetera — mitte ehk igas üksikasjas, küll aga üldistes kontuurides.

​*Loomismüüt.* Maailma loomiseks tarvitses Jumalal vaid fikseerida äsja loetletud faktid. Jõukulu kokkuhoiu maksimeerimiseks fikseeris ta kõigepealt loodusseadused — füüsikaseadused ja kõik füüsikat ja teadvuselamust omavahel suhestavad seadused. Seejärel fikseeris ta rajatingimused: võib-olla konkreetsed füüsikalised faktid mingil ajalõigul[[132]](#footnote-133), võib-olla ka juhuslike arvude generaatori algväärtused. Need tingimused, ühenduses seadustega, fikseerisid ülejäänud füüsikalised ja fenomenaalsed faktid. Ja lõpetuseks teatas Jumal: “Ja see on kõik.”

​*Epistemoloogiline müüt.*​ Esialgselt on mul käepärast ainult minu teadvuselamuse faktid. Nende põhjal tuletan ma faktid keskmise suurusega objektide kohta maailmas ja pikapeale ka mikrofüüsikalised faktid. Nende faktide reeglipäradest tuletan ma füüsikaseadused ning leian seekäigus ka uusi füüsikalisi fakte. Reeglipärastest vastavustest minu teadvuselamuste ja füüsikaliste faktide vahel tuletan ma psühhofüüsika seadused ja seega ka teiste olendite teadvuselamuse faktid. Kuna abduktsioon on nüüd jõudnud viimase piirini, siis püstitan ma hüpoteesi: see on kõik. Ja kuna maailm on palju suurem kui algul tundus, siis tõstan kõige muu seast esile esialgsed teadvuselamused ja nimetan need *minu elamusteks*​.

Pangem tähele, kui erinev on toimingute järjekord neist kahest perspektiivist vaadates. Võiks peaaegu öelda, et epistemoloogia on ontoloogia rekapitulatsioon tagurpidi. Pangem tähele ka, et minu indeksikaalse fakti fikseerimine ei ole nähtavasti Jumala võimuses. Võib-olla annab see veel ühe põhjuse selles faktis kahtlemiseks.

Järeldus, mille kohaselt enamik kõrgtaseme fenomene on loogiliselt superventsed, ei ole leidnud nii laialdast tunnustust, kui oleks pidanud — isegi mitte autorite seas, kes on superventsuse üle arutlenud. Ehkki seda teemat ei võeta jutuks kuigi sageli, ei ole paljud mõtlejad kuigi meelsasti pidanud kontseptuaalset modaalsust sobivaks vahendiks, ​**[lk.88]**​ mille varal arutleda superventsussuhete üle. Niipalju kui mina otsustada oskan, on säärasel kõhklevusel olnud mitu põhjust, millest ükski ei tundu olevat küllalt kaalukas.

Esiteks, raskused loogiliselt võimalike füüsiliselt identsete maailmadega, kus on ka mingit mittefüüsilist sisu (inglid, ektoplasma), on viinud mõned autorid oletusele, et superventsussuhted ei saa olla loogilised (Haugeland 1982; Petrie 1987); kuid me oleme juba näinud​, kuidas sellega toime tulla. Teiseks näitavad aposterioorset paratamatust puudutavad kaalutlused paljude autorite arvates, et superventsussuhteid ei ole võimalik väljendada​[[133]](#footnote-134) tähenduste kaudu (Brink 1989; Teller 1984); meie aga oleme näinud, et superventsussuhteid, mille aluseks on aposterioorne paratamatus, võib vaadelda kui loogilise superventsuse üht erikuju. Kolmandaks on Quine’i mõjul maad võtnud üldine skepsis kontseptuaalse tõe mõiste suhtes; meie aga nägime, et siin on see küll täielik valejälg. Neljandaks, mure “redutseeritavuse” pärast on viinud mõned autorid arvamusele, et superventsus ei ole üldjuhul kontseptuaalne seos (Hellman ja Thompson 1975); kuid ei ole selge, kas redutseeritavuse vastu leidub häid argumente, mis töötaksid hästi ka loogilise superventsuse vastu. Viiendaks on mõnikord püütud teadvuselamuse fenomenile endale viidates näidata, et superventsussuhted ei saa üldiselt olla loogilised (Seager 1988); meie aga nägime, et teadvuselamus on peaaegu unikaalne näide loogilise superventsuse puudumisest. Ja lõpuks on väidet, et superventsussuhted ei ole üldjuhul loogilised, sageli esitatud ilma ühegi põhjenduseta, nähtavasti arvates, et iga mõistlik inimene peabki nii arvama (Bacon 1986; Heil 1992).[[134]](#footnote-135)

Usutavasti on kõrgtaseme omaduste superventsus füüsikaliste omaduste suhtes viimselt alati kas 1) primaarset või sekundaarset liiki loogiline superventsussuhe, või 2) kontingentne looduslik superventsussuhe. Kui mõne suhte puhul, mida tahetakse pidada superventsuseks, selgub, et see ei mahu kummagi juhu alla, siis on põhjust uskuda, et kõnealust liiki objektiivseid kõrgtaseme fakte ei ole üldse olemas. (Moraalifaktidega võibki see nii olla.) Allpool, ​4. peatükis​ püüan tuua veelgi argumente näitamaks, et pole olemas ühtki olulist superventsuse liiki, mis oleks loogilise ja loodusliku superventsuse vahevorm.

Põhimõtteliselt annabki see kõik meile ühtse seletusliku pildi. Peaaegu iga fenomen on ülalkirjeldatud nõrgas mõttes reduktiivselt seletatav; erandiks on teadvuselamused ja võib-olla ka indeksikaalsus ühes mikrofüüsikaliste alusfaktidega ja seadustega, mida tuleb tunnistada fundamentaalseteks.

Siinkohal tasub hetkeks peatuda, et vastata küsimusele, mida on esitanud Blackburn (1985) ja Horgan (1993): kuidas seletada superventsussuhteid endid? Kui jutt on loogilisest superventsussuhtest mõiste primaarsete intensioonide alusel, siis tuleb selleks vaid sooritada mõiste asjakohane, s.t. näiteks funktsionaalne või struktuuriline analüüs, ning panna tähele, et füüsiliselt identsetes maailmades selle mõiste referendid ühtivad. Siin on superventsuskonditsionaal ise aprioorne kontseptuaalne tõde. Sekundaarsetele intensioonidele toetuva loogilise superventsussuhte seletamiseks ​**[lk.89]**​ paneme tähele, et mõiste primaarne intensioon osutab aktuaalses maailmas referendile, mis projitseerub (rigidiseerimise teel) invariantselt kõigisse füüsiliselt identsetesse maailmadesse. Seletamiseks piisab sel juhul aprioorsest kontseptuaalset analüüsist ühenduses aktuaalse maailma kontingentsete faktidega.[[135]](#footnote-136) Mis aga puutub pelgalt looduslikku superventsussuhtesse, siis see on ise kontingentne seadus. Parimal juhul saab teda seletada fundamentaalsemate seaduste kaudu; halvimal juhul on superventsusseadus ise fundamentaalne. Mõlemal juhul saab maailma reeglipärasid seletada fundamentaalseadustega, nagu seda tehakse füüsikas, ja igasugune seletamine lakkab, nagu ikka, fundamentaalseaduste endi juures. Nagu nägime, on pelgalt looduslik superventsus ontoloogiliselt kulukas, ja seega on meie õnn, et loogiline superventsus on reegel ja looduslik superventsus vaid erand.

# ​**[lk.91]**​ II OSA: Teadvuse mitteredutseeritavus

## ​**[lk.93]**​ 3. peatükk. Kas teadvust on võimalik reduktiivselt seletada?

### 1. Kas teadvus on füüsika suhtes loogiliselt superventne?

Peaaegu kõik maailma nähtused on seletatavad füüsika mõistetes, ja on loomulik oodata, et nii peaks saama seletada ka teadvust. Käesolevas peatükis püüan ma aga näidata, et teadvus lipsab reduktiivsete seletuste võrgust välja. Ainuüksi füüsikalistes mõistetes pole iial võimalik seletada teadvuselamuse ilmumist. See järeldus võib tunduda negatiivsena, kuid ta viib ka mõnede tugevate positiivsete tulemusteni, millest tuleb põhjalikumalt juttu edasistes peatükkides.

Demonstreerimaks reduktiivsete seletuste võimatust, tuleb näidata, et teadvus ei ole füüsika suhtes loogiliselt superventne. Põhimõtteliselt tuleb näidata, et teadvus ei ole superventne globaalselt — see tähendab, teda ei saa tuletada maailma mikrofüüsikaliste faktide täielikustki kogumist. Praktikas on kergem esitada seda mõttekäiku lokaalsel kujul — näidata, et mitte ühegi indiviidi teadvuse faktid ei ole tuletatavad mikrofüüsikalistest faktidest. Võib arvata, et teadvuse vallas on globaalsel ja lokaalsel superventsusel ühesugune saatus, mistõttu pole kuigi oluline, kumbapidi me oma argumendi üles ehitame: kui teadvus on üldse superventne, siis on ta peaaegu kindlasti ka lokaalselt superventne. Kui keegi püüab seda kahtluse alla seada, piisab lihtsatest ümbersõnastustest kõigis argumentides, nii et nad kehtiksid globaalsel tasemel.

Kuidas näidata, et teadvus ei ole füüsika suhtes loogiliselt superventne? Selleks on mitmesuguseid võimalusi. Me võime uurida, mis on üldse mõeldav, ning jõuda otsesel teel järeldusele, et olukord, kus füüsikalised faktid oleksid samad, elamuse faktid aga teistsugused, on loogiliselt võimalik. Me võime ka apelleerida epistemoloogiale ning püüda näidata, et ​**[lk.94]**​ pole olemas sobivat liiki ühenduslüli, mis viiks teadmistelt füüsikaliste faktide kohta teadmisteni teadvuse kohta. Ja lõpuks, me võime apelleerida otseselt teadvuse mõistele ja näidata, et ükski selle mõiste analüüs ei anna mingit alust fenomenaalsete faktide tuletamiseks füüsikast. Järgnevas püüan ma oma seisukohta põhjendada kõigi nende kolme strateegia varal. Põhimõtteliselt toetuvad kaks esimest argumenti mõeldavusele, kaks järgmist epistemoloogiale ja viies analüüsile. Need viis argumenti dubleerivad üksteist mõnevõrra, kuid annavad minu vaatele üheskoos tugevat kinnitust.

Võib talitada ka sirgjoonelisemalt: võib põhjendada reduktiivse seletamise võimatust, mainimatagi seejuures otsesõnu loogilist superventsust. Olen valinud selle tee juba teisal, siin aga annan üksikasjalikuma analüüsi, et mõttekäik saaks täielikum. Sellegipoolest peaksid reduktiivse seletamise vastased argumendid ja seniste reduktiivsete seletuste kriitika (allpool, [2. jaos](#ORG.3.2)) olema arusaadavad ka ilma selle analüüsita. Lugeja võibki liikuda kohe sinna edasi, vähemalt esmakordsel lugemisel.

(Tehniline märkus: käesoleva peatüki ülesanne on püüda tõestada, et füüsikalistest faktidest ei ole võimalik ​*a priori* tuletada fenomenaalseid fakte. Paratamatus, millele vastav superventsussuhe toetub, on loogilise paratamatuse aprioorne versioon, kus kesksel kohal on primaarsed intensioonid. Nagu ​2. peatükis nägime, on just niisugune suhe relevantne, kui kõne all on seletamine; nüansid, mis kerkivad seoses aposterioorse paratamatusega, võib kõrvale jätta. Järgmises peatükis, kus seletuslike küsimuste asemel tõusevad esiplaanile ontoloogilised küsimused, püüan ma sõltumatult näidata, et füüsikaliste faktide ja fenomenaalsete faktide vahel pole ka mingit aposterioorselt paratamatut seost.)

#### Esimene argument: sombide loogiline võimalikkus

Kõige ilmsem (kuigi mitte ainus) viis teadvuse loogilise superventsuse uurimiseks on vaadelda *sombide* loogilist võimalikkust. (Sombiks nimetatakse kedagi või midagi, kes​/​mis on minuga (või mõne teise teadvusliku olendiga) füüsikaliselt identne, kuid kellel​/​millel teadvuselamused täielikult puuduvad.[[136]](#footnote-137)) Globaalsel tasandil võime kaaluda *sombimaailma* loogilist võimalikkust — maailma, mis on meie omaga füüsikaliselt identne, ent kus teadvuselamused täielikult puuduvad. Niisuguses maailmas on kõik sombid.

Vaatame siis minu sombist teisikut. See olend on minu molekulhaavaline koopia, minuga identne kõigilt madaltaseme omadustelt, mida postuleeriks lõpuleviidud füüsika, kuid teadvuselamused puuduvad tal täielikult.[[137]](#footnote-138) Piltlikkuse mõttes võime kujutleda, et ma vaatan praegusel hetkel aknast välja, naudin rohelist elamust, mida pakuvad puud akna taga, mälun šokolaaditahvlit, mis annab mulle meeldivaid maitseelamusi, ning tunnen oma paremas õlas tuima kipitust.

​**[lk.95]**​ Mis toimub minu sombist teisikus? Ta on minuga füüsikaliselt identne ja me võime väga hästi oletada, et ta viibib ka identses keskkonnas. Kindlasti on ta minuga identne funktsionaalselt: ta töötleb samasugust informatsiooni, reageerib samamoodi stiimulitele, mis muudavad samamoodi ka tema seesmisi konfiguratsioone, ja tulemuseks on täpselt samasugune käitumine. Ta on minuga identne psühholoogiliselt — selles mõttes, millest oli juttu [1. peatükis](#PARA.22-2). Ta tajub — funktsionaalses mõttes — puid akna taga, ja tunneb — psühholoogilises mõttes — šokolaadi maitset. Tänu psühholoogiliste mõistete funktsionaalsele analüüsitavusele tuleneb see kõik loogiliselt tema füüsikalisest identsusest minuga. Tal on ülalkirjeldatud funktsionaalses mõttes isegi “teadvus”  — ta on ärkvel, suudab oma seesmiste seisundite sisu kohta ütlusi anda, oskab koondada tähelepanu siia või sinna, jne. Kuid mitte ühegagi neist protsessidest ei kaasne reaalset teadvuselamust. Puudub igasugune fenomenaalne tunne. Ei ole olemas tunnet olla sombi.

Seesugune sombi on väga erinev sombidest, keda võib näha Hollywoodi filmides ja kellel on enamasti tõsiseid funktsionaalseid kahjustusi (joonis 3.1). Kõige ilmsemini jääb Hollywoodi sombidel vajaka teadvusest selle sõna psühholoogilises mõttes: tavaliselt on neil puudulik introspektsioonivõime ja nad ei suuda oma käitumist kuigi hästi tahteliselt juhtida. Fenomenaalne teadvus võib neil puududa, kuid võib ka olemas olla; nagu osutab Block (1995), on mõistlik oletada, et on olemas maitse, mida nad tunnevad oma ohvreid süües. Me võime nimetada neid *psühholoogilisteks sombideks*​; mind aga huvitavad *fenomenaalsed sombid*​, kes on on füüsikaliselt ja funktsionaalselt identsed, kuid kel puudub elamus. (Võib-olla polegi ime, et fenomenaalsed sombid pole olnud Hollywoodis kuigi populaarsed, sest nende kujutamisega oleks arusaadavatel põhjustel raskusi.)

​**[lk.95]**​ **Joonis 3.1.** Calvin ja Hobbes sombidest. (Calvin and Hobbes © Watterson. Distributed by Universal Press Syndicate. Reprinted with permission. All rights reserved)​

Aga tõelisi sombisid ei aja teineteise vaatamine kunagi naerma.

​**[lk.96]**​ Ettekujutus sombidest, nagu ma neid äsja kirjeldasin, on kaunis kummaline. Kõigepealt on ebatõenäoline, et sombid oleksid looduslikult võimalikud. Reaalses maailmas oleks igal minu koopial arvatavasti teadvus. Sel põhjusel on väga loomulik kujutleda teadvuseta olendeid teadvusega olenditest füüsiliselt erinevatena — neil võiksid olla näiteks käitumishäired. Kuid küsimus ei ole selles, kas sombide olemasolu meie maailmas on usutav, ega isegi mitte selles, kas idee inimese koopiast, kes oleks sombi, on loomulik; küsimus on selles, kas sombi mõiste on kontseptuaalselt sidus. Ja piisab juba ainuüksi selle mõiste arusaadavusest, et mõista, et ta seda on.

Loogilist võimalikkust ei saa põhjendada päris otsesel teel. Kuidas näiteks tõestada, et miilikõrgune monoratas on loogiliselt võimalik? See näib olevat lihtsalt ilmselge. Kuigi reaalses maailmas midagi niisugust ei eksisteeri, ei ole meil kahtlust kirjelduse loogilises sidususes. Kui keegi vaidleb vastu, et niisugune asi ei ole loogiliselt võimalik, vaid see ainult tundub meile nii, siis pole meil palju muud öelda kui korrata kirjeldust ja kinnitada, et sidusus on ilmne. Tundub olevat täiesti selge, et kirjelduses ei peitu mingeid märkamata jäänud vastuolusid.

Tunnistan, et sombide loogiline võimalikkus tundub mulle sama ilmsena. Sombi on lihtsalt miski, mis on minuga füüsiliselt identne, kuid ilma teadvuselamusteta — tema sisemuses on kõik pime. Kuigi see ei ole arvatavasti empiiriliselt võimalik, pole raske uskuda, et kirjeldatud olukord on koherentne; ma ei suuda leida selles mingit vastuolu. Mõnes mõttes taandub väide, et selline olukord on loogiliselt võimalik, toorele intuitsioonile, kuid mitte rohkem kui monoratta näite puhul. Mulle tundub, et peaaegu igaüks on võimeline niisugust asja ette kujutama. Mõni autor leiab võib-olla, et seda võimalust tuleb eitada lihtsalt selleks, et mõnd teooriat klappima saada, kuid selliste teooriate põhjendused peaksid toetuma küsimusele võimalikkusest, mitte vastupidi.

Üldjuhul lasub mõningane tõendamiskohustus[[138]](#footnote-139) sellel, kes väidab, et üks või teine kirjeldus on loogiliselt võimatu. Kui keegi tõesti usub, et miilikõrgune monoratas on loogiliselt võimatu, siis peab ta meile näitama, kus on selles idees otsene või kaudne vastuolu. Kui ta ei suuda osutada mõistete “miilikõrgune” ja “monoratas” intensioonides millelegi, mis võiks viia vastuoluni, siis ei ole tema väide veenev. Samas ei ole veenvam ka kõnealuste mõistete ilmselgelt ekslik analüüs — näiteks kui keegi kinnitaks, et selleks, et mingit objekti saaks pidada monorattaks, peab see olema lühem kui Vabadussammas. Kui ükski kõnealuste mõistete mõistlik analüüs ei leia vastuolu või ei tee vastuolulisust vähemasti tõenäoliseks, siis on loomulik jääda pooldama loogilist võimalikkust.

Siiski on loogilise võimalikkuse pooldajatel oma positsiooni toetamiseks ka positiivseid võimalusi. Nad võivad esitada mitmesuguseid kaudseid argumente, apelleerides meie teadmistele kõnealustest fenomenidest või kirjeldades, kuidas me mõtleme hüpoteetilistest juhtudest, kus mingi niisugune fenomen figureerib, ​**[lk.97]**​ ning seega näidates, et see, mida nad peavad ilmseks loogiliseks võimalikkuseks, on tõepoolest loogiline võimalikkus ja et see on tõepoolest ilmne. Võib veeretada fantaasiaid, kus tavaline inimene sõidab monorattaga ja korraga paisub kogu süsteem tuhat korda suuremaks, või kirjeldada monorataste rida, kus iga monoratas on eelmisest suurem. Mõnes mõttes apelleerivad kõik need võimalused intuitsioonile ja oponent võib kõigil juhtudel väita, et intuitsioon on meid eksitanud, kuid kirjelduse ilmselgus töötab meie kasuks ja aitab tõendamiskohustust veelgi enam vastase poolele veeretada.

Sombide loogilise võimalikkuse kaudseks põhjendamiseks võib näiteks vaadelda minu funktsionaalse organisatsiooni *mittestandardseid realiseeringuid*​.[[139]](#footnote-140) Minu funktsionaalne organisatsioon  — põhjusliku organisatsiooni muster, mida realiseerivate mehhanismide töö produtseerib minu käitumise — võib põhimõtteliselt realiseeruda väga mitmesugustel kummalistel viisidel. Võtame tuntud näite (Block 1978): mõne suure riigi, näiteks Hiina elanikkonna saaks organiseerida nii, et see realiseeriks põhjusliku organisatsiooni, mis on isomorfne minu aju põhjusliku organisatsiooniga; seejuures simuleeriks iga inimene ühe neuroni käitumist, sünapsitele vastaksid raadioühendused. Nii saaks elanikkond juhtida sensoorsete andurite ja motoorsete täituritega robotkeha.

Paljud autorid peavad ebatõenäoliseks, et säärases süsteemis võiks sündida teadvuselamus — et selles ilmuks mingil moel “kollektiivne vaim”. Ma ei hakka siinkohal arutama, kas teadvuselamused selles süsteemis tõepoolest tekiksid või mitte; ma ise oletan, et tekiksid, ja püüan seda näidata [7. peatükis](#ORG.7). Oluline on esialgu ainult see, et mõte teadvuselamuse puudumisest sellisel süsteemil ei vii vastuoluni. Jutt on mõttekast võimalusest; küsimus, kas teadvus süsteemis tõepoolest ilmub või mitte, jääb lahtiseks. Analoogilisele tulemusele jõudmiseks võib vaadelda ka minu ränist isomorfi, kelle organisatsioon ühtiks minu omaga, kuid neuronite asemel oleksid ränikiibid. On vaieldav, kas sellisel isomorfil oleks tõepoolest teadvus, kuid enamiku autorite arvates väljendavad selle võimaluse eitajad loogiliselt koherentset võimalust. Selletaolistest näidetest võib järeldada, et minu teadvuselamuse olemasolu ei ole loogiliselt tuletatav minu funktsionaalse organisatsiooni faktidest.

Ent kui tunnistada teadvuselamuse puudumise võimalikkus grupivaimuga struktuuril või minu ränist isomorfil kontseptuaalselt koherentseks, siis järeldub sellest, et minu sombist teisiku võimalikkus on samavõrra koherentne. On ju selge, et kontseptuaalselt ei ole teadvus sugugi hõlpsamini tuletatav biokeemiast kui ränist või homunkuluste kollektiivist. Kui teadvuselamuseta ränist isomorf on mõeldav, siis tarvitseb sellekohases kujutluses vaid panna räni asemele neuronid, funktsionaalset organisatsiooni samaks jättes, ja olemegi saanud minu sombist teisiku. Selles asenduses ei ole midagi, mis nõuaks, et kujuteldaval olendil oleksid elamused; need implementatsioonilised erinevused ei saa ju kuidagi kontseptuaalselt mõjutada elamust ennast. Seega ei ole teadvus füüsika suhtes loogiliselt superventne.

​**[lk.98]**​ Sombide võimalikkuse kasuks saab argumenteerida ka ilma mittestandardsete realiseeringuteta, kuid viimaste heuristiline väärtus on selles, et nad vabastavad meid ühest kontseptuaalsete segaduste allikast. Teadvuseta füüsilise teisiku loogiline võimalikkus ei tarvitse olla igaühele esialgu intuitiivselt päris ilmne — võib-olla sellepärast, et harjumus seostada teadvust biokeemiaga on viinud neid mõttele, et nende kahe vahel on kontseptuaalne seos. Vähem harjumuspäraste näidete puhul langevad need empiirilised korrelatsioonid ära, mistõttu otsuseid loogilise võimalikkuse kohta saab teha sirgjoonelisemalt.[[140]](#footnote-141) Ent niipea kui me tunnistame funktsionaalsed koopiad, kel teadvus puudub, loogiliselt võimalikuks, ei ole enam võimalik vältida analoogilist järeldust ka füüsilise koopia puhul.

Mõeldavusargumente võidakse pidada ebausaldusväärseteks. Näiteks on väidetud, et me ei suuda tegelikkuses üksikasjalikult ette kujutada inimaju miljardeid neuroneid. See on muidugi tõsi, kuid meie argument ei nõuagi, et me peaksime kõiki neuroneid endale ükshaaval ette kujutama. Teadvust ei saa kontseptuaalselt tuletada ainuüksi neuronstruktuuri keerukusest; kui sel struktuuril on mingi seos teadvusega, siis peab see tulenema mõnest kõrgtaseme omadusest, mille ta võimalikuks teeb. Seega tarvitseb meil süsteemi vaid laias laastus ette kujutada ja näha, et meie kujutlus sisaldab vajaliku keerukusastmega mehhanisme tajumiseks, liigitamiseks, suure läbilaskevõimega infotöötluseks, raporteerimiseks ja muuks taoliseks. Ükskõik kui keerukatena me neid mehhanisme kujutleme, sombidega stsenaarium jääb ikka ühevõrra koherentseks. Oponent võib muidugi väita, et kõik neuronstruktuuri üksikasjad, mida me ei suutnud kujutleda, on kuidagi kontseptuaalselt olulised sõltumatult nende panusest süsteemi töösse; kuid säärasel juhul võlgneb ta meile selgituse, mis nad nii oluliseks teeb — ja seda ta anda ei saa. Need implementatsioonilised üksikasjad lihtsalt ei kuulu sellele tasemele, kus nad võiksid teadvuse jaoks kontseptuaalselt olulised olla.

Mõnikord öeldakse ka, et mõeldavus ei saa olla usaldusväärne võimalikkuse näitaja​. Peamine viis, kuidas mõeldavus ja võimalikkus võivad teineteisest lahku minna, tuleneb aposterioorse paratamatuse fenomenist: näiteks hüpoteesis, et vesi ei ole H₂O, ei tundu kontseptuaalselt olevat mingit vastuolu, kuid samas on vesi kõigis võimalikes maailmades üsna kindlasti H₂O. Kuid aposterioorne paratamatus ei puutu siinse peatüki probleemidesse. Nagu eelmises peatükis nägime, on seletuslike seoste aluseks kõrgtaseme faktide aprioorne tuletatavus füüsika faktidest. Vastavat liiki võimalikkuse hindamisel ei tule lähtuda mitte asjaomaste mõistete sekundaarsetest intensioonidest, mis mängivad rolli aposterioorses paratamatuses, vaid primaarsetest. Nii et isegi kui sombimaailm on mõeldav ainult samas mõttes, nagu on mõeldav, et vesi ei oleks H₂O, piisab juba sellestki tõestamaks, et teadvus ei ole reduktiivselt seletatav.

Kui need kaalutlused kõrvale jätta, siis võivad mõeldavusargumendid ekslikeks osutuda peamiselt raskesti tabatava kontseptuaalse segaduse tõttu: ebapiisava ​**[lk.99]**​ refleksiooni korral võib meil jääda märkamata koherentsuse puudumine selles, mida me pidasime võimalikuks, kuna me kirjeldasime kujuteldud olukorda valesti. Näiteks võib keegi arvata, et ta suudab ette kujutada olukorda, kus Fermat’ viimane teoreem on väär; selleks kujutleks ta olukorda, kus juhtivad matemaatikud teatavad, et nad on leidnud vastunäite. Ent juhul, kui teoreem on tegelikult tõene, on olukorra kirjeldus vale; tegelikult on kirjeldatud olukorda, kus Fermat’ viimane teoreem on tõene ja mõned matemaatikud eksivad. Oluline on aga, et sedaliiki vead on alati aprioorsed, sest nende allikaks on meie mõistete primaarsete intensioonide ekslik rakendamine kujuteldud olukorrale. Korralikult järele mõeldes selgub, et mõisteid on rakendatud valesti ning et väide, nagu oleks vastav olukord loogilisest võimalik, on alusetu.

Seega jääb meie vastastele siin üle üksainus võimalus: öelda, et sombimaailma sombimaailmana kirjeldades oleme me mõisteid ekslikult kohaldanud ning et meie kirjelduses peitub tõepoolest kontseptuaalne vastuolu. Piisava selgusega järele mõeldes mõistaksime, et kujutledes füüsiliselt identset maailma, kujutleme me automaatselt maailma, kus teadvuselamused on olemas. Kuid sellisel juhul on juba oponendi kohus anda meile mingi ettekujutus sellest, kus võiks meie pealtnäha täiesti sidusas kirjelduses peituda vastuolu. Kui ta ei suuda leida ühtki seesmist lahkheli, siis on see väga tugev argument sombimaailma loogilise võimalikkuse kasuks.

Nagu ennegi, ei näe ma siin mingit seesmist vastuolu; mul on selge pilt sellest, mida ma nimelt kujutlen, kujutledes sombit. On siiski ka autoreid, kes leiavad, et mõeldavusargumente on raske hinnata, eriti kui jutt on säärastest kummalistest ideedest. Seepärast oleme õnnega koos, et kõiki argumente, kus figureerivad sombid, on võimalik esitada ka ilma sombideta, uurides asja näiteks epistemoloogia või analüüsi küljelt. Paljudele autoritele tunduvad viimasena mainitud liiki argumendid (nagu näiteks alltoodud 3.–5. argument allpool) sirgjoonelisematena ja seega kaalukama alusena loogilise superventsuse eitamiseks. Kuid sombid annavad meile vähemalt võimaluse kujukalt illustreerida olulisi asjasse puutuvaid küsimusi.

#### Teine argument: invertspekter

Isegi loogilise superventsuse vastu mõeldavusargumenti esitades ei ole rangelt võttes tingimata tarvilik eeldada sombide või sombimaailma loogilist võimalikkust. Piisab niisuguse maailma loogilisest võimalikkusest, mis oleks füüsiliselt identne meie maailmaga ja kus ei tarvitse täielikult puududa ka teadvuselamus, ent teadvuselamuse faktid erineksid vastavatest faktidest meie maailmas. Kui leidub positiivne fakt teadvuselamuse kohta, mis peab paika meie maailmas, kuid mitte sellega füüsiliselt identses maailmas, siis ei ole teadvus loogiliselt superventne.

​**[lk.100]**​ Seega piisab nentimast, et me võime koherentselt kujutleda füüsiliselt identset, kuid invertsete teadvuselamustega maailma, või (lokaalsel tasandil) olendit, kes on minuga füüsiliselt identne, kuid kelle teadvuselamused on invertsed. Võib näiteks kujutleda, et seal, kus minul on punane elamus, on minu invertteisikul sinine elamus, ja ümberpöördult. Loomulikult nimetab ta oma siniseid elamusi punasteks, kuid see ei puutu asjasse. Oluline on see, et asjad, mida me mõlemad nimetame punasteks — veri, tuletõrjeautod jne. — kutsuvad temas esile samasuguse elamuse nagu minul asjad, mida me mõlemad nimetame sinisteks, näiteks meri ja taevas.

Analoogiliselt punase–sinise inversiooniga on ka tema ülejäänud värvuselamused minu omade suhtes süstemaatiliselt invertsed. Seda on vahest kõige parem ette kujutada meie kolmemõõtmelise värvusruumi kahe telje niisuguse ümbervahetusena, et seal, kus meil on punase–rohelise telg, on temal kollase–sinise telg, ja ümberpöördult.[[141]](#footnote-142) Säärase inversiooni saavutamiseks aktuaalses maailmas tuleks arvatavasti närviprotsessid mingil sobival viisil ümber korraldada. Ent kui pidada silmas loogilist võimalikkust, siis tundub oletus, et täpselt sama füüsilise struktuuri juures võivad elamused olla invertsed, igati koherentsena. Neurofüsioloogias pole midagi, mis dikteeriks, et mingit kindlat liiki protsessidega peaksid kaasnema just punased, aga mitte kollased elamused.

Mõnikord vaieldakse sellele vastu (Harrison 1973; Hardin 1987), öeldes, et inimese värvusruum on asümmeetriline viisil, mis säärast inversiooni ei võimalda. Näiteks assotsieerub mõne värvusega soojus või jahedus, need aga seostuvad otseselt mitmesuguste funktsionaalsete rollidega (näiteks tajutakse soojust kui midagi “positiivset” ja külmust kui “negatiivset”). Sooja ja jaheda värvuse ümbervahetamine irrutaks “sooja” fenomenaalne tunne soojuse funktsionaalsest rollist: “jahe” roheline elamus tunduks pigem positiivse kui negatiivsena jne. Tundub ka, et punasel värvusel on eristatavaid varjundeid rohkem kui kollasel, mistõttu punase ja kollase elamuse ümbervahetamine võib tuua kaasa kummalise olukorra, kus subjekt suudab funktsionaalselt eristada rohkem kollase varjundeid, kui see fenomenoloogiliselt võimalik oleks. Võib-olla on asümmeetriate hulk värvusruumis nii suur, et iga seesugune inversioon tooks kaasa fenomenaalse tunde kummalise irdumise oma “õigest” funktsionaalsest rollist.

Sellele argumendile võib anda kolm vastust. Esiteks ei tundu ettekujutuses sellisest dissotsiatsioonist (näiteks jahedast fenomenoloogiast ühenduses soojade reaktsioonidega), nii kummaline kui see mööndavasti ongi, olevat mingit vastuolu.[[142]](#footnote-143) Teiseks ei tarvitse punane värvus olla täpses vastavuses sinisega ja ümberpöördult: võib ka ette kujutada, et neile vastaksid pisut teistsugused värvused. Näiteks võib punasega olla vastavuses sinise värvuse “soe” erikuju (nagu soovitab Levine (1991)) või isegi mõni niisugune värvus, mida meie värvusruumis ei olegi. Mis puutub punasesse ja kollasesse, ​**[lk.101]**​ siis võib punasele vastata kollaste elamuste laiem vahemik, kus on võimalik teha rohkem eristusi. Ei ole mingit põhjust võtta invertspektrisse värvusi ainult meie tavalisest värvusruumist. Kolmas ja vahest kõige veenvam vastus on öelda (koos Shoemakeriga (1982)), et olgugi meie endi värvusruum asümmeetriline, on kindlasti võimalikud ka sümmeetrilise värvusruumiga olendid. Näiteks olend, kes näeks (ja elaks läbi) täpselt kahte värvust ​*A* ja ​*B*​, mis vastaksid kahele selgelt erinevale lainepikkusvahemikule, ja kelle värvusruum ammendub nende kahe värvusega, oleks arvatavasti looduslikult võimalik. Ei tundu olevat mingit vastuolu kujutluses kahest olendist, kes oleksid füüsiliselt identsed, kuid kelle elamused värvustest ​*A* ja ​*B* oleksid teineteise suhtes invertsed. Ja sellest meie järelduse kinnituseks piisab.

Isegi paljud reduktiivsed materialistid (näiteks Shoemaker (1982) on möönnud, et ettekujutuses värvuselamuste inverteerumisest funktsionaalse organisatsiooni endiseks jäädes pole mingit vastuolu. Pole midagi võimatut selles, et süsteemil, mille neurofüsioloogilised omadused oleksid teistsugused või millel kogu neurobioloogiat asendaks räni, oleksid teistsugused värvuselamused. Aga niipea kui see võimalikuks tunnistada, järeldub siit automaatselt, et idees elamuste inversioonist füüsilises koopias ei ole vähemasti kontseptuaalselt mingit vastuolu. Need neurofüsioloogilised omadused, mis sellisel juhul realiseerumata jäävad​[[143]](#footnote-144), ei ole samuti midagi niisugust, mis võiks elamuse iseloomu loogiliselt determineerida. Isegi kui mingeid neurofüsioloogilisi struktuure saabki kuidagi aposterioorselt samastada mingite kindlate elamustega (nagu usub Shoemaker), tuleb ikkagi tunnistada teistsuguste seosemustrite mõeldavust, kui mõista mõeldavust reduktiivsele seletamisele relevantses mõttes.

Kuigi nii invertspektrite kui ka sombide võimalikkus kinnitavad, et teadvus ei ole loogiliselt superventne, võimaldab esimene jõuda rangelt nõrgemale järeldusele kui teine: on mõeldav pidada loogiliselt võimalikuks invertspektreid, kuid mitte sombisid. Kui asi olekski tegelikult nii, siis oleks võimalik reduktiivselt seletada teadvuse olemasolu, kuid mitte konkreetsete teadvuselamuste spetsiifilist iseloomu.

#### Kolmas argument: episteemiline asümmeetria

Nagu [ülal](#PARA.8-8) nägime, on teadvuses kui universumi omaduses midagi üllatavat. Meil pole teadvusse uskumiseks mingit alust peale omaenda teadvuselamuse. Isegi kui me tunneksime universumi füüsikat viimsegi üksikasjani — kogu konfiguratsiooni, kõiki põhjuslikke seoseid, kõigi väljade ja osakeste liikumisi aegruumilises muutkonnas —, ei oleks kogu selles informatsioonis midagi, mis võiks viia meid mõttele postuleerida teadvuselamuse olemasolu. Minu kõige autoriteetsem teadmine teadvuse kohta ​**[lk.102]**​ pärineb minult endalt, mitte välistest vaatlustest. Probleemi sunnib mulle peale minu esimese isiku teadvuselamus.

Kui meie käsutuses oleksid kõik madaltaseme faktid füüsikaliste konfiguratsioonide ja põhjuslike seoste kohta, siis saaksime neist põhimõtteliselt tuletada kõikvõimalikud kõrgtaseme faktid makroskoopiliste süsteemide, nende organisatsiooni ja nendevaheliste põhjuslike seoste kohta. Nii saaks kindlaks teha kõik faktid inimese bioloogilise talitluse, samuti tema käitumise ja seda sünnitavate ajumehhanismide kohta. Kuid selles tohutus põhjuslikus pajatuses ei oleks midagi, mis võiks panna uskuma teadvuse olemasolusse seda, kes pole teadvust ise otseselt läbi elanud. Juba selline mõtegi oleks ebamõistlik, kui mitte lausa müstiline.

On tõsi, et maailma füüsikalised faktid võivad anda mõningaid kaudseid tõendeid teadvuse olemasolu kohta. Näiteks saaks nende faktide põhjal kindlaks teha, et on palju organisme, mis ütlevat endil olevat teadvuse ning kinnitavad, et neil on mingid salapärased subjektiivsed elamused. Kuid need oleksid üsna ebaveenvad tõendid, ja kõige loomulikum oleks vististi teha eliminativistlik järeldus — et tegelikult ei ole neil olenditel mingit teadvuselamust, vaid ainult palju sõnu.

Eliminativism on teadvuselamuse küsimuses ebamõistlik positsioon ühelainsal põhjusel: meil on see elamus endil olemas. Kui seda otsest teadmist ei oleks, võiks teadvus minna elujõu teed. Teisiti öeldes, meie teadmistes teadvuse kohta *episteemiline asümmeetria*​, mida pole meie teadmistes teiste fenomenide kohta.[[144]](#footnote-145) Meie teadmine teadvuselamuse olemasolust pärineb eeskätt meilt endilt, välistel tõenditel on parimal juhul teisejärguline roll.

Selle mõtte selgitamiseks võib osutada ka *teiste vaimude probleemile*​. Isegi kui meil oleksid teiste olendite kohta kõik füüsikalised teadmised, ei ole meil ikkagi võimalik kindlalt teada, et neil on teadvus või missugused on nende elamused (kuigi meil on võib-olla põhjust uskuda elamuste olemasolusse). Silmatorkaval kombel pole olemas analoogilist “teiste elude” või “teiste majanduste” või “teiste kõrguste” probleemi. Nendes juhtudes episteemilist asümmeetriat ei ole, ja seda just sellepärast, et need fenomenid on füüsika suhtes loogiliselt superventsed.

Episteemiline asümmeetria teadmistes teadvusest näitab selgesti, et teadvus ei saa olla loogiliselt superventne. Vastasel korral niisugust episteemilist asümmeetriat ei oleks; loogiliselt superventsed omadused on välise tõendusmaterjali põhjal hõlpsasti tuvastatavad ja esimese isiku juht ei mängi nende puhul mingit erilist rolli. Muidugi on mõnd superventset omadust, võib-olla näiteks mälu, hõlpsam tuvastatada esimese isiku positsioonilt. Kuid niisugustel juhtudel on küsimus ainult vajaliku töö mahus. Põhimõtteliselt on mälu olemasolu ühtviisi tuvastatav nii kolmanda kui esimese isiku positsioonilt. Teadvus toob kaasa palju fundamentaalsema episteemilise asümmeetria: ühestki ​**[lk.103]**​ faktide kogumist keerukate põhjuslike seoste kohta füüsikalistes süsteemides ei saa tuletada ainsatki fakti teadvuse kohta.

#### Neljas argument: teadmine

Kõige ilmekama argumendi teadvuse loogilise superventsuse vastu on Nageli (1974) ja teiste autorite samalaadseid mõtteid edasi arendades välja pakkunud Jackson (1982). Kujutage ette, et me elame lõpuleviidud neuroteaduse ajastul: me teame ajus kulgevate füüsikaliste protsesside kohta, mis meie käitumist juhivad, kõike, mida neist üldse teada annab. Mary on üles kasvanud must-valges toas ja pole kordagi näinud muid värve peale musta, valge ja mitmesugustes varjundites halli.[[145]](#footnote-146) Samas on ta üks maailma juhtivaid neuroteadlasi, kelle erialaks on värvusnägemise neurofüsioloogia. Visuaalse informatsiooni töötlemises osalevatest neuronprotsessidest, optiliste protsesside füüsikast ja keskkonna objektide füüsilisest struktuurist teab ta kõike, mida teada annab. Kuid ta ei tea, mis tunne on näha punast. Mitte ükski füüsikalistest faktidest lähtuv mõttekäik seda teadmist talle ei anna.

Sellest järeldub, et subjektiivsete värvuselamuste fakte ei saa tuletada füüsikalistest faktidest. Vastasel korral võiks Mary põhimõtteliselt kindlaks teha, mis tunne on näha punast, lähtudes oma teadmistest füüsikaliste faktide kohta. Kuid see pole võimalik. Võib-olla saaks Mary selle välja uurida mingi kaudse meetodi abil, näiteks oma ajuga sobival viisil manipuleerides. Argumendi mõte on aga selles, et see teadmine ei tulene ainuüksi füüsikalistest teadmistest. Kõigi füüsikaliste faktide tundmine võimaldab Maryl põhimõtteliselt tuletada kõik faktid süsteemi reaktsioonide, oskuste ja kognitiivsete võimete kohta, kuid punane elamus jääb talle täiesti tundmatuks.

Selle argumenti selgitamiseks võib vaadelda ka süsteeme, mis on meist täiesti erinevad, palju lihtsamad — näiteks hiiri või nahkhiiri —, ning panna tähele, et nende süsteemide füüsikalised faktid ei ütle meile, missugused on nende teadvuselamused, kui neid üldse on. (Nagel asetabki põhirõhu just sellistele küsimustele.) Isegi kui me teame mõne hiire kohta kõiki füüsikalisi fakte, jääb tema teadvuselamuste iseloom endiselt lahtiseks: füüsikaliste faktidega ei ole vastuolus ei see, et tal teadvuselamus on, kui ka, et tal seda ei ole. Füüsikaliste faktide põhjal nahkhiirte kohta võib kindlaks teha kõik faktid ühe konkreetse nahkhiire kohta — kõik peale faktide tema teadvuselamuste kohta. Ka kõiki füüsikalisi fakte teades ei tea me ikkagi, mis tunne on olla nahkhiir.

Võime vaadelda ka arvutit, mis on programmeeritud toimima lihtsa kognitiivse agendina (võib-olla umbes koera tasemel intelligentsusega), millel on meiega mõningaid sarnasusi: ta eristab oma tajusid samamoodi ning liigitab ka värvusstiimuleid üsna sarnaselt, paigutades asjad, mida meie nimetaksime “punaseks”, ühte kategooriasse, ja asjad, mida me nimetaksime “roheliseks”, teise. Isegi kui me tunneme arvuti elektroonikat viimse üksikasjani, jääb meile ikka kaks vastuseta küsimust: 1) kas arvutil on roose vaadates mingisuguseidki elamusi? 2) kui jah, siis kas tema värvusaistingu kvaliteet on seejuures sama mis meil roose vaadates, või hoopis teistsugune? Need on täiesti mõttekad küsimused, ja ka kõigi füüsikaliste faktide tundmine ei anna meile alust eelistada üht vastust teisele. Seega ei saa teadvuselamuse fakte loogiliselt tuletada füüsikalistest faktidest.

Jackson ise ei polemiseerinud oma argumendiga mitte teadvuse reduktiivse seletamise, vaid materialismi vastu. Tema argumendile on esitatud palju vastulauseid; ma võtan need arutuse alla järgmises peatükis, kus kõne all ei ole reduktiivne seletamine, vaid materialism. Siinkohal on aga huvitav märkida, et enamik materialismivastase argumendi vastaseid on tunnistanud õigeks ühe punkti, mis puudutab oluliselt ka reduktiivse seletamise vastast argumenti: et teadmine sellest, mismoodi punane värvus välja näeb, on faktiline teadmine, mis ei ole ​*a priori* tuletatav füüsikalistest faktidest. Sellest järeldusest on võimalik pääseda ainult ühel viisil: öeldes, et teadmine sellest, mismoodi punane värvus välja näeb, ei ole üldse faktiline teadmine. Nii ongi teinud Lewis (1990) ja Nemirow (1990), kes väidavad, et Maryl puudub vaid mingi võime, näiteks võime punaseid asju ära tunda. Ma võtan selle idee arutluse alla [järgmises peatükis](#ORG.4.3.1); siinkohal märgin vaid, et kui on ilmne, et esimest korda punast värvust nähes omandab Mary mingi teadmise maailma kohta, siis peaks sama ilmne olema ka see, et see teadmine on mingi fakti teadmine.

#### Viies argument: analüüsi puudumine

Kui reduktiivse seletamise pooldajad tahavad jätta endile vähegi lootust ülaltoodud argumente ümber lükata, peavad nad andma meile mingi ettekujutuse, kuidas võiks teadvuse olemasolu kuidagigi tuleneda füüsikalistest faktidest. Poleks õiglane nõuda neilt kõiki üksikasju, kuid nad peaksid vähemalt kirjeldama, kuidas selline tuletuskäik võiks välja näha. Kuid iga selline katse kukub paratamatult läbi. Selleks, et teadvus oleks tuletatav mingist füüsikaliste faktide kogumist, on tarvis teadvuse mõiste analüüsi — niisugust, mille realiseering võiks tuleneda füüsikalistest faktidest —, ja säärast analüüsi pole kusagilt võtta.

Ainus teadvuse analüüs, mis selleks otstarbeks vähegi sobida võiks, on funktsionaalne analüüs. Niisugune analüüs näitaks, et kui me ütleme mingil objektil olevat teadvuse, siis ei tähenda see midagi muud kui seda, et see objekt mängib teatavat funktsionaalset rolli. Näiteks võidakse öelda, et teadvus seisneb ainult sõnalises raporteeritavuses, ​**[lk.105]**​ tekib mingite tajueristuste tagajärjena, teeb informatsiooni mingil kindlal viisil kättesaadavaks hilisematele protsessidele vms. Kuid paistab, et analüüsidena kukuvad kõik need variandid haledalt läbi. Nad lihtsalt ei taba seda, mida teadvuselamus tähendab. Teadvusseisundid võivad mängida mitmesuguseid põhjuslikke rolle, kuid põhjuslikud rollid ei määratle neid; teadvusseisunditeks teeb nad nende juurde kuuluv fenomenaalne tunne, ja seda tunnet ei saa funktsionaalselt olematuks defineerida.

Et mõista, kui mitterahuldavad on kõik niisugused analüüsid, paneme tähele, kui lamedaks teevad nad teadvuse seletamise probleemi. Korraga selgub, et teadvuse seletamiseks piisab, kui seletada ära meie võime anda mingeid sõnalisi teateid, sooritada mingit liiki eristusi või ilmutada mõnd muud võimet. Tundub aga olevat täiesti mõeldav, et kõigi nende fenomenide seletamiseks ei ole tarvis seletada kõige vähematki teadvuse enda kohta, s.t. pole tarvis seletada teatamise või eristamisega kaasnevat elamust ennast. Teadvuse analüüsimine mingi funktsionaalse mõiste kaudu eeldab alati kas teemavahetust​[[146]](#footnote-147) või probleemi olematuks defineerimist. Samasuguse eduga võiks defineerida “rahu kogu maailmas” “singivõileivaks”. Niiviisi oleks rahu saavutamine kogu maailmas palju lihtsam, kuid see oleks vaid näiline saavutus.

Teadvuse funktsionaalseid analüüse võib vaidlustada ka konkreetsemate mõttekäikudega. Näiteks on igale funktsionaalselt analüüsitavale mõistele omane teatav semantiline määramatus. Kas hiirtel on uskumusi? Kas bakterid õpivad? Kas arvutiviirus on elus? Õige vastus niisugustele küsimustele kõlab tavaliselt nii: mingis mõtte jah, mingis mõttes ei. Kõik oleneb sellest, kuidas me mõisteid piiritleme, piirid aga on kõigil funktsionaalsetel kõrgtaseme mõistetel ebamäärased. Ent küsime nüüd võrdluseks: Kas hiirtel on teadvuselamusi? Aga viirustel? Need ei ole küsimused, mille vastus oleneks mõistete piiritlemisest. Tunne olla hiir kas on olemas või ei ole, ja me ei saa teha hiire elamust olevaks või olematuks mõisteid nii- või teistpidi defineerides. Muidugi moodustavad teadvuselamused arvatavasti kontiinumi, mis ulatub väga nõrkadest elamustest väga külluslikeni; ent kui millelgi on teadvuselamused, olgu need ükskõik kui tuhmid, siis ei saa neid olematuks postuleerida. Säärase selge alternatiivini ei ole võimalik jõuda teadvusega seonduvate mõistete funktsionaalse analüüsi teel, sest kõik säärased funktsionaalsed mõisted on mingil määral laialivalguvad. Aga kui see on nii, siis järeldubki siit, et teadvuse mõiste ei ole funktsionaalselt analüüsitav.

Teine vastuväide on, et funktsionalistlikus analüüsis läheb kaduma oluline erinevus teadlikkuse ja teadvuse mõiste vahel, millest oli juttu [1. peatükis](#ORG.1.5.2). Kui teadvus oleks funktsionaalselt analüüsitav, siis käiks see analüüs arvatavasti umbes nii, nagu me analüüsisime seal teadlikkust kui informatsiooni mingis mõttes kättesaadavust hilisemaks töötluseks ja käitumise juhtimiseks. Teadlikkus on väga hea mõiste, kuid teadvuselamuse mõistest on ta täiesti erinev. Funktsionalistlikus tõlgenduses langevad need kaks mõistet — teadvus ja teadlikkus — ühte, moonutades seega meie kontseptuaalset süsteemi.

​**[lk.106]**​ Funktsionaalse analüüsi alternatiivid näivad olevat veelgi halvemad. Pole sugugi kindel, kas mõnd teist reduktiivseks seletamiseks sobivat analüüsimisviisi leidubki. Kõne alla võiks tulla ehk veel struktuurne analüüs — mis püüaks teadvust analüüsida näiteks kui mingit biokeemilist struktuuri —, kuid niisugune analüüs oleks ilmselt veelgi puudulikum. Olenemata sellest, kas teadvus on biokeemiline struktuur või mitte, mõeldakse teadvuse all tegelikkuses ikkagi midagi muud. Teadvuse analüüsimine niisugusel viisil teeks seletamisprobleemi jällegi lamedaks ja hiiliks teemast mööda. Tundub, et teadvuse mõiste ei ole redutseeritav, vaid ainult kirjeldatav mõistete kaudu, mis eeldavad omakorda jällegi teadvust.

Pange tähele, et seesugune mitteredutseeritavus on täiesti erinev mitteredutseeritavusest, mis on, nagu mõnikord kinnitatakse, omane kõrgtaseme mõistetele üldse. Me oleme näinud, et paljusid kõrgtaseme mõisteid ei saagi selgepiiriliselt defineerida ning et neid ei saa analüüsida tarvilike ja piisavate tingimuste kaudu. Siiski saab nende kohta, nagu [eelmises peatükis](#ORG.2.5.0.3) nägime, anda vähemalt ligikaudsegi analüüsi, mis annab sageli enam-vähem õigeid tulemusi, kuigi hägustab vältimatult üksikasju. Kõige tähtsam on aga, et elu, õppimist jts. omadusi on, nagu kerge näha, võimalik analüüsida kui funktsionaalseid omadusi, kuigi ei tarvitse olla kerge sõnades seletada, missugused need omadused nimelt on. Isegi kui neil omadustel selged funktsionaalsed definitsioonid puuduvad, ühilduvad nad siiski üsna hästi tuletatavusega füüsikalistest faktidest.

Teadvuse probleemid kuuluvad hoopis teise kaalukategooriasse. Siin ei anna analüüsid, mida keegi pakkuda võiks, isegi mitte umbkaudselt õigeid tulemusi. Asi on hullemgi: nad ei suuda isegi kirjeldada seda, mida on tarvis seletada. Erinevalt analoogilistest olukordadest elu ja õppimisega ei tule siin kiusatustki püüda funktsionaalset analüüsi, mida teadvuse kohta välja pakutakse, epitsüklite lisamise teel paremaks teha. Teadvust ei ole lihtsalt võimalik iseloomustada kui funktsionaalset omadust. Sama kehtib ka teadvuse struktuursete või veel mingit muud liiki reduktiivsete analüüside kohta. Seega pole teadvuse tuletamist füüsikalistest faktidest isegi kusagilt alustada.

### 2. Reduktiivsete seletuste läbikukkumine

Teadvuse loogiline mittesuperventsus füüsika suhtes tähendab, et ühelgi katsel anda teadvusele reduktiivset seletust ei saa olla edu. Iga kord, kui keegi pakub teadvuse alusena välja mingi füüsikalise protsessi, kerkib küsimus: miks kaasneb nende protsessidega teadvuselamus? Enamiku teiste fenomenide puhul leidub analoogilisele küsimusele lihtne vastus: fenomeni olemasolu tuleneb nende protsesside füüsikalistest faktidest. Näiteks kui fenomeniks on bioloogiline elu, tuleneb füüsikalistest faktidest mingite kindlate funktsioonide talitlemine, ja nende funktsioonide talitlemisest ​**[lk.107]**​ piisab elu seletamiseks. Kuid teadvuse puhul samasugune vastus ei sobi.

Füüsikalised seletused sobivad hästi struktuuri ja funktsiooni seletamiseks. Struktuurseid ja funktsionaalseid omadusi saab otseselt tuletada füüsikalisest madaltaseme kirjeldusest, mistõttu nendele on ilmaselgelt sobiv kohaldada reduktiivseid seletusi. Ja peaaegu kõik kõrgtaseme fenomenid, mis nõuavad seletust, taanduvad viimselt ikka struktuurile või funktsioonile: mõelgem näiteks koskede, planeetide, seedimise, paljunemise või keele seletamisele. Kuid teadvuse seletamine ei taandu struktuuri ja funktsiooni seletamisele. Kui me oleme seletanud ära kõik ajuga seonduvad füüsikalised struktuurid ning aju kõikvõimalike funktsioonide talitluse, jääb üks *explanandum* ikka veel, nimelt teadvus ise. Miks peaks kogu sellest struktuurist ja kõigi nende funktsioonide talitlusest sündima elamus? Lugu füüsikalistest protsessidest meile seda ei ütle.

Seda mõtet võib selgitada ka ülalkirjeldatud mõtteliste eksperimentide abil. Iga lugu füüsikalistest protsessidest võib käia nii minu kui ka minu sombist teisiku kohta. Järelikult pole selles loos midagi, mis seletaks, miks minus sünnib teadvus, aga minu teisikus mitte. Analoogiliselt saab kõiki füüsikaliste protsesside kirjeldusi rakendada ka minu invertspektriga teisikule, kes näeb sinist seal, kus mina näen punast: järelikult pole selles kirjelduses midagi, mis seletaks, miks on minu elamus on just niisugune nagu minul, aga mitte nagu minu teisikul. Juba see, et olukord, kus füüsikalised faktid oleksid samad, teadvuse faktid aga erinevad, on loogiliselt võimalik, näitab meile, et füüsikalise tasandi ja teadvuselamuse vahel on, nagu Levine (1983) seda nimetab, *seletuslünk*​.

Kui see on õige, siis on teadvuse kaasnemine ühe või teise füüsikalise protsessiga uus fakt, mille seletamiseks ei piisa arutlemisest füüsikaliste faktide üle. Seda protsessi ja teadvuse kooslust tuleb pidada mingis mõttes lihtsalt elementaarseks faktiks. Me võime püüda neid elementaarseid fakte süstematiseerida ja seletada mingi lihtsa alusskeemi abil, kuid sellesse jääb alati element, mis on sellest füüsikalisest kirjeldusest loogiliselt sõltumatu. Midagi seletusetaolist võiks ehk välja tulla, kui kombineerida neid füüsikalisi alusfakte mingite lisaprintsiipidega, mis seoksid füüsikalised faktid teadvusega, kuid säärane seletus poleks enam reduktiivne. Juba vajadus neid siduvaid lisaprintsiipe tarvitusele võtta näitab, et me ei seleta teadvust enam reduktiivselt, vaid temast endast lähtudes.

Muidugi ei tahtnud ma selle kõigega öelda, nagu ei oleks teadvuse seletamisel füüsikaliste faktidega üldse midagi peale hakata. Me võime endiselt eeldada, et füüsikalised faktid mängivad teadvuseteoorias olulist rolli, andes meile näiteks informatsiooni teadvuse füüsikalise aluse kohta ja lubades võib-olla ka kujutada üksikasjalikult ette vastavust füüsiliste protsesside ja teadvuselamuse mitmesuguste aspektide vahel. Sellised kirjeldused võivad olla eriti kasulikud teadvuse struktuuri — elamuste sarnasus- ja erinevusmustrite, fenomenaalsete​**[lk.108]**​ väljade geomeetrilise struktuuri jne. — mõistmiseks. Nendest ja paljust muust, mida füüsikalised seletused meile elamuse kohta mittereduktiivses kontekstis öelda võivad, tuleb palju põhjalikumalt juttu [6. peatükis](#ORG.6.4). Füüsikalisest seletusest üksi igatahes ei piisa.

Siinkohal kerkib hulganisti loomulikke vastuväited.

#### Esimene vastuväide: võib-olla on meie latt liiga kõrge?

Mõni ütleb võib-olla, et kõrgtaseme fenomenide seletused postuleerivad lisaks madaltaseme kirjeldusele alati ka “sildseadusi” ning et kõrgtaseme fenomenide üksikasju saabki tuletada ainult selliste sildseaduste abiga. Kuid nagu nägime eelmises peatükis (ja nagu väga hoolikalt argumenteerib Horgan (1978)), ei ole need sildseadused sellistel juhtudel lisafaktid maailma kohta, vaid nad on ise madaltaseme faktide suhtes loogiliselt superventsed. Äärmuslikuks näiteks sellisest sildprintsiibist on superventsuskonditsionaal, mis on, nagu nägime, tavaliselt mingi kontseptuaalne tõde. Teisi, “lokaalsemaid” sildprintsiipe, näiteks seost molekulaarse liikumise ja soojuse vahel, on võimalik füüsikalistest faktidest vähemasti tuletada. Seevastu teadvuse puhul tuleb selliseid sildprintsiipe käsitleda kui midagi elementaarset.

Huvitav on jälgida, kuidas mõne tüüpilise kõrgtasemelise omaduse, näiteks elu puhul ei kehti argumendid, mida me esitasime teadvuse puhul. Esiteks on täiesti mõeldamatu, et elusolendi täpne füüsiline koopia ei oleks samuti elus. Võiks arvata, et asi on kontekstist sõltuvates omadustes (kas inimese teisik, mis oleks iseeneslikult tekkinud kusagil soos, oleks elus, või kas ta oleks inimene?), kuid keskkonna faktide fikseerimine elimineerib sellegi võimaluse. Teiseks ei saa siin analoogiliselt invertspektritega kõnelda “invertelust”. Kolmandaks, niipea kui meile on teada kõik füüsikalised faktid mingi organismi (ja võib-olla ka tema keskkonna) kohta, on meil piisavalt materjali, et saada teada ka kõik bioloogilised faktid. Neljandaks, elu küsimustes episteemilist asümmeetriat ei ole: faktid teiste inimeste elususe kohta on meile põhimõtteliselt sama kättesaadavad kui faktid meie endi elususe kohta. Viiendaks on elu mõistet täiesti võimalik analüüsida funktsionaalselt: elusus tähendab laias laastus teatavaid võimeid — kohanemis-, paljunemis- ja metabolismivõimet. Üldiselt võib öelda, et enamikku kõrgtaseme fenomene saab taandada füüsikalisele struktuurile ja füüsilistele funktsioonidele, ning on põhjust uskuda, et struktuurilised ja funktsionaalsed omadused on füüsika suhtes loogiliselt superventsed.

#### Teine vastuväide: kas vitalist ei ütleks sedasama bioloogilise elu kohta?

Kõigest sellest hoolimata on üheks tüüpiliseks vastuseks minu argumentidele olnud, et vitalist ütleks sedasama elu kohta.[[147]](#footnote-148) Näiteks võiks vitalist, kes tahaks tõendada, et elu ei ole reduktiivselt seletatav, ​**[lk.109]**​ väita, et minu füüsiline teisik, kes ei oleks elus, on loogiliselt võimalik. Samuti võiks ta väita, et elusus on lisafakt, mida ei saa seletada füüsikaliste faktide kirjeldusega. Kuid vitalist eksiks. Kas ei eksi analoogiliselt ka teadvuse reduktiivse seletamise vastased?

Ma arvan, et need, kes nii ütlevad, ei ole aru saanud, milles on vitalistlike vastuväidete allikas. Vitalismi ajendiks on peamiselt kahtlus, kas füüsikalistest mehhanismidest piisab kõigiks eluga kaasnevateks keerulisteks funktsioonideks: adaptiivseks käitumiseks, paljunemiseks jms. Tollal olid teadmised biokeemiliste mehhanismide äärmisest keerukusest väga napid, mistõttu seesugused kahtlused olid täiesti loomulikud. Kuid kõigi nende kahtluste taga oli kontseptuaalne eeldus, et elu seletamine tähendab just mitmesuguste funktsioonide seletamist. Ja maksab tähele panna, et sedamööda, kuidas neile funktsioonidele hakkasid ilmuma füüsikalised seletused, kadusid suuremalt jaolt ka vitalistlikud kahtlused. Seevastu teadvuse probleem jääb alles ka siis, kui kõigile funktsioonidele on seletus leitud.

Kui esitada mõistlikule vitalistile täielik füüsikaline seletus, mis näitab ära, kuidas kõik need funktsioonid toimivad füüsikaliste protsesside varal, tunnistaks ta, et bioloogiline elu on leidnud seletuse. Isegi kontseptuaalselt pole mingit võimalust, et need funktsioonid toimiksid, aga süsteem ei oleks elus. Mõni äärmuslik vitalist eitaks võib-olla isegi seda ja väidaks, et elu funktsionaalne seletus jätab midagi ikkagi seletamata, võib-olla näiteks eluvaimu. Kuid sellele on ilmne ja veenev vastus: erinevalt elamusest ei ole eluvaim midagi niisugust, millesse uskumiseks meil oleks sõltumatut alust. Niivõrd kui selleks on üldse alust olnud, on eluvaim olnud seletuslik konstrukt: “me peame ju tunnistama millegi olemasolu, mis nii hämmastavad asjad võimalikuks teeb”. Ent niipea kui meil on elu funktsioonidele parem seletus leitud, võib eluvaimust kui seletuskonstruktist loobuda. Seevastu teadvuselamus sunnib end meile ise peale kui *explanandum*​, millest pole nii lihtne lahti saada.

Üks põhjus, miks vitalist võiks leida, et elu funktsionaalses seletuses on midagi puudu, ongi selles, et füüsikalises kirjelduses pole midagi, mis selgitaks meile, miks on olemas tunne olla elus. Võib-olla tuli mingi osake usust eluvaimusse just meie siseelu fenomenidest. Elususe ja elamuse mõistete vahel on sageli aimatud seost, ja ka tänapäeval võib õigusega öelda, et elu juures nõuab seletust muu hulgas ka asjaolu, et paljudel elusolenditel on teadvus. Kuid seesuguste vitalistlike kahtluste olemasolu ei tule teadvuse reduktiivse seletamise pooldajatele kasuks, sest keegi pole suutnud näidata, et need kahtlused oleksid alusetud.

#### Kolmas vastuväide: kas mõeldavus võib viia võimalikkuseni?

Argumendid, milles pearolli mängib mõeldavus, on filosoofide silmis sageli kahtlased; nende sagedane vastuväide on, et mõeldavusest ei piisa võimalikkuseks. See on peen küsimus, mille üle ma olen arutlenud [varemgi](#PARA.46-9) ja ​**[lk.110]**​ mille juurde ma tulen ka allpool veel tagasi, kuid siinkohal ei ole peensused nii olulised. Kui jutt käib seletamisest, on mõeldavusel selgelt keskne roll. Kui me peame oma mõtiskluste tulemusena võimalikuks, et kõik need füüsikalised protsessid võiksid toimuda ka ilma teadvuseta, siis ei ole teadvusele võimalik leida rahuldavat seletust: alati jääb veel küsimus, miks oleme olemas meie, aga mitte sombid. Kuigi inimlikud võimed seavad mõeldavusele omad piirid, seavad nad sarnased piirid ka seletamisele.

Seda mõtet saab sõnastada ka teisiti, kui panna tähele,et fenomenide reduktiivne seletamine füüsika mõistete kaudu eeldab asjaomaste kõrgtaseme faktide aprioorset tuletatavust füüsikalistest faktidest (või, nagu ma seda ülal nimetasin, loogilist superventsust primaarsete intensioonide alusel). Sellise seose puududes võime alati esitada järgmise küsimuse: miks sünnib füüsikalistest protsessidest teadvus? Me oleme näinud, et peaaegu kõigis valdkondades on niisugune seos olemas ja reduktiivne seletamine seega võimalik, kuid teadvuselamuse puhul ei paista niisugust seost olevat. Võib kahelda, kas ontoloogilised vaated, näiteks materialism, selliseid aprioorseid seoseid eeldavad, — võtan selle küsimuse arutusele järgmises peatükis —, ent kui kõne all on reduktiivne seletamine, omandavad need seosed otsustava tähtsuse.

#### Neljas vastuväide: kas ei ole see kõik vaid rida intuitsioonile toetuvaid tsirkulaarseid väiteid?

Edasi võidakse mulle ette heita, et kõik ülaltoodud argumendid ei ole midagi enamat kui hulk intuitiivseid ettekujutusi. Muidugi toetuvadki kõik need argumendid mingis mõttes intuitsioonile, kuid ma olen püüdnud näidata, kui loomulikud ja selged need intuitiivsed arusaamad on ja kui vägivaldne oleks neid eitada. Kõige tähtsam neist on, et siin on midagi, mis vajab seletust: esimese isiku elamusega seoses tekib mingi probleem, mida ei kerki kognitsiooni vaatlemisel kolmanda isiku positsioonist. Niipea kui me tunnistame, et esimese isiku elamus sunnib meile peale mingi *explanandum*​i, mis kolmanda isiku positsioonilt ei avaldu, ilmub enamik ülaltoodud argumente iseenesest. Näiteks järeldub siit otsekohe, et seda, mis vajab seletust, ei saa analüüsida mingi funktsionaalse rollina, kuna viimane avaneb meile ka kolmanda isiku positsioonilt vaadeldes ning on palju lihtsam fenomen.

Just see intuitiivne arusaam ongi kogu teadvuse probleemi *raison d’être*​. Neist arusaamadest järjekindlaks jäädes mööda vaadata saab ainult kogu probleemi ja fenomeni täielikult eitades. Vähemasti “filosoofilistes” arutlustes on alati võimalik täielikult eitada nii kõnealuseid intuitiivseid ettekujutusi kui seda, et siin on lisaks mitmesugustele funktsioonidele veel midagi, mis nõuab seletust. Aga kui võtta teadvust tõsiselt, siis tulevad järeldused, mida ma siin põhjendada püüan, iseenesest.

#### ​**[lk.111]**​ Viies vastuväide: kas ei pea kõik seletused kusagil lõppema?

Viimane vastuväide on, et ükski seletus ei alga tühjalt kohalt: seletamine peab alati kusagil lõppema. Näiteks planeetide liikumise seletamisel võetakse gravitatsiooniseadused ja massi olemasolu aksioomideks. Võib-olla tuleks ka teadvuse küsimuses midagi lihtsalt aksioomiks võtta? See mõte on mulle sümpaatne; ma arvan tõepoolest, et me peame teadvuse seletamisel midagi aksioomiks võtma. Kuid sellega väljume me paratamatult reduktiivse seletamise piiridest. Tõepoolest, toodud analoogia kinnitab mittereduktiivset positsiooni, mida ma siin kaitsen. Me võtame füüsikaseadused aksioomideks, sest nad on fundamentaalsed. Kui võtta aksioomiks füüsikaliste protsesside ja teadvuselamuse vaheline ühenduslüli, jõuame mõtteni, et seda ühenduslüli tuleks samuti pidada millekski fundamentaalseks. Ma tulen selle küsimuse juurde tagasi [järgmises peatükis](#PARA.82-8).

### 3. Kognitiivne modelleerimine

Siin ja järgnevates jagudes toon mõned näited reduktiivsete seletuste läbikukkumisest, vaadeldes kriitiliselt mõningaid arusaamu teadvusest, mida on väljendanud teadlased mitmelt erialalt. Mitte kõiki neid kontseptsioone ei ole välja pakutud teadvuselamuse reduktiivsete seletustena, kuigi neid on sageli sellistena tõlgendatud; igal juhul on aga õpetlik näha, kuhu nende kontseptsioonidega jõuda võib ja kuhu mitte. Arutluse käigus on huvitav panna tähele erinevusi nende teadlaste hoiakutes rasketele küsimustele teadvuselamuse kohta.

Võtan kõigepealt vaatluse alla kontseptsioonid, mille aluseks on *kognitiivne modelleerimine*​. Kognitiivne modelleerimine on sobiv meetod enamiku kognitiivteaduse probleemide lahendamisel. Kognitiivsete protsesside põhjusliku dünaamika mudeli abil on võimalik seletada kognitiivse agendi käitumise põhjuslikke mehhanisme. Sel viisil võib leida häid seletusi mitmesugustele psühholoogilistele fenomenidele, näiteks õppimisele, mälule, tajumisele, tegevuse juhtimisele, tähelepanule, kategoriseerimisele, keelelisele käitumisele jne. Kui meil on näiteks mudel, mis kajastab põhjuslikku dünaamikat kelleski, kes parajasti õpib, siis võib iga süsteemi, milles see dünaamika sobivas keskkonnas realiseerub, nimetada õppivaks süsteemiks. Mudel näitab meile, kuidas mõningad funktsioonid toimivad, ja nende funktsioonide seletamisest piisab õppimise seletamiseks. Kuid teadvuse seletamiseks sellest ei piisa. Iga mudeli puhul, mille me välja pakume, jääb veel üks küsimus: miks peaks selle mudeli realiseerumisega kaasnema teadvus. Mudeli kirjeldamine ja analüüs üksi sellele küsimusele vastust ei anna.

Mõnikord öeldakse, et pakutud teadvusemudelid on mittekontrollitavad, sest pole võimalik kindlaks teha, kas mudelit realiseerival süsteemil on teadvus või mitte. See on tõsi, kuid on ka üks sügavam raskus. Isegi kui meil ​**[lk.112]**​ oleks (​*per impossibile*​) “elamusmeeter”, mille abil saaks piiluda süsteemi sisse ja teha kindlaks, kas tal on teadvus, saaksime tuvastada vaid korrelatsiooni. Me teaksime, et selle mudeli realiseerumisega kaasneb iga kord teadvus. Kuid see ei annaks seletust teadvusele selles mõttes, nagu taolised mudelid annavad seletusi teistele vaimsetele fenomenidele.

Muidugi on selliste mudelite abil on võimalik seletada “teadvust” psühholoogilises mõttes, teadvust kui kognitiivset või funktsionaalset võimet. Väga heatahtliku suhtumise korral võib niisuguses valguses mõista paljusid olemasolevaid “teadvuse mudeleid”. Me võime oodata neilt seletusi raporteerimisvõime, tähelepanu, introspektsioonivõimete jne. kohta. Kuid ükski neist ei seleta ligilähedaltki, miks peaks nende protsessidega kaasnema teadvuselamus. Toome selgituseks mõned näited.

Vaatleme kõigepealt Bernard Baarsi (1988) kognitiivset mudelit — üht osa tema kognitiivpsühholoogilisest teadvusekäsitlusest, millele ta on pühendanud terve raamatu. Baars püüab kõikvõimalikele katseandmetele toetudes põhjendada oma põhiteesi, mille järgi teadvus on omamoodi *globaalne töölaud* mõistuslike infoprotsessorite hajussüsteemis. Globaalne töölaud on otsekui klassitahvel, millel protsessorid edastavad teateid kogu süsteemile. Selle globaalse töölaua sisu ongi teadvuse sisu.

Oma mudeli abil seletab Baars märkimisväärselt paljusid inimpsüühikas toimuvate protsesside omadusi. Mudel annab väga hea lähtekoha selgitamaks subjekti ligipääsu informatsioonile ning informatsiooni rolli tähelepanus, raporteerimisvõimes, käitumise tahtelises juhtimises ja isegi minamõiste kujunemises. Seega sobib globaalse töölaua mudel hästi teadvuse seletamiseks igas selle sõna psühholoogilises mõttes. Vähemasti teadlikkusele annab see mudel üldise teoreetilise seletuse.

Kuid elamuse reduktiivset seletust me siit ei leia. Küsimust, miks peaks neist protsessidest sündima elamus, lihtsalt ei esitata. Võib oletada, et elamuse sisu ühtib selle teooria järgi täpselt töölaua sisuga. Ent isegi kui see on nii, ei ole teoorias endas midagi, mis seletaks, miks kaasneb globaalsel töölaual sisalduva informatsiooniga elamus. Parimal juhul võib see teooria öelda, et informatsiooniga kaasneb elamus sellepärast, et ta on globaalselt kättesaadav. Kuid nüüd kerkib sama küsimus teisel kujul: miks peaks globaalse kättesaadavuse pinnalt sündima teadvuselamus? Seda sildküsimust Baarsi töö ei käsitle.

Seda puudust mainib möödaminnes ka Baars ise: “Skeptiline lugeja võib … küsida, kas me kirjeldame tõepoolest teadvuselamust, või saame me tegelda vaid teadvusega kaasnevate kõrvaliste fenomenidega” (lk. 27). Vastuseks ütleb ta, et teaduslikud teooriad püüavad “asjale endale” ​**[lk.113]**​ vähemalt läheneda; näiteks bioloogia seletab pärilikkust ennast, mitte ainult sellega kaasnevaid fenomene. Kuid see tähendab, nagu nägime, et kõnealune teooria lihtsalt ignoreerib teadvuse põhimõttelist erinevust nendest fenomenidest. Pärilikkuse puhul nõuavadki seletust ainult funktsioonid. Teadvuse puhul aga on üks *explanandum* veel — elamus ise. Seega võib Baarsi teooriat pidada huvitavaks vaateviisiks teadvuse kognitiivsetele alusprotsessidele, mis edendab kaudselt tublisti meie arusaamist teadvusest, kuid jätab puutumata põhiküsimused: miks teadvus eksisteerib ja kuidas ta kognitiivsetes protsessides tekib?

Teadvuse kognitiivse mudeli, tegelikult isegi kaks mudelit, on esitanud ka Daniel Dennett. Esimene neist (Dennett 1978c) on “kastmudel”, mis kirjeldab informatsiooni liikumist moodulite vahel (joonis 3.2). Selle mudeli komponentideks on 1) tajumoodul, 2) lühimälusalv M, mis saab ​**[lk.114]**​ informatsiooni tajumoodulilt, 3) juhtsüsteem, mis küsitleb mälusalve ja võib suunata tähelepanu tajumooduli sisule, ja 4) “avalike suhete” keskus, mis teisendab juhtimissüsteemilt saadud kõneaktikäsud avaliku keele lausungiteks.

​**[lk.113]**​ **Joonis 3.2.** Dennetti kognitiivne teadvusmudel. (Ümber joonistatud raamatust Daniel C. Dennett, “Brainstorms: Philosophical Essays on Mind and Psychology” [“Ajurünnakud: filosoofilisi esseid vaimust ja psühholoogiast”], joonis 9.1, lk. 155). (​The MIT Pressi loal.​)

Unenäod

Tajude analüüs; Hüpotees; Testimine; Paralleeltöötlus; Tunnuste detektorid; Ikooniline mälu

Meeleelundid

Tähelepanu

Juhtsüsteem: 1) Introspektsioon; 2) Küsimus mälule; 3) Mälult saadud vastuse hindamine (tsenseerimine, tõlgendamine, järelduste tegemine, edastamine); 4) Kõneaktikäsu saatmine avalike suhete keskusele

K

V

M

Probleemide lahendamine

Avalike suhete kes­kus: Kõne­akt; Süva­struktuur; Pind­struktuur; Fono­loogia; Grafo­loogia; Mo­toorsed protse­duurid; Mo­toorsed protse­duurid

​**[lk.114]**​ Mida see mudel seletada võiks? Olgugi ta vormilt väga jämedakoeline (küllap tunnistaks seda ka Dennett), saaks teda täpsustada nii, et ta annaks seletuse *raporteerimisvõimele*​, s.t. meie võimele anda teistele teateid oma seesmiste seisundite sisust. Mudel annab ka skemaatilise seletuse meie võimele kasutada tajudest saadavat infot käitumise juhtimiseks, võimele introspektiivselt mõista oma seesmisi seisundeid jne. Kuid ta ei ütle meile midagi selle kohta, miks peaks olema olemas tunne olla süsteem, milles need protsessid kulgevad.

Raamatus “Teadvuse seletus” [​*Consciousness Explained*​] (1991) esitab Dennett detailsema kontseptsiooni, mis toetub paljudele uuematele kognitiivteaduste saavutustele. Mudel, mille ta siin välja pakub, kujutab põhimõtteliselt “pandeemoniumi”, mis koosneb paljudest väikestest agentidest, kes kõik konkureerivad tähelepanu pärast; edasiste protsesside dirigeerimisel mängib põhirolli agent, kes karjub kõige valjemini. Selles mudelis ei ole keskset “staapi”, vaid palju üheaegselt tegutsevaid juhtkanaleid. Dennett viitab oma kontseptsioonis neuroteadusele ja evolutsioonibioloogiale, tehisintellekti alalt konnektsionistlikele mudelitele ja reegelsüsteemidele​[[148]](#footnote-149)

Kontseptsiooni keerukusest hoolimata vaatleb see teine mudel peamiselt neidsamu fenomene nagu esimene. Hea õnne korral suudaks ta seletada raporteerimisvõimet ning mitmesugust liiki informatsiooni mõju käitumise juhtimisele. Samuti saaks temaga võib-olla seletada tähelepanu fookustamist​. Kontseptsioon annab provokatiivse seletuse mõnedele meie kognitiivsetele võimetele, ent küsimusele, miks peaks nende võimetega kaasnema teadvuselamus, ei vasta ta midagi enamat kui eelmine mudel.

Erinevalt enamikust teistest autoritest, kes on esitanud kognitiivseid mudeleid, kinnitab Dennett otsesõnu, et tema mudelite abil on võimalik seletada elamuse kohta kõike, mis seletamist vajab. Muu hulgas leiab ta, et teadvuse seletamiseks piisab, kui seletada ära raporteerimisvõime, juhtimine ja teised funktsionaalsed fenomenid; iga fenomen, mis näib olevat käsitlemata jäänud, on kimäär. Kohati tundub, et ta lähtub põhimõttelisest eeldusest, et kui kõik need mitmesugused funktsioonid on ära seletatud, siis ongi kõik seletatud (vt. näit. Dennett 1993a, lk. 210), kohati aga esitab ta selle kasuks ka argumente, millest mõned tulevad vaatluse alla allpool.[[149]](#footnote-150)

Analoogiliselt saab kritiseerida ka Churchlandi (1995), Johnson-Lairdi (1988), Shallice’i (1972, 1988a, 1988b) ja paljude teiste autorite kognitiivseid teadvusmudeleid. Kõik nad pakuvad põnevaid seletusi kognitiivsetele funktsioonidele, kuid vaatavad mööda tõeliselt rasketest küsimustest.

### ​**[lk.115]**​ 4. Neurobioloogiline seletus

Uuemal ajal on hakanud populaarsust võitma teadvuse neurobioloogilised käsitlused. Sarnaselt kognitiivsetele mudelitele võivad ka need psühholoogiliste fenomenide, sealhulgas mitmesuguste teadlikkuse vormide seletamisel paljuski abiks olla. Samuti võivad nad öelda meile üht-teist teadvusega korreleeruvate ajuprotsesside kohta. Kuid ükski neist käsitlustest ei anna sellele korrelatsioonile seletust: me ei saa teada, miks peaks ajuprotsesside pinnalt üldse sündima elamus. Neuroteaduse seisukohalt on sellised korrelatsioonid lihtsalt toores fakt.

Metodoloogilises mõttes ei ole selge, kuidas neuroteadusliku teooria väljatöötamise kallale asudagi tuleks. Missugused katsed tuvastaksid korrelatsiooni mingite neuronprotsesside ja teadvuse vahel? Tavaliselt toetuvad teoreetikud vaikivalt mõnele teadvuse psühholoogilisele kriteeriumile, näiteks tähelepanu fookusele, käitumise juhtimisele, kõige sagedamini aga võimele oma seesmise seisundi kohta sõnalisi ütlusi anda. Seejärel pannakse tähele, et nende kriitiliste tunnuste ilmumisel realiseerub mingi neurofüsioloogiline omadus. Ja teadvuseteooria ongi käes.

Kuid juba sellistele kaudsetele kriteeriumidele toetumise fakt üksi näitab selgesti, et teadvuse reduktiivset seletust siin ei anta. Parimal juhul võib neurofüsioloogiline käsitlus seletada, miks asjaomane psühholoogiline omadus realiseerub. Küsimus, miks peaks ühe või teise psühholoogilise omadusega kaasnema teadvuselamus, jääb vastuseta. Kuna nende teooriate edu sõltub juba ette eeldusest, et psühholoogiliste omaduste ja teadvuselamuse vahel on mingi seos, siis on selge, et nad ei püüa seda seost seletada. See saab ilmseks, kui vaadelda mõningaid uuemaid neuroteaduslikke teadvusekäsitlusi.

Viimasel ajal on neuroteaduses palju tähelepanu pööratud teatavatele 40-hertsistele võnkumistele nägemiskorteksis ja teistes ajupiirkondades. Francis Crick ja Christof Koch (1990) on püstitanud hüpoteesi, et need võnkumised ongi see fundamentaalne neuronprotsess, millest sünnib teadvuselamus, ning on soovitanud lähtuda neurobioloogilise teooria väljatöötamisel sellest ideest[[150]](#footnote-151).

Miks just 40-hertsised võnkumised? Eelkõige sellepärast, et olemasolevatel andmetel mängivad just need võnkumised olulist rolli mitut eri liiki informatsiooni sidumisel ühtseks tervikuks. Kui meil on mingi olukorra kohta kaht eri tüüpi informatsiooni — näiteks objekti kuju ja tema asukoha kohta —, siis võib kummalgi olla omaette representatsioon, kuid teooria ütleb, et süsteem suudab erinevaid representatsioone, millel on ühine võnkesagedus ja -faas, edasiste protsesside käigus ühendada ja üheskoos töömällu salvestada. Sel viisil võivad kõige erinevamat liiki infokillud moodustada ühtse “teadvuse sisu”.

Selline teooria aitaks võib-olla tõepoolest paremini neurobioloogiliselt mõista informatsiooni ühtsustamist ja töömälu ning pikapeale ​**[lk.116]**​ ehk ka seda, kuidas süsteem informatsiooni integraalselt käitumise juhtimiseks kasutab. Kuid endiselt jääb vastamata põhiküsimus: miks peaks nende võnkumistega kaasnema teadvuselamus? Kõnealune teooria annab osalise vastuse: sellepärast, et nende võnkumiste varal toimub informatsiooni ühtsustamine. Kuid küsimust, miks peaks ühtsustamisprotsessiga kaasnema elamus, ta ei püstita. Teooria edu sõltub eeldusest, et informatsiooni ühtsustamisel on mingi seos teadvusega, ja sellepärast ei annagi ta midagi, et seda seost seletada.

Teadvuse “suur probleem” on Crickile ja Kochile nähtavasti südamelähedane: nad nimetavad seda “peamiseks mõistatuseks vaimu neurontasemelise mõistmise teel”. Nad väidavad, et puhtalt kognitiivseid käsitlusviise ootab kindel läbikukkumine ning et tarvis on neurontasemel teooriaid. Kuid nad ei anna meile põhjust uskuda, et nende teooria suudab tõeliselt rasketele küsimustele vastata paremini kui kognitiivsed teooriad. Ja tõepoolest ei ütlegi nad, nagu püüaks nende projekt lahendada elamuse probleemi. Ühes intervjuus on Koch väljendanud teooria kitsendusi üsna selgelt:

Alustuseks unustagem tõeliselt rasked aspektid, näiteks subjektiivsed tunded, sest neile ei pruugi teaduslikku lahendust leidudagi. Subjektiivsed seisundid — mäng, valu, rõõm, sinise värvuse nägemine, roosilõhna haistmine — tundub, et seda subjektiivset taset lahutab materialistlikust tasemest, kus nähtusi seletatakse molekulide ja neuronite kaudu, tohutu hüpe. Koondagem tähelepanu sellele, mida on kergem uurida, näiteks visuaalsele teadlikkusele. Te ajate praegu minuga juttu, kuid ei vaata mind, vaid oma kaputšiinotassi ja olete tassist seega teadlik. Te võite öelda: “Siin on tass ja tassis mingi vedelik.” Kui ma tassi teile ulatan, sirutate te käe välja ja võtate selle, reageerite arukalt. Seda ma nimetangi teadlikkuseks.[[151]](#footnote-152)

Teise neurofüsioloogilise teadvuseteooria on raamatus “Olevik kui mälestus” [​*The Remembered Present*​] (1989) ning oma teistes raamatutes ja artiklites, visandanud Gerald Edelman. Tema teooria keskseks elemendiks on korduvläbitavad neuronahelad, mis liigitavad tajusignaalid enne nende mällu salvestamist kontseptuaalsete kategooriate alusel. Tajuinformatsiooni ja seesmise seisundi keerukates vastastoimetes (mida kujutab joonis 3.3) tekib “primaarne teadvus”. Tema “kõrgemat järku teadvuse” mudeli uudseks mäluelemendiks on “semantiline alglaadimine”, mille tagajärjel ilmuvad “mina”, mineviku ja tuleviku mõisted. Kõik see toimub koostöös keeleproduktsiooniga Broca ja Wernicke piirkondades.

​**[lk.117]**​ **Joonis 3.3.** Kõrgemat järku teadvuse skeem Edelmani järgi. (Joonise 12-4 järgi raamatus Edelman 1992 (lk. 132)). By permission of BasicBooks, a division of HarperCollins, Inc.​)

MINA

Seesmised homöostaatilised süsteemid

Ajutüvi, hüpotaalamus, autonoomsed keskused

Uute seesmiste seisundite ja väärtuse registreerimine

Korrelatsioon hüpokampuses; mandeltuum, septum​

Kontseptuaalne kategoriseerimine

Eriline väärtus- ja kategooriamälu otsmiku-, oimu- ja kiirukorteksis

(eeldab varasemaid kogemusi ja neurongrupivalikut)

MITTEMINA

Signaalid maailmast, sealhulgas lihastaju

Primaarne ja sekundaarne nägemis-, kuulmis-, kompimis- jne. korteks

Tajude jooksev kategoriseerimine

Semantiline alglaadimine

Broca keskus ja Wernicke keskus

KÕRGEMAT JÄRKU TEADVUS

PRIMAARNE TEADVUS

Korduvläbitav silmus, mis ühendab väärtuste ja kategooriate mälu uute tajukategoriseeringutega

​**[lk.116]**​ Üldiselt püüab Edelman oma raamatus seletada rohkem tajumist, mälu ja keelt kui teadvust. Niipalju kui ta teadvusest kõnelebki, on tema jutt sageli ebamäärane, kuid tundub, et mingi sarnane mudel võib ehk pikapeale aidata mõista *pertseptiivset teadlikkust* — see tähendab, tajuprotsesside mõju edasistele protsessidele ja käitumise juhtimisele — ning eneseteadvuse aspekte, eriti minamõiste päritolu.

​**[lk.117]**​ Edelman ei selgita, kuidas peaks kõigi nende protsesside tagajärjel sündima teadvuselamus, vaid lihtsalt nendib korrelatsiooni olemasolu. Ta ei varjagi seda, vaid möönab, et fenomenaalne elamus ongi teadvuseteooria jaoks kõige raskem probleem ning et ükski füüsikaline teooria ei suuda kvaale lõpuni seletada:

Siit selgub üks võimalus kvaalide probleemi mõistmiseks. Mõistlik on lähtuda teadvuseteooria rajamisel eeldusest, et ka teistel teadvusega inimestel on kvaalid nagu meilgi, olenemata sellest, kas me võtame neid kui teadlastest vaatlejaid või kui katseisikuid​[[152]](#footnote-153). … Seejärel võime tunnistada inimesed parimaks kanooniliseks objektiks teadvuse uurimisel. Seda õigustab asjaolu, et kõigile inimeste subjektiivsetele ütlustele (sealhulgas kvaale puudutavatele) on võimalik leida korrelaate aju struktuurides ja funktsioonides. Pärast teooria loomist, mis lähtub eeldusest, et inimestel on kvaalid, võime nende korrelatsioonide põhjal uuesti uurida mõningaid kvaalide omadusi. Meie suutlikkus kvaale individuaalselt läbi elades nende kohta ütlusi anda ja korrelatsioone leida avabki võimaluse teadvuse teaduslikuks uurimiseks. (Edelman 1992, lk. 115.)

​**[lk.118]**​ Kuna see teooria eeldab korrelatsioone juba ette, siis on selge, et ta ei saa anda elamuse reduktiivset seletust. Enamasti ütleb Edelman, et ta seletab vaid protsesse, mis on teadvuselamuse aluseks, ega väidagi, nagu seletaks ta elamust ennast.[[153]](#footnote-154)

### 5. Apelleerimine uuele füüsikale

Mõnikord oletatakse, et teadvuse seletuse võti peitub mingis uutmoodi füüsikalises teoorias. Meie väite taga, et teadvus ei tulene meie maailma füüsikast, võis olla vaikiv eeldus, et see füüsika sarnaneb füüsikaga, nagu me seda praegu mõistame — füüsikaga, mis kõneleb osakeste ja väljade jaotusest aegruumilises muutkonnas ning selles toimuvatest keerukatest põhjuslikest ja evolutsioonilistest protsessidest. Mõni oponent võib nõustuda, et niisuguses füüsikas ei ole tõepoolest midagi, millest saaks tuletada teadvuse olemasolu, kuid ühtlasi kinnitada, et teadvus võib ilmuda järeldusena mingit uut liiki füüsikateooriast.

Senikaua kui keegi pole midagi konkreetsemat välja pakkunud, on seda väidet raske hinnata. Enne tahaks siiski näha mõnd näidet, mis annaks meile ligikaudseltki aimu, kuidas see uus füüsika võiks välja näha. See näide ei pea tänapäeva teooriate valguses isegi usutav olema, kuid ta peab olema mingis mõttes äratuntav füüsikana. Otsustav küsimus on: kuidas oleks mingist äratuntavalt füüsikalisest teooriast võimalik tuletada teadvuse olemasolu? Teooria puhul, mis seisneks väljade, lainete, osakeste jms. struktuuri ja dünaamika kirjeldustes, kerkiksid endiselt kõik probleemid, mida me juba tunneme. Ja pole selge, mismoodi peaks välja nägema füüsikaline teooria, mis oleks piisavalt teistsugune, et neid probleeme ei tekiks.

Häda on selles, et füüsikaliste teooriate põhielemendid taanduvad nähtavasti alati kahele asjale: füüsikaliste protsesside struktuurile ja nende dünaamikale. Eri teooriad kõnelevad erinevatest struktuuridest. Newtoni füüsikas on struktuuriks eukleidiline aegruum, relatiivsusteoorias diferentseeruv mitteeukleidiline muutkond​, kvantteoorias lainefunktsioonide Hilberti ruum. Samuti on dünaamika neis struktuurides eri teooriate järgi erisugune: Newtoni seadused, relatiivsusprintsiibid, lainevõrrandid. Kuid struktuur ja dünaamika ei anna meile midagi enamat kui veelgi rohkem struktuuri ja dünaamikat. Nii võib rahuldavalt seletada kõikvõimalikke struktuurilisi ja funktsionaalseid kõrgtaseme omadusi, kuid see kõik ei puuduta teadvuselamust. Ükski faktide kogum füüsikalise struktuuri ja dünaamika kohta ei saa anda kokku fenomenoloogilist fakti.

Muidugi, kui defineerida füüsikat kui fundamentaalteadust, mille faktidest ja seadustest tuleneb kõik muu, siis selles mõttes peabki füüsikast tulenema ka teadvuse olemasolu. Ent füüsika niisuguse tõlgenduse korral minetab meie küsimus kogu huvitavuse. Kui võtta füüsikasse ​**[lk.119]**​ juurde teooriaid, mis on välja töötatud spetsiaalselt teadvuse fenomeni uurimiseks, ilma et selleks oleks kaalukamat alust, võib füüsika anda teadvusele küll mingi seletuse, kuid see ei ole siis kindlasti mitte reduktiivne seletus. Meie eesmärkideks kõige sobivam on mõista füüsikat kui fundamentaalteadust, mis peab seletama välise maailma vaatluste tulemusi. Kui niisugusest füüsikast oleks võimalik tuletada teadvuse faktid, ilma et teadvusel endal temas keskset rolli oleks, siis leiduks teadvusele tõepoolest reduktiivne seletus. Kuid kaalutlused, millest ma juba kõnelesin, annnavad tõsist alust selliste reduktiivsete seletuste võimalikkust eitada.

Peaaegu kõik senised ideed teadvuse seletamiseks füüsika abil koondavad tähelepanu füüsika kõige mõistatuslikumale harule, nimelt kvantmehhaanikale. See on ka arusaadav: teadvuse seletuses füüsika abil peaks olema midagi erakordset, ja kvantmehhaanika ongi tänapäeva füüsika kõige erakordsem haru. Lähemalt uurides tundub aga, et ta ei ole siiski piisavalt erakordne.

Näiteks Penrose arvab (1994), et teadvuse mõistmise võti võib peituda teoorias, mis ühendaks kvantteooria üldrelatiivsusteooriaga. Ta oletab, et gravitatsioonilised efektid, mida praegu veel ei mõisteta, võivad kutsuda esile lainefunktsiooni kollapsi, mis toob loodusseadustesse mittealgoritmilise elemendi. Hameroffi (1994) ideedele toetudes oletab ta, et inimkognitsioonis mängivad olulist rolli kvantkollapsid *mikrotuubulites* — valkstruktuurides, mida leidub neuroni tsütoskeletis. Penrose’i ja Hameroffi arvates võibki kvantkollaps mikrotuubulites olla teadvuselamuse füüsikaliseks aluseks.

Need on äärmiselt spekulatiivsed ideed, kuid on vähemasti mõeldav, et nad võiksid kaasa aidata inimese kognitiivsete funktsioonide mõningate elementide seletamisele. Penrose arvab, et kollapsi mittealgoritmiline element võib seletada meie matemaatilise intuitsiooni mõningaid aspekte, mis tema arvates ületavad iga algoritmilise süsteemi võimalused. Hameroff oletab, et superpositsioonilise lainefunktsiooni kollaps võib aidata seletada mõningaid inimlike otsustusprotsesside aspekte. Kuid selles kõiges ei paista olevat midagi, millest võiks olla abi teadvuselamuse seletamisel. Miks peaks kvantprotsessidest mikrotuubulites sündima teadvus? See küsimus pole sugugi kergem kui vastav küsimus klassikaliste protsesside kohta klassikalises ajus. Kui jutt läheb elamuse probleemile, on mittealgoritmilised ja algoritmilised protsessid ühes paadis.

On oletatud ka, et teadvuseteooria võti peitub kvantmehhaanika mittelokaalsuses, mille võimalikkusele viitavad hiljutised eksperimendid seoses Einstein–Podolsky–Roseni paradoksi ja Belli teoreemiga (sellesisulisi ideid on esitanud Lahav ja Shanks 1992). Ent olgu füüsika pealegi mittelokaalne, ikkagi on raske näha, mis abi oleks sellest teadvuse seletamisel. Ka mittelokaalsete füüsikaliste protsesside puhul jääb endiselt kehtima loogiline võimalus, et need protsessid toimuksid ilma teadvuseta. Seletuslünk on sama suur nagu ennegi.

​**[lk.120]**​ Kõige sagedamini mainitakse teadvuse seost kvantmehhaanikaga seoses asjaoluga, et on olemas kvantmehhaanika interpretatsioonid, mille järgi lainefunktsiooni kollapsi kutsub esile mõõtmine, mille sooritab teadvusega vaatleja. See tähendaks, et teadvusel on füüsikalise teooria dünaamikas keskne roll. Need tõlgendused on väga vaieldavad, kuid igal juhul tuleb tähendada, et nad ei anna teadvusele seletust, vaid eeldavad teadvuse olemasolu lihtviisil juba ette ning kasutavad seda eeldust mõnede füüsikaliste fenomenide seletamiseks. Mõned teadvuseteooriad püüavad seda seost ära kasutada (näiteks Hodgson 1988; Stapp 1993), kuid need ei ole kindlasti mitte reduktiivsed teooriad.[[154]](#footnote-155)

Võimalust, et fundamentaalsed füüsikateooriad, näiteks kvantmehhaanika, hakkavad teadvuseteoorias kunagi otsustavat rolli mängima, ei saa välistada. Näiteks võib selguda, et teadvus on kuidagi ühenduses mingite fundamentaalsete füüsikaliste omadustega või nende omaduste mingite konfiguratsioonidega. Seos võib olla ka peenem. Kuid sellegipoolest pole palju lootust, et selline teooria annaks teadvusele puhtalt füüsikalise seletuse. Füüsikapõhistel teooriatel ei ole teadvuse reduktiivsel seletamisel suuremat edulootust kui neurobioloogilistel ja kognitiivsetel teooriatel.

### 6. Evolutsiooniline seletus

Isegi neid, kes võtavad teadvust tõsiselt, ahvatleb sageli idee seletada teadvust evolutsiooniliselt. On ju teadvus niivõrd levinud ja tsentraalne nähe, et tundub, et tema tekkimisel evolutsiooni käigus pidi olema mingi põhjus. Täpsemalt öeldes on loomulik oletada, et teadvus tekkis selleks, et täita mingit funktsiooni, mis oleks ilma teadvuseta võimatu. Kui selles funktsioonis küllaldaselt selgusele jõuda, jõuaksime ka mingi ettekujutuseni selle kohta, miks teadvus eksisteerib.

Paraku hindab see idee evolutsiooniliste seletuste võimalusi üle. Loodusliku valiku protsess ei saa teha vahet minu ja minu sombist teisiku vahel. Omaduste evolutsiooniline valik toimub vastavalt nende funktsionaalsele rollile, ent minu sombist teisikul talitlevad kõik samad funktsioonid ja sama hästi nagu minul, ja see tähendab muu hulgas, et ta jätab ka oma geenidest sama palju koopiaid. Seepärast ei suuda evolutsioon üksi seletada, miks tekkisid teadvusega olendid, aga mitte sombid.

Lugejal võib tulla kiusatus vastata: “Aga sombi ei saa ju teha kõike, mida võin teha mina.” Ent kuna minu sombist teisik on kogu oma olemasolu vältel minuga füüsiliselt identne, siis ei ole ka tema käitumine minu omast eristatav. Seega peab igaüks, kes tahab sombide võimeid kahtluse alla seada, leidma vea [käesoleva peatüki algul](#ORG.3.1) toodud argumentides, selle asemel et hakata esitama vastuväiteid alles nüüd.

​**[lk.121]**​ Et vaadata sama asja teise nurga alt, pangem tähele, et teadvuse reaalseks probleemiks on seletada ära printsiibid, mille alusel füüsikalistes süsteemides sünnib teadvus. Eeldatavasti on need printsiibid — olgu nad siis kontseptuaalsed tõed, metafüüsilised paratamatused või loodusseadused — ühesugused kogu aegruumis: kui minu füüsiline koopia oleks tekkinud ka miljon aastat tagasi, ei oleks ta ikkagi vähemal määral teadvusega olend kui mina. Seega on printsiibid, mis seovad teadvust füüsikaga, ise evolutsiooniprotsessist sõltumatud. Kuigi evolutsioonist võib olla palju kasu selgitamisel, miks on välja arenenud just niisugused füüsikalised süsteemid, ei ole temaga midagi teha sildprintsiipide seletamisel, tänu milledele mõnel neist süsteemidest on teadvus.

### 7. Kui kaugele võib reduktiivse seletamisega jõuda?​

Üsna sageli nõustutakse konkreetsete reduktiivsete kontseptsioonide kriitikaga, kuid tehakse siis mööndus: “Jah, see muidugi veel teadvust ei seleta, aga ootame natuke, küll tuleb ka seletus.” Ma loodan, et siinsetest arutlustest on saanud selgeks, et raskused teadvuse seesuguse seletamise teel on põhimõttelisemad. Ülalkirjeldatud mudelite ja teooriate puudused ei ole üldsegi detailides; igatahes ei ole meil kordagi olnud vaja uurida detaile, et näha, mis neis teooriates ei klapi. Probleem on seletamise üldises strateegias. Need mudelid ja teooriad lihtsalt ei ole niisugust liikigi, et suuta seletada teadvust.

Teadvuse üha keerukamate reduktiivsete “seletuste” ilmumine on vältimatu, kuid kõik need on vaid üha keerukamad seletused kognitiivsetele funktsioonidele. Isegi sellised “revolutsioonilised” saavutused nagu konnektsionistlikud võrgud, mittelineaarne dünaamika, tehiselu ja kvantmehhaanika annavad vaid aina võimsamaid funktsionaalseid seletusi. See võib tuua paljugi huvitavat kognitiivteadusse, kuid teadvuse müsteerium sellest ei kao.

Iga puhtalt füüsikalistes mõistetes antud seletus kannatab sama puuduse all. Viimselt formuleeritakse niisugused seletused füüsikaliste protsesside struktuursete ja dünaamiliste omaduste kaudu, ja ükskõik kui peeneks ja keerukaks selline kirjeldus ka ajada, on tulemuseks vaid veelgi rohkem struktuuri ja dünaamikat. Kuigi sellest piisab enamiku looduslike fenomenide käsitlemiseks, jääb teadvuse probleem kõigi struktuuri ja funktsiooni seletamise teel lahenduvate probleemide piiridest väljapoole, mistõttu otsitav seletus peab olema mingit uut liiki.

Võiks ju oletada, et kunagi õnnestub leida reduktiivne seletamismeetod, mis seletaks veel midagi muud peale struktuuri ja funktsiooni, ent kuna viimselt formuleeritakse füüsikaseadused ikkagi struktuuri ja dünaamika kaudu, siis on väga raske näha, kuidas see võiks võimalik olla. ​**[lk.122]**​ Teadvuse olemasolu on struktuuriliste ja dünaamiliste faktide suhtes alati lisafakt, ja seega ei hakka kunagi olema võimalik seletada teda füüsikaliste kontseptsioonidega.

Järelikult tuleb teadvuse seletust otsida mujalt. Muidugi ei pea me loobuma katsetest teadvust seletada; me peame vaid loobuma katsetest seletada teadvust reduktiivselt. Teadvuse mittereduktiivse seletamise võimalikkus jääb lahtiseks. See oleks hoopis teistsugune seletus, mis nõuaks radikaalseid muudatusi meie ettekujutuses maailma struktuurist. Ent kui me oleme need muudatused teinud, võib silmapiiril hakata terendama ka teadvuseteooria.

## ​**[lk.123]**​ 4. peatükk. Naturalistlik dualism

### 1. Materialismivastane argument

Eelmises peatükis huvitas mind rohkem seletuslik küsimus: “Kas teadvust saab seletada füüsikaliste teooriatega?” ja vähem ontoloogilise küsimus: “Kas teadvus on ise füüsiline?”. Kuid need kaks küsimust puutuvad tihedalt kokku, ja käesolevas peatükis hakkan ma kirjeldama ontoloogilisi järeldusi eelmises peatükis esitatud argumentidest. Eelkõige tuleneb loogilise superventsuse puudumisest otseselt materialismi väärus: maailmas on nähtusi, mis ei mahu füüsikaliste nähtuste piiridesse. Väite tõestuskäik on põhijoontes järgmine:

1. Meie maailmas eksisteerivad teadvuselamused.
2. Maailm, mis oleks meie maailmaga füüsiliselt identne, ent milles puuduksid meie maailma positiivsed faktid teadvuse kohta, on loogiliselt võimalik.
3. Järelikult on teadvuse faktid meie maailmas lisafaktid, mis on midagi muud ja enamat kui füüsikalised faktid.
4. Järelikult on materialism väär.

Kui füüsiliselt identne sombimaailm on loogiliselt võimalik, siis järelikult on teadvuse olemasolu meie maailma lisafakt, mida füüsikalised faktid üksi tagada ei suuda. Meie maailma olemus ei ammendu sellega, mis tuleneb ainuüksi füüsikalistest faktidest; tema olemuses on veel midagi, mis tuleneb teadvuse olemasolust. Või, nagu ütleb Lewis (1990): teadvus kannab fenomenaalset informatsiooni. Füüsikalised faktid ei määratle maailma olemisviisi täielikult; teadvuse faktid piiritlevad maailma veelgi.

​**[lk.124]**​ Sarnase järelduse võib tuletada ka invertsete teadvuselamustega maailma loogilisest võimalikkusest. Niisugune maailm oleks füüsiliselt identne meie maailmaga, kuid mõned meie maailma teadvuselamuse faktid selles maailmas ei kehtiks. Sellest järeldubki, et meie maailma teadvuselamuse faktid on lisafaktid, midagi muud ja enamat kui füüsikalised faktid, ning et materialism on väär.

Nii või teisiti tuleb välja, et kui teadvus ei ole füüsika suhtes loogiliselt superventne, siis on materialism väär. Loogilise superventsuse puudumisest järeldub, et meie maailmas leidub positiivseid fakte, mis temaga füüsikaliselt identses maailmas ei kehti ja mis on seega uued faktid, midagi enamat ja suuremat kui füüsikalised faktid. Nagu [2. peatükis](#PARA.36-3), mõtlen ma materialismi all õpetust, mille järgi kõik faktid ammenduvad maailma füüsikaliste faktidega, s.t. iga positiivne fakt on tuletatav füüsikalistest faktidest. Kui sombidega maailmad ja invertmaailmad on võimalikud, siis ei tulene kõik positiivsed faktid füüsikalisest faktidest, ja materialism on väär.

Siinkohal võime kasutada Kripke kujundit. Maailma luues pidi Jumal pärast füüsikaliste faktide kehtestamist tegema veel midagi: ta pidi tagama ka teadvuse faktid. Sombimaailmade ja invertmaailmade võimalikkusest järeldub, et tal oli võimalus valida. Elamused võinuksid maailmas puududa või olla teistsugused, isegi kõigi füüsikaliste faktide samaks jäädes. Tagamaks, et teadvuse faktid oleksid sellised, nagu nad on, tuli maailmale lisada veel midagi.

#### Missugune dualism?

Materialismi läbikukkumine toob meid mingitliiki *dualismi* juurde: maailmal on nii füüsikalisi kui ka mittefüüsikalisi omadusi. Loogilise superventsuse väärusest tuleneb, et elamus on midagi põhimõtteliselt erinevat ükskõik millisest füüsikalisest omadusest. Kuid dualismil on palju erikujusid ja on oluline täpselt mõista, kuhu äsjatoodud argument meid viib.

[Eelmises peatükis](#ORG.3.1) toodud argumendid näitavad, et teadvus ei ole füüsika suhtes loogiliselt superventne, kuid see ei tähenda, et ta ei ole üldse superventne. Tundub, et meile tuttavatel juhtudel on teadvuselamus süstemaatilises sõltuvuses füüsikalisest struktuurist, ja eelmise peatüki argumentides pole midagi, mis ütleks vastupidist. Näiteks jääb endiselt usutavaks, et kui aktuaalses maailmas leiduks olend, kelle füüsikaline struktuur oleks minu struktuuri koopia, siis oleks tal ka teadvuselamus samasugune nagu minul. Seega võib ka nüüd pidada tõepäraseks, et teadvus on füüsika suhtes looduslikult superventne. Seda pilti — looduslik superventsus ilma loogilise superventsuseta — hakkangi ma nüüd edasi arendama.

Need argumendid ei vii meid kartesiaanliku dualismi juurde, mille järgi füüsikalisi protsesse ​**[lk.125]**​ mõjutab mingi omaette vaimne substants. Tänapäeva teaduse kõige usaldusväärsemad andmed ütlevad meile, et füüsiline maailm on põhjuslikult enam-vähem kinnine: igal füüsikalisel sündmusel on piisav füüsikaline põhjus. Kui see on tõsi, siis ei ole kohta “vaimule masinas” [​*ghost in the machine*​], mis teeks mingit põhjuslikku lisatööd. Väikese pilu avab võib-olla kvantmääramatus, kuid, nagu ma allpool näidata püüan, selle abiga mittefüüsilisele vaimule põhjuslikku rolli anda arvatavasti ei saa. Nii või teisiti ei tee ükski eelmises peatükis toodud argumentidest vähem usutavaks võimalust, et vähemasti kõik füüsikalised sündmused on seletatavad füüsikalistes mõistetes, mistõttu pole mingit tarvidust võtta omaks kartesiaanliku dualismi.

Dualism, millest siin juttu tuleb, on *omadusdualism*​: teadvuselamus eeldab indiviidi omadusi, mis ei tulene tema füüsikalistest omadustest, kuigi võivad olla neist seaduspärases sõltuvuses. Teadvus on maailma nähtus, mis jääb tema füüsikaliste nähtuste piiridest välja. See ei tähenda, et teadvus oleks omaette substants; küsimus, mis tingimustel oleks alust kõnelda substantsidualismist, tundub mulle üpris ebaselgena. Me teame ainult niipalju, et siin maailmas on indiviididel ka fenomenaalsed omadused, mis on füüsikalistest omadustest ontoloogiliselt sõltumatud.

On olemas ka nõrgemat liiki omadusdualism, mida ei tohiks segi ajada positsiooniga, millest siin juttu on. Mõnikord öeldakse, et omadusdualism kehtib igas valdkonnas, kus omadused ei ole ei ise füüsikalised ega otseselt taandatavad füüsikalistele omadustele. Selles mõttes ei ole isegi bioloogiline kohastumus füüsikaline omadus. Kuid see on väga nõrgavormiline dualism. Kohastumuses ja teistes taolistes omadustes ei ole ontoloogilises mõttes midagi põhimõtteliselt uut, sest nad on mikrofüüsikaliste omaduste suhtes ikkagi loogiliselt superventsed. Niisugune omadusdualism on materialismiga täiesti ühitatav. Seevastu omadusdualism, mida pooldan mina, toob maailma põhimõtteliselt uusi omadusi. Kuna need omadused ei ole loogiliselt superventsed isegi mitte mikrofüüsikaliste omaduste suhtes, siis on nad mittefüüsikalised palju tugevamas mõttes. Mina pean omadusdualismist ja mittefüüsikalistest omadustest kõneldes silmas just seda tugevamat vaadet ja sõna “mittefüüsikaline” tugevamat tähendust.

Siiski jääb tõenäoliseks, et teadvus sünnib mingi füüsilise aluse pinnal, kuigi ta ei ole sellest alusest tuletatav. Nii jõuame seisukohale, et teadvus sünnib füüsilisest substraadist teatavate kontingentsete loodusseaduste varal, mis ise ei tulene füüsikaseadustest. Niisugust vaadet pooldavad kaudselt paljud, kes peavad end materialistideks. Sageli võib kuulda sõnu: “Loomulikult olen ma materialist; teadagi sünnib vaim ajus”. Sõna “sünnib” tuleks siin võtta kui vihjet. Näiteks ei öelda tavaliselt, et “õppimine sünnib ajus”, või kui mõnikord öeldaksegi, siis on sõnal “sünnib” siin ajaline mõte. Loomulikum on öelda, et õppimine on protsess, mis toimub ajus. Juba asjaolu, et vaim “sünnib” ajus (mitte ei “toimu”), ​**[lk.126]**​ osutab, et siin on mängus veel midagi, mis ulatub füüsikaliste faktide piiridest väljapoole.[[155]](#footnote-156)

Mõned arvavad, et seda vaadet tuleks pidada pigem materialistlikuks kui dualistlikuks, kuna ta eeldab fenomenaalsete faktide niivõrd tugevat seaduspärast sõltuvust füüsikalistest faktidest ning kuna ta jätab füüsika vallale autonoomia. Muidugi pole erilist mõtet vaielda nimetuste üle, kuid mulle tundub, et füüsikaliste faktide piiridest väljuvate kontingentsete faktide lisamine muudab üldtunnustatud materialistlikku maailmavaadet siiski küllalt oluliselt, et tulemus vääriks omaette nimetust. Kui materialismiks piisab juba kõigi faktide seaduspärasest seosest füüsikaliste faktidega, siis on ta küll väga väheütlev õpetus.

Kuigi see vaade on üks dualismi teisendeid, ei ole temas midagi ebateaduslikku ega üleloomulikku. Kõige parem on kujutada asja ette järgmisel viisil. Füüsika postuleerib maailmas mingi hulga fundamentaalseid omadusi: aegruum, mass, energia, laeng, spinn jne. Samuti tunnistab ta tervet hulka fundamentaalseid seadusi, mis määratlevad nende fundamentaalsete omaduste omavahelised seosed. Fundamentaalseid omadusi ei saa seletada mingite veel fundamentaalsemate omaduste kaudu, ja fundamentaalseid seadusi ei saa seletada mingite veel elementaarsemate seaduste kaudu — neid tuleb endid võtta kui elementaarseid. Kui aga fundamentaalsed seadused ja fundamentaalsete omaduste jaotus on paigas, siis tuleneb neist peaaegu kõik, mis maailmas toimub. Sellepärast nimetataksegi fundamentaalteooriat füüsikas mõnikord “kõige teooriaks” [​*theory of everything*​]. Kuid asjaolu, et teadvus ei ole füüsikaliste omaduste suhtes superventne, näitab meile, et see füüsikaline teooria ei hõlma siiski päris kõike. Selleks, et tuua teadvus fundamentaalteooria käsitlusalasse, tuleb tuua viimasesse juurde uusi fundamentaalseid omadusi ja seadusi.

Oma raamatus “Unistused lõplikust teooriast” (1992) tähendab füüsik Steven Weinberg, et füüsika fundamentaalteooria teeb eriliseks see, et ta liigub seletuste ahelaga kõige kõrgemale välja ja seletab kokkuvõttes kõike. Kuid ta peab möönma, et selline teooria ei pruugi seletada teadvust. Parimal juhul, ütleb ta, on võimalik seletada teadvuse “objektiivseid korrelaate”. “See ei ei ole võib-olla teadvuse seletus, kuid jõuab sellele üsna lähedale” (lk. 45). Kuid muidugi mitte küllalt lähedale. Ta ei seleta ikkagi kõike, mis maailmas toimub. Et jääda järjekindlaks, peame tunnistama, et tõeliselt lõplik teooria vajab veel üht komponenti.

Selleks on kaks võimalust. Võib tunnistada elamuse enda maailma fundamentaalseks omaduseks ja asetada ta ühte ritta aegruumiga, spinniga, laenguga jne. See tähendab, et mõned fenomenaalsed omadused tuleb tunnistada baasomadusteks. Teine võimalus on eeldada mingit teist uudsete fundamentaalsete omaduste klassi, milledest saab fenomenaalsed omadused tuletada. Nagu ülaltoodud argumendid on näidanud, ei saa nendeks olla füüsikalised omadused. Võib-olla on need mingit uut tüüpi mittefüüsikalised omadused, mille suhtes fenomenaalsed omadused on loogiliselt superventsed. Sellistel omadustel oleks ​**[lk.127]**​ elamusega samasugune seos nagu füüsikalistel baasomadustel mittebaasomadustega, näiteks temperatuuriga. Neid võiks nimetada *protofenomenaalseteks* omadusteks, sest nad ei ole ise fenomenaalsed, kuid fenomenaalsus võib sugeneda nende ühendusest. Muidugi on väga raske ette kujutada, missugune üks protofenomenaalne omadus võiks välja näha, kuid nende olemasolu ei saa välistada. Edaspidi kõnelen ma siiski nii, nagu oleksid otsitavad fundamentaalsed omadused ise fenomenaalsed.

Kus on uued fundamentaalsed omadused, seal on ka uued fundamentaalsed seadused. Meie juhul on fundamentaalseteks seadusteks *psühhofüüsika*​ seadused, mis ütlevad, kuidas fenomenaalsed (või protofenomenaalsed) omadused sõltuvad füüsikalistest omadustest. Kuna füüsikalised seadused moodustavad juba ise kinnise süsteemi, siis ei sekku need seadused füüsikaseadustesse, vaid on *superventsusseadused*​, mis ütlevad, kuidas füüsikalistest protsessidest sünnivad elamused. Kuna elamuse sõltuvust füüsikast ei saa, nagu nägime, tuletada füüsikaseadustest, siis peavad iga lõpliku teooria koosseisu kuuluma ka seda liiki seadused.

Loomulikult on meie ettekujutus sellest, kuidas asjaomane fundamentaalne teooria võiks välja nägema hakata või missugused need fundamentaalsed psühhofüüsikaseadused oleksid, praegusel etapil väga udune. Kuid meil on põhjust uskuda, et niisugune teooria on olemas. On vägagi põhjust uskuda, et füüsikaliste protsesside ja teadvuselamuste vahel on seaduspärane seos. Kuid iga seaduspärane seos peab toetuma fundamentaalsetele seadustele. Füüsika näide ütleb meile, et fundamentaalseadused on tavaliselt lihtsad ja elegantsed; sama tuleks oodata ka teadvuseteooria fundamentaalseadustelt. Kui meil on fundamentaalse füüsikateooria kõrvale kunagi ka fundamentaalse teadvuseteooria, siis on “kõige teooria” meil võib-olla tõepoolest käes. Me võime oodata, et füüsika ja psühhofüüsika alusseadustest ja fundamentaalsete omaduste jaotusest tulenevad kõik maailma faktid. Sellise teooria väljatöötamine ei kujune kindlasti lihtsaks, kuid peaks olema põhimõtteliselt võimalik.

See, mis toimub siin teadvusega, on mingis mõttes analoogiline sellega, mis juhtus XIX sajandil elektromagnetismiga. Elektromagnetilisi fenomene püüti seletada juba teada olevate füüsikaseaduste abil, kasutades mehhaanika printsiipe jms., kuid edutult. Selgus, et elektromagnetiliste fenomenide seletamiseks tuleb tuua sisse uusi fundamentaalseid omadusi nagu elektrilaeng[[156]](#footnote-157) ja elektromagnetjõud, ja Maxwell kehtestas uute fundamentaalseadustena elektromagnetismi seadused. Ainult nii sai uute fenomenide seletamine võimalikuks. Samamoodi ei piisa teadvuse seletamiseks füüsikalise teooria omadustest ja seadustest. Teadvuseteooriaks on tarvis uusi fundamentaalseid omaduseid ja seadusi.

See seisukoht on täies kooskõlas tänapäevase teadusliku maailmavaatega ja täiesti naturalistlik; maailm on endiselt fundamentaalsete omaduste võrgustik, neid omadusi suhestavad omavahel alusseadused, ja kõike tuleb viimselt ​**[lk.128]**​ seletada alusseaduste kaudu. Nagu Maxwelli juhul, nii ka siin ei ole toimunud muud, kui et omaduste ja seaduste varamu on täienenud. Veel enam, selles kontseptsioonis ei ole midagi, mis füüsikalisele teooriale kuidagigi vastu räägiks, ta vaid täiendab seda. Füüsikateooria on füüsikaliste protsesside teooria, psühhofüüsikateooria aga ütleb meile, kuidas nendest protsessidest sünnib elamus.

Ma nimetan vaadet, mida ma siin kaitsen, *naturalistlikuks dualismiks* — see nimetus tabab minu arvates asja vaimu. Ta on naturalistlik, kuna ta postuleerib, et kõik tuleneb baasomaduste ja -seaduste võrgustikust, ja kuna ta on ühitatav kõigi tänapäeva teaduse saavutustega. Ja nagu naturalistlikud teooriad teistes valdkondadeski, nii tunnistab ka see vaade, et teadvust on võimalik seletada looduse baasseaduste kaudu. Ta ei nõua, et teadvuses oleks midagi eriliselt transtsendentaalset; teadvus on looduslik fenomen nagu iga teinegi. Meie ettekujutus loodusest on vaid avardunud. Mõnikord peetakse sõna “naturalism” “materialismi” sünonüümiks, kuid mulle tundub, et maailma naturalistlikule mõistmisele võib truuks jääda ka siis, kui materialism alt veab. (Juhuks, kui lugeja selles kahtleb, viitan kinnituseks raamatu ülejäänud osale.) Mõni lugeja näeb selle positsiooni nimetuses võib-olla mõningat irooniat, kuid peaasi on, et selles kajastub põhimõtteline sõnum: dualismi omaksvõtt ei tähenda tingimata müstikasse langemist.

Mingis mõttes sarnanevad seesuguse dualismi pooldajad temperamendilt võib-olla pigem materialistidega kui muudliiki dualistidega: nad väldivad igasugust transtsendentaalset elementi ning püüavad alati leida looduslikke seletusi, eeldades, et käitumise põhjused on füüsikalised, samas kõigiti eitades “vaimu masinas” ning tänu sellele vältides traditsiooniliste dualistlike vaadete kõige ebatõepärasemaid järelmeid. Sageli võib kuulda, et kognitiiv- ja neuroteaduse edu diskrediteerib dualismi. Kuid see ei kehti ühtviisi kõigi dualismi variantide kohta. Kõigi nende edusammude aluseks on käitumise ja teiste füüsikaliste fenomenide füüsikalised seletused, ning seega pole naturalistliku dualismi ja materialismi vahel nende jaoks erinevust.

Lõpuks veel kaks märkust. Mõni ei saa võib-olla aru, miks ei saa elamust pidada füüsikaliseks omaduseks, kui ta on fundamentaalne. Kas füüsika ei olegi ju viimaks just teadus sellest, mis on tõeliselt fundamentaalne? Vastus: jah, kui defineerida füüsikat nii, siis kvalifitseerub elamus muidugi füüsikaliseks omaduseks ja superventsusseadused osutuvad füüsikaseadusteks. Ent kui mõista sõnu “füüsika” ja “füüsikaline” loomulikumal viisil, siis elamus füüsika mõisteks ei sobi. Elamus ei ole fundamentaalne omadus, mida postuleerimata füüsikud ei saaks luua oma teooriat välismaailmast; füüsika moodustab kinnise ja vastuoludeta teooria ka ilma elamuseta. Kui tunnistada sombimaailma võimalikkust, saab ilmseks, et füüsika oma tavapärses mõttes teadvust ei vaja. Seepärast on loomulikum pidada elamust fundamentaalseks, kuid mitte ​**[lk.129]**​ füüsikaliseks omaduseks ning psühhofüüsika seadusi fundamentaalseteks loodusseadusteks, kuid mitte füüsikaseadusteks. Ent senikaua kui asja sisu on selge, ei ole terminoloogilised küsimused kuigi olulised.

Tuleb ka mainida, et kuigi ma nimetan seda vaadet dualistlikuks, pole võimatu, et ta osutub hoopis mingitliiki monismiks. Võib-olla osutuvad füüsika ja fenomenaalsus ühe ja sama mõlemat hõlmava kategooria kaheks aspektiks, umbes samamoodi, nagu aine ja energia osutuvad ühe kategooria kaheks aspektiks. Selles, mida ma olen öelnud, pole midagi, mis seda välistaks, ja mingis mõttes on see mõte mulle isegi sümpaatne. Kuid tõsiasjaks jääb, et kui mingi monismi variant osutubki tõeseks, siis igatahes mitte materialistlik monism. Tarvis on midagi avaramat.

#### Vastuväited

Peatüki alguses esitatud materialismivastasele argumendile võidakse esitada mitmesuguseid vastuväiteid. Mõned neist ründavad eeldust 2), s.t. loogilise superventsuse eitust; neid vastuväiteid vaatlesin ma eelmises peatükis. Nüüd võtan ette vastuväited sammule, mis viis loogilise superventsuse puudumiselt materialismi vääruseni. Kõige tõsisemad neist toetuvad aposterioorsele paratamatusele. Nendest tuleb juttu järgmises peatükis. Siinses jaos vaatleme vähem tõsiseid vastuväiteid.

Mõnikord öeldakse, et teadvus võib olla *emergentne* omadus mingis niisuguses mõttes, mis on siiski ühitatav ka materialismiga. Hiljutistes töödes keerukate süsteemide​[[157]](#footnote-158) ja tehiselu kohta on sageli väidetud, et emergentseid omadusi pole võimalik ennustada madaltaseme omaduste põhjal, kuid et nad on sellegipoolest füüsikalised omadused. Näideteks võib tuua eneseorganisatsiooni emergeerumise bioloogilistes süsteemides või linnuparvele omaste käitumismustrite emergeerumise lihtsatest reeglitest silmulatsioonides (Langton 1990; Reynolds 1987). Kuid seesuguseid emergentseid omadusi ei saa pidada teadvuse analoogideks. Need juhtumid teeb huvitavaks asjaolu, et otseselt ei ole kõnealuste omaduste tulenevus madaltaseme seadustest ilmne, kuid madaltaseme faktide suhtes on nad sellegipoolest loogiliselt superventsed. Kui meil on olemas kõik füüsikalised faktid selletaoliste süsteemide ajalise muutumise kohta, siis saab neist otsesel teel tuletada ka iseenesliku organiseerumise fakti. See on ka ootuspärane, sest iseorganiseerumine ja parvede moodustumine on otseselt funktsionaalsed ja struktuursed omadused.

Kui teadvus on emergentne omadus, siis on ta seda palju tugevamas mõttes. Emergentsuse tugevama käsitluse järgi, millest kõnelevad briti emergentistid (näiteks Broad (1925)), ei ole emergentsed omadused ennustatavad, isegi kui meil oleks käepärast täielik komplekt füüsikalisi madaltaseme fakte. On mõistlik öelda (nagu briti emergentistid ütlesidki), et teadvuselamus on selles mõttes emergentne. Kuid seesugust emergentsust on kõige sobivam vaadelda ​**[lk.130]**​ kui üht omadusdualismi erikuju. Erinevalt ülaltoodud “süütumatest” emergentsuse näidetest nõuab see tugev kuju emergentsete omaduste ilmumiseks uusi fundamentaalseid seadusi.

Teine vastuväide ütleb, et teadvus ja füüsiline maailm võivad olla ühe ja sama asja kaks aspekti, nii nagu Koidutäht ja Ehatäht on Veenuse kaks aspekti. Kui see on nii, siis võib teadvus olla mingis mõttes füüsiline. Kuid taas tuleb küsida: kas fenomenaalne aspekt tuleneb füüsilisest aspektist? Jaatav vastus annab meile mingitliiki materialismi, kuid sellisel juhul jõuame uuesti ​3. peatükis toodud argumentide juurde. Eitava vastuse korral toob fenomenaalne aspekt maailma uut, füüsika piiridest kaugemale ulatuvat kontingentsust, ja aspektide duaalsus annab meile mingitliiki omadusdualismi. Pole võimatu, et neid kaht duaalset aspekti — füüsilist ja fenomenaalset — annaks ühendada mingi avarama monismi alla, kuid see ei oleks siis enam puhtalt füüsikaline monism.

Kolmanda vastuväite võib välja lugeda Searle’i raamatust (1992). Nagu minu, nii ka Searle’i arvates on teadvus füüsika suhtes superventne vaid looduslikult. Olles seisukohal, et aju seisundid on vaid teadvuse põhjus, möönab ta sombidest teisikute loogilist võimalikkust. Kuid ta ei ole nõus pidama seda vaadet dualismiks, isegi mitte omadusdualismiks. See võib näida puhtalt terminoloogilise küsimusena, ent kuna Searle rõhutab, et teadvusel on samasugune ontoloogiline staatus nagu füüsikalistel omadustel, näiteks vedelusel, siis ei ole asi siin ainult terminites. Põhjenduseks, miks säärane vaade ei ole dualistlik, ütleb Searle, et sarnast pilti võib näha ka mujal: näiteks H₂O põhjustab vedelust, kuid mitte keegi ei ole dualist vedeluse küsimuses.

See analoogia on aga üsna ilmselt väär. On loogiliselt võimatu, et kõik mikrofüüsikalised faktid mingi konkreetse H₂O kogumi kohta kehtiksid, ilma et ühtlasi realiseeruks vedelus. Mõte vedela H₂O koguse mittevedelast koopiast on lihtsalt mittekoherentne. Sellest järeldub, et mikrofüüsikaliste faktide ja vedeluse vaheline suhe on midagi palju enamat kui lihtne põhjuslik suhe. Mikrofüüsikalised omadused mitte ei põhjusta vedelust, vaid konstitueerivad selle. See on hoopis midagi muud kui see, mis leiab aset teadvuse puhul, ja järelikult ei pea analoogia paika. Teadvus on ontoloogiliselt palju olulisemas mõttes uudne nähtus kui vedelus.[[158]](#footnote-159)

Ja lõpuks võidakse öelda, et minu argument dualismi kaitseks meenutab Descartes’i argumenti. Descartes arutles, et kuna ta suudab kujutleda oma vaimu eksisteerivat kehast lahus, siis ei saa tema vaim olla identne tema kehaga. Üldise arvamuse järgi on seesugune mõttekäik vigane: ainuüksi sellest, et keegi võib endale ette kujutada, et A ja B ei ole identsed, ei järeldu, et A ja B ei ole identsed. (Mõelge näiteks Koidutähele ja Ehatähele.) Kas minu mõttekäik ei ole ehk samamoodi vigane? Sombimaailm näitab ainult seda, et on mõeldav, et keegi viibiks mingis füüsilises seisundis, ilma et tal oleks teadvust, kuid mitte seda, et füüsiline seisund ja teadvus ei ole identsed.

​**[lk.131]**​ Kui keegi nii mõtleb, siis on ta minu argumendist valesti aru saanud. Otsustav erinevus on selles, et minu sõnastuses ei käi see argument identsuse, vaid superventsuse kohta. Minu argument ei ole: “Füüsilist seisundit ​*P* on võimalik ette kujutada ilma teadvuseta, ja järelikult ei ole teadvus füüsiline seisund ​*P*​”, vaid: “On võimalik ette kujutada, et kõik füüsikalised faktid kehtivad, ilma et kehtiksid teadvuse faktid, ja järelikult ei ammenda füüsikalised faktid kõiki fakte.” See on hoopis teistsugune argument. Üldse on identsuse mõistest lähtuvad modaalsed argumendid dualismi kaitsmisel vähem veenvad kui superventsuse mõistest lähtuvad modaalsed argumendid; see on üks põhjusi, miks ma olen kogu aeg formuleerinud oma väiteid superventsuse mõiste abil ja pole peaaegu üldse maininud identsust. Mulle tundub, et just superventsust puudutavad küsimused on siin kõige fundamentaalsemad.

Selle argumendi vastu võidakse sellegipoolest püüda käiku lasta Descartes’ile antud vastusega analoogset strateegiat. Näiteks võidakse märkida, et minu strateegia toetub ikkagi järelduskäigule, mis liigub mõeldavuselt võimalikkusele, ja püüda seda loogikat kahtluse alla seada. Võtan selletaolised vastuväited vaatluse alla järgmises jaos​.

### 2. Vastuväited seoses aposterioorse paratamatusega

Üks populaarseid vastuväiteid seesugustele argumentidele on, et need tõendavad ainult, et sombimaailm on loogiliselt võimalik, see on aga midagi hoopis muud kui olla metafüüsiliselt võimalik. Kui loogiliseks võimalikkuseks piisab kontseptuaalsest koherentsusest, siis metafüüsiline võimalikkus esitab rangemaid nõudeid. Sama mõtet sõnastatakse mõnikord ka teisiti, öeldes, et mõeldavus ei ole sama mis tõeline võimalikkus. Sombimaailm on võib-olla tõepoolest mõeldav, kuid selleks, et näidata tema võimalikkust materialismi kummutamiseks tarvilikus metafüüsilises mõttes, on tarvis midagi enamat.

Enamasti kaasneb selle vastuväitega viide Kripke raamatule “​Nimetamine ja paratamatus” [​*Naming and Necessity*​] (1980) kus autor näitab, et on olemas paratamatud tõed, mille paratamatust on võimalik teada ainult *a posteriori* — näiteks “vesi on H₂O”. Nende vastuväidete esitajate sõnul on loogiliselt võimalik, kuid mitte metafüüsiliselt võimalik, et vesi ei ole H₂O. Ja täiesti loomulik on öelda analoogiliselt sombide kohta, et kui nad oleksidki loogiliselt võimalikud, metafüüsiliselt oleksid nad ikkagi võimatud. Ja sellest võiks materialismi päästmiseks juba piisata.

Materialistid, kes usuvad, et fenomenaalsed mõisted ei ole tuletatavad füüsikalistest mõistetest, rakendavad valdavas enamikus just seda strateegiat. Selle vaate järgi võib olla olemas kontseptuaalne lünk ilma metafüüsilise lüngata. Selline arusaam on ahvatlev selle poolest, et ta võimaldab jääda materialistiks, kuid võtta tõsiselt ka teadvust. Lähemal uurimisel osutub see vaade aga täiesti vigaseks. Aposterioorse paratamatuse mõiste ei suuda kanda koormat, mida see argument nõuab, ja juhib siinses kontekstis õigupoolest vaid tähelepanu kõrvale.[[159]](#footnote-160)

​**[lk.132]**​Kõige paremini võib seda taibata [2. peatüki 4. jaos](#ORG.2.4) välja töötatud aposterioorse paratamatuse kahemõõtmelise mudeli abil. Tuletame meelde, et selles mudelis on igal mõistel kaks intensiooni (s.t. kaks funktsiooni võimalike maailmade hulgast referentide hulka): ​*a priori*​ määratletav primaarne intensioon, mis fikseerib referendi aktuaalses maailmas, ja *a posteriori*​ määratletav sekundaarne intensioon, mis osutab referendile kontrafaktilistes maailmades. Sõna “vesi” primaarne intensioon on umbes nagu “vesijas ollus”. Sekundaarne intensioon on “H₂O”, mis on tuletatav primaarsest intensioonist Kaplani *dthat*​-operaatori abil: “​*dthat*​(vesijas ollus)”, osutab kõigis võimalikes maailmades H₂O-le, kuna vesijas ollus on aktuaalses maailmas H₂O.

Loogiline võimalikkus tähendab väite võimalikku tõesust, kui hinnata väite tõeväärtust mõistete primaarsete intensioonide alusel. (See on see võimalikkuse liik, mida ma [2. peatükis](#PARA.46-9) nimetasin 1-võimalikkuseks). Sõnade “vesi” ja “H₂O” primaarsed intensioonid on erinevad, seega on selles mõttes loogiliselt võimalik, et vesi ei oleks H₂O. “Metafüüsiline võimalikkus” tähendab väite tõesuse võimalikkust, kui hinnata seda mõistete sekundaarsete intensioonide alusel (s.t. 2-võimalikkust). Sõnade “vesi” ja “H₂O” sekundaarsed intensioonid ühtivad, ja seega on metafüüsiliselt paratamatu, et vesi on H₂O.

Kogu vastuväite mõte on seega, et me oleme mõeldavuse argumendiga ja teiste argumentidega näidanud sombimaailma võimalikkust asjaomaste mõistete primaarsete, kuid mitte sekundaarsete intensioonide järgi, mis on siin olulisemad. Kuigi fenomenaalsetel mõistetel ei pruugi primaarne intensioon vastata ühegi füüsikalise mõiste primaarsele intensioonile, võivad nende sekundaarsed intensioonid ühtida. Ja sellisel juhul võivad fenomenaalsed ja füüsikalised​/​funktsionaalsed mõisted oma aprioorsest erinevusest hoolimata ​*a posteriori* osutada samadele omadustele. Niisuguse vastuväite võib esitada *psühhofunktsionalismi* (vt. Block 1980) pooldaja, kes samastab fenomenaalsed omadused aposterioorselt funktsionaalsete omadustega, või niisuguse seisukoha pooldaja, mis samastab nad aposterioorselt mingite neurofüsioloogiliste omadustega.

Minu argumenti dualismi kasuks see kõik ei mõjuta, nagu võib hõlpsasti näha sellest, et argument töötab, kui võtta arvesse ainult primaarseid intensioone ja ignoreerida sekundaarseid. [2. peatükis](#PARA.45-9) nägime, et just primaarne intensioon on seletamisel kõige relevantsem, kuid ta teenib meid hästi ka dualismi kasuks argumenteerimisel. Sest pangem tähele, et olenemata sellest, kas primaarne ja sekundaarne intensioon ühtivad või mitte, on objektide omadused, mida primaarne intensioon võimalikes maailmades defineerib, täiesti laitmatud. Omadus olla vesijas ollus on täiesti mõistlik omadus, kuigi mitte sama mis omadus olla H₂O. Kui me suudame näidata, et mõnes meie maailmaga identses võimalikus maailmas mõni primaarse intensiooni alusel defineeritud omadus puudub, jõuamegi dualismini.

Just seda ongi tehtud teadvusega. Me oleme näinud maailmu, mis on füüsiliselt täpselt samasugused nagu meie oma, ent kus teadvust, kui mõista seda mõistet ​**[lk.133]**​ primaarse intensiooni alusel, ei ole. Sellest maailmade erinevusest piisab näitamaks, et meie maailmal on omadusi, mis ulatuvad füüsikaliste omaduste piiridest kaugemale. Kui me suudaksime analoogiliselt näidata, et on maailmu, mis on meie maailmaga füüsikaliselt identsed, ent kus pole vesijat ollust, siis oleksime kehtestanud vee suhtes dualismi niisama hästi nagu siis, kui oleksime tõestanud, et on olemas maailmad, mis on füüsiliselt identsed meie omaga, aga kus ei ole H₂O-d. Ja mis peaasi, erinevus primaarse intensiooni alusel on tuvastatav aposterioorsetest teguritest sõltumatult, mis tähendab, et aposterioorset paratamatust puudutavad kaalutlused ei puutu asjasse.

(Siinkohal tuleb teha kaks tehnilist märkust. Esiteks, kuna primaarne intensioon rakendub võimalikele keskpunktiga maailmadele, siis määrab ta objekti omaduse mingis võimalikus maailmas rangelt võttes selle maailma keskpunkti (või objektide ja keskpunktide) suhtes. Kuid meie oponent ei saa seda suhtelisust enda kasuks pöörata. Kui keskpunkti asukoht on kindlaks määratud, defineerib primaarne intensioon täiesti laitmatu mitteindeksiaalse omaduse, ja kõik [3. peatükis](#ORG.3.1) toodud argumendid töötavad isegi siis, kui superventsusbaas sisaldab keskpunkti asukohta. Näiteks isegi kui faktide hulka, mida Mary maailma kohta teab, kuuluvad faktid, mis puudutavad tema asukohta, ei ole tal ikkagi võimalik teada, mis tunne on näha punast.

Teiseks võib meid muretsema panna asjaolu, et sombimaailma keskpunktis teadvuse mõiste nähtavasti puudub, kuigi primaarse intensiooni rakendamine võib eeldada vastava mõiste olemasolu maailma keskmes. (Võiks isegi hakata kahtlema sombi mõiste rakendamise legitiimsuses!) Mina arvan, et olukord on veidi keerulisem — primaarsed intensioonid ei tarvitse nõuda esialgse mõiste olemasolu —, kuid igal juhul tarvitseb selleks, et neist kahtlustest täielikult vabaneda, vaid vaadelda *osalist sombimaailma*​, s.t. maailma, kus mina olen keskpunktis ja mul on teadvus koos kõigi selle juurde kuuluvate mõistetega, ent kus mõned teised inimesed on sombid.)

Aposterioorse paratamatuse irrelevantsust kinnitab veelgi tähelepanek, et teadvuse mõistel primaarne ja sekundaarne intensioon ühtivad. Selleks, et mingi seisund oleks aktuaalses maailmas teadvuselamus, peab temaga kaasnema fenomenaalne tunne, ja nii on see ka kontrafaktilistes maailmades. Vee mõiste primaarse ja sekundaarse intensiooni erinevus vastab asjaolule, et mõnes kontrafaktilises maailmas võib leiduda midagi, mis näeb välja nagu vesi ja on tajutav veena, kuid ei ole tegelikult vesi, vaid ainult vesijas ollus. Aga see, mis on tajutav teadvuselamusena, ongi teadvuselamus, ka kontrafaktilistes maailmades. Kõigis võimalikes maailmades tähendab teadvuselamus teatavat tunnet. (Midagi sarnast on öelnud ka Kripke, kuigi tema ei kõnele tähendusest, vaid olemuslikest omadustest.)

Ent isegi kui oponent kinnitab visalt, et teadvuse mõiste primaarne ja sekundaarne intensioon on erinevad, jääb argument ikkagi kehtima. Me lihtsalt koondame pilgu primaarsele intensioonile, mille abil viide fikseeritakse, nagu ülal kirjeldatud. Näiteks ​**[lk.134]**​ mõistes teadvust kui “​*dthat*​(fenomenaalne tunne on olemas)”, me lihtsalt koondame tähelepanu intensioonile “fenomenaalne tunne on olemas”. [3. peatükis](#ORG.3.1.1) toodud argumendid näitavad, et on olemas võimalik maailm, kus minu teisikul fenomenaalne tunne puudub. Seega on fenomenaalse tunde olemasolu fakt, mis ulatub füüsikaliste faktide ringist kaugemale ja kõrgemale, ja argument dualismi kaitseks jääb kehtima.[[160]](#footnote-161)

Kõige üldisemalt saab seda mõtet väljendada öeldes, et Kripke aposterioorses paratamatuses ei ole midagi, mis võiks ainsagi loogiliselt võimaliku maailma võimatuks teha, vaid ta ütleb meile ainult seda, et me oleme mõnda neist maailmadest valesti kirjeldanud, võttes sõnatarvituse aluseks vastavate mõistete primaarsed intensioonid sekundaarsete asemel, mis on siin kohasemad. *A priori* saaks ju pidada võimalikuks, et vesi ei oleks H₂O, vaid XYZ. Selleks tuleks kujutleda maailma, kus ookeane ja järvi täitev vedelik oleks XYZ. Kripke analüüs näitab meile aga, et asjaseisu tõttu aktuaalses maailmas on selle kujuteldava maailma kirjeldus maailmana, kus vesi on XYZ, vigane, sest me võtsime kirjelduse aluseks primaarsed intensioonid sekundaarsete asemel, mis on siin kohasemad. Rangelt võttes on see maailm, kus XYZ ongi vesijas ollus. Need kaalutlused ei saa näidata selle võimalikuna näiva maailma võimatust; nad lihtsalt näitavad, kuidas on õige seda maailma kirjeldada.

Nagu nägime [2. peatükis](#PARA.45-6), näitavad Kripke mõttekäigud, et sekundaarne intensioon ​$F\_{a}:W\rightarrow R$ võib erineda primaarsest intensioonist $f:W^{}\rightarrow R$. See asjaolu seab mõningaid aposterioorseid kitsendusi mõistete kohaldatavustingimustele, kuid relevantne maailmade ruum jääb alati konstantseks; ainus erinevus nende kahe funktsiooni argumentide vahel puudutab keskpunkti asukohta. Seega kuigi lausete võimalikkusel võibki olla kaht liiki, on maailmade võimalikkusi ainult üks relevantne liik.

Sellest järeldub, et kui on olemas mõeldav maailm, mis on füüsiliselt identne meie maailmaga, ent kus mingid meie maailma positiivsed omadused puuduvad, siis ei saa ükski mõistete, kasvõi teadvuse mõiste designaati puudutav kaalutlus välistada selle maailma metafüüsilist võimalikkust. Me võime lihtsalt unustada nende terminite semantika ning panna tähele, et vastavas võimalikus maailmas on midagi — nimetame me seda siis teadvuseks või mitte — selgelt puudu. Kripke ideed võivad meile parimal juhul öelda, kuidas seda maailma ja tema relevantseid omadusi tuleks kirjeldada, kuid selle maailma võimalikkust nad ei mõjuta; ent selleks, et dualismi argument töötaks, piisab juba ainuüksi sellise maailma võimalikkusest, olenemata sellest, kuidas me teda kirjeldaksime.

#### Alternatiivne strateegia

 Leidub ka hoopis teine viis, kuidas saaks püüda aposterioorsele paratamatusele toetudes dualismi vältida. Me võiksime öelda, et kui keegi väidab, et sombimaailm on meie maailmaga füüsiliselt identne, siis kirjeldab ta seda valesti. Nagu XYZ-maailm näib sisaldavat vett, kuid tegelikult ei sisalda, just samuti näib ka ​**[lk.135]**​ sombimaailm meie maailmaga füüsiliselt identsena, kuid ei ole seda. See kõlab võib-olla kummaliselt, kuid väidet on võimalik selgemini sõnastada. Oponent võib kinnitada, et on omadusi, millel on maailma füüsikalises koosseisus oluline koht, kuid mis ei ole füüsikalisele uurimisele kättesaadavad. Mõeldes “füüsikaliselt identsest” maailmast, kujutleme me tegelikult maailma, mis vaid tundub meie maailmaga identsena, kui rakendada talle füüsikalisi uurimismeetodeid, kuid on sellest erinev mingite tabamatute olemuslike omaduste poolest — nendesamade, mis tagavad teadvuse olemasolu.

Näiteks võib olla, et selleks, et mõnd objekti saaks pidada kontrafaktilises maailmas elektroniks, ei piisa sellest, et tal on teiste füüsikaliste entiteetidega samasugused põhjuslikud seosed nagu elektronil, vaid tarvilik on ka mingi tabamatu olemuslik “elektronsus”. Selle vaate järgi on elektroni mõiste midagi umbes niisugust nagu “​*dthat*​(entiteet, mis mängib elektroni rolli)”. Viite elektronidele fikseerib väline kirjeldus, seejärel aga tuleb viide rigidiseerida nii, et ta osutaks kontrafaktilistes maailmades samasuguse seesmise olemusega entiteetidele, olenemata sellest, kas need mängivad neis maailmades vastavat rolli või mitte, ja samas nii, et entiteedid, mis seal seda rolli mängivad, kvalifitseeruksid elektronideks ainult juhul, kui neil on ka sellekohane seesmine olemus. Sama võib kehtida ka teiste omaduste, näiteks massi kohta, mida võiks mõista kui “​*dthat*​(omadus, mis mängib massi rolli)”. Sellisel juhul jääks füüsikalisele teooriale, mis iseloomustab elektrone ja massi ainult väliselt, nende seesmine olemus kättesaamatuks. Ja siis võib selguda, et relevantsed olemuslikud omadused on ise fenomenaalsed või protofenomenaalsed omadused, millede realiseerumine võiks tagada teadvuse olemasolu meie maailmas.

Kui asi oleks nii, siis neid olulisi peitomadusi meie kujuteldavas sombimaailmas ei oleks, ja seega ei oleks see maailm meie maailmaga füüsiliselt identne. Sombimaailm osutuks samasuguseks nagu meie maailm, kui lähtuda teda vaadeldes väliste suhete alusel kohaldatavate füüsikaliste predikaatide primaarsetest intensioonidest, kuid meie maailmast erinevaks, kui võtta hindamise aluseks sekundaarsed intensioonid, mis eeldavad seesmist peitolemust. Sellisel juhul võib teadvuselamus viimaks ikkagi olla füüsikaliste omaduste suhtes superventne “metafüüsiliselt”. (Üsna sarnase argumendi on esitanud Maxwell (1978), ja millegi taoliseni viib välja ka Lockwoodi (1989) käsitlusviis. Maxwelli järgi on peamine idee selles, et kui fenomenaalsete mõistete kohta võib-olla ei saagi anda temaatiliselt neutraalseid​[[161]](#footnote-162) analüüse, mis osutaksid nende aluseks olevatele füüsikalistele omadustele, siis see-eest saab füüsikalistele mõistetele anda temaatiliselt neutraalsed analüüsid, mis osutavad nende aluseks olevatele fenomenaalsetele omadustele.[[162]](#footnote-163))

See vastuväide on mitmes mõttes huvitavam kui eelmine. Muidugi toetub ta spekulatiivsele metafüüsikale, kuid see ei tee takistusi tema koherentsusele. Otsesem vastus oleks, et see mõttekäik toetub ebaõigele arusaamale füüsika terminite semantikast. Pole võimatu, et füüsikaliste predikaatide kohaldamine aluseks on isegi aposterioorselt​ füüsikaliste entiteetide vahelised välised suhted, sõltumatult ükskõik millistest peitomadustest. See on puhtalt kontseptuaalne küsimus: kui elektronidel on meie maailmas mingid protofenomenaalsed ​**[lk.136]**​ peitomadused, kas me siis nimetaksime elektroniks kontrafaktilist entiteeti, millel need omadused puuduvad, kuid mis on elektroniga muidu igati identne? Minu arvates küll. Viidet elektronidele ei fikseeri mitte ainult roll, mida nad teoorias etendavad; see roll määratleb juba elektroni mõistegi ja determineerib selle mõiste rakendatavuse kõigis maailmades. Mõte elektronist, mille välised omadused ühtiksid aktuaalsete prootonite väliste omadustega, ei tundu olevat koherentne, nagu ka ettekujutus maailmast, kus mass etendaks rolli, mida aktuaalses maailmas etendab laeng. Ülal formuleeritud semantiline käsitlus ennustaks, et säärased ettekujutused on koherentsed, ja käsitleks neid mõisteid seega ebaõigesti.

Semantika vallas võib intuitsioon meile mitmesuguseid asju öelda, kuid, nagu tavaliselt ikka, leidub ka vastus, mis küünib sellest kõigest sügavamale. Isegi kui möönda, et füüsikaliste omaduste hulka võivad kuuluda ka peitomadused, ei erine niisugune vaade kuigi palju omadusdualismist, mida mina siin kaitsen. Endiselt jääb õigeks, et maailmal on fenomenaalseid omadusi, mida füüsikalise uurimise teel leitavad omadused ei fikseeri. Pärast seda, kui Jumal garanteeris, et üks või teine maailm on meie füüsikateooriate mõttes meie maailmaga identne, pidi ta tegema lisatööd, et teha see maailm meie omaga identseks igas suhtes. Füüsikaliste ja mittefüüsikaliste omaduste dualismi asendab selles kontseptsioonis “kättesaadavate” ja “tabamatute” füüsikaliste omaduste dualism, kuid asja olemus sellest ei muutu.

Arusaama juurde, et füüsikalistel entiteetidel on seesmine protofenomenaalne olemus, tulen ma [allpool](#PARA.96-7) veel tagasi, kuid selle vaate metafüüsiline külg ei olene kuigivõrd füüsikaliste predikaatide semantika tõlgendamisest. Nagu ennegi, tuleneb sekundaarsetest intensioonidest ja aposterioorsest paratamatusest ainult semantilisi, mitte metafüüsilisi erinevusi. Ükskõik kuidas seda arusaama ka formuleerida, tunnistab ta fenomenaalsed või protofenomenaalsed omadused fundamentaalseteks ja on seega lähem mingile dualismi versioonile (või ehk mingile idealismi või neutraalse monismi vormile — sellest tuleb [allpool](#PARA.96-8) veel juttu) kui ühelegi materialismi versioonile.

#### Tugev metafüüsiline paratamatus

[Ülal](#PARA.84-4) käsitletud kahemõõtmeline analüüs näitab, et Kripke aposterioorne paratamatus on superventsuse argumendi vastu jõuetu. Selline paratamatus ei sea võimalike maailmade ruumile aposterioorseid kitsendusi, vaid kitsendab ainult mõnede terminite kasutust nende kirjeldamisel. Ja see tähendab, et kui on olemas loogiliselt võimalik maailm, mis on meie maailmaga identne kõigis füüsikalistes, kuid mitte kõigis positiivsetes aspektides, siis ei ole nende kaalutluste põhjal võimalik eitada sellise maailma metafüüsilist võimalikkust.

Selle vastu võidakse aga öelda, et vastavad maailmad võivad ikkagi olla metafüüsiliselt võimatud; et on olemas loogilisest võimalikkusest erinev, kitsam modaalsus — metafüüsiline võimalikkus —, ​**[lk.137]**​ ning et selle modaalsuse olemasolu tuleneb asjaoludest, mis on Kripke ideedest sõltumatud. Selle seisukoha järgi on metafüüsiliselt võimalikke maailmu vähem kui loogiliselt võimalikke maailmu, ja mingite väidete aposterioorne paratamatus võib tuleneda teguritest, mis on asjaomaste terminite semantikast täiesti sõltumatud. Me võime nimetada seda hüpoteetilist modaalsust *tugevaks metafüüsiliseks paratamatuseks*​, vastupidiseks *nõrgale metafüüsilisele paratamatusele*​, millega opereerib Kripke mudel.

Selle seisukoha järgi võib olla maailmu, mis on isegi kõige rangemate kriteeriumide järgi mõeldavad, kuid sellegipoolest täiesti võimatud. Mõeldavuse ja võimalikkuse vaheline lünk on palju suurem kui kõik teised lüngad. Mingis mõttes võib väidete, nagu näiteks “vesi on XYZ”, tõesus olla mõeldav, kuid mitte võimalik, kuid sellised näited ei välista kunagi ühegi mõeldava maailma võimalikkust — neil juhtudel on vaid maailma kirjeldus vigane. Tugev metafüüsiline paratamatus läheb kaugemale. Siin võib sõna “sombimaailm” olla mõeldava maailma õige iseloomustus, ja seda isegi sekundaarsete intensioonide järgi. Sellegipoolest ei ole see maailm metafüüsiliselt võimalik.[[163]](#footnote-164)

Lühike vastus sellele vastuväitele on, et meil puudub alus uskuda niisuguse modaalsuse olemasolusse. Sellised “metafüüsilised paratamatused” seavad võimalike maailmade ruumile karme ja seletamatuid kitsendusi. Karme ja seletamatuid fakte võib olla mõistlik tunnistada meie maailma piires, kuid selliste faktide olemasolu võimalike maailmade ruumis oleks üpris veider. Võimalikkuse valdkonnas (erinevalt looduslikust valdkonnast) ei ole sellistele meelevaldsetele kitsendustele kohta.

Seda seisukohta ei saa põhjendada ka ühegi analoogiaga, sest sobivaid analoogiaid ei leidu kusagil.[[164]](#footnote-165) Me nägime juba, et analoogia lausete “vesi on H₂O”, “Hesperos on Phosphoros” jne. paratamatusega ei pea paika, sest kõik need näited eeldavad ühtainsat maailmade ruumi. Tõepoolest, kui mõni maailm ongi loogiliselt võimalik, kuid metafüüsiliselt võimatu, siis ei ole meil vist küll mingit võimalust seda teada saada: eelduse kohaselt ei ole see informatsioon ​*a priori* kättesaadav, aposterioorset informatsiooni saab meil aga olla ainult meie endi maailma kohta; see informatsioon võib aidata meil lokaliseerida meie maailma võimalike maailmade ruumis, kuid pole selge, kuidas ta võiks anda meile informatsiooni võimalike maailmade ruumi mahu kohta. Iga väidetav lisakitsendus metafüüsilisele võimalikkusele näiks meelevaldse postulaadina; sama hästi võiks postuleerida, et kui kivist lahti lasta, siis on metafüüsiliselt võimatu, et ta hakkaks liikuma ülespoole.

Lisaks toob see seisukoht kaasa modaalsuste *ad hoc* vohamise. Säärast vaadet omaks võttes tuleks aktsepteerida lausungite lausa nelja liiki võimalikkust ja paratamatust, isegi kui looduslik modaalsus kõrvale jätta: võimalikkus ja paratamatus vastavalt primaarsetele või sekundaarsetele intensioonidele nii loogiliselt võimalike kui ka metafüüsiliselt võimalike maailmade ruumis. Aga kui vaadelda väidete võimalikkuse asemel maailmade võimalikkust, saaksime kolm objektiivset võimalike maailmade klassi: loogiliselt võimalikud maailmad, metafüüsiliselt võimalikud maailmad ja looduslikult võimalikud maailmad. Meil on põhjust uskuda esimesse ja viimasesse neist klassidest, kuid väga vähe põhjust uskuda erilisse kolmandasse, metafüüsilisse klassi.

See, kes peab sombimaailma loogiliselt võimalikuks, kuid metafüüsiliselt võimatuks, peab vastama võtmeküsimusele: miks ei oleks Jumal saanud luua sombimaailma? Võiks ju arvata, et Jumala võimuses on teha maailma luues kõike, mis on loogiliselt võimalik. Kuid metafüüsilise paratamatuse pooldaja peab kas ütlema, et võimalikkus on loogiliselt koherentne, kuid Jumal ei oleks saanud seda luua, või et Jumal oleks saanud seda luua, kuid et see on ikkagi metafüüsiliselt võimatu. Esimene seisukoht on täiesti põhjendamatu, teine aga täiesti meelevaldne. Ent isegi kui oletada, et teine võimalus peab paika, jääb materialismivastane argument ikkagi kehtima: pärast maailma füüsikaliste faktide fikseerimist pidi Jumal tegema veel midagi.

Isegi kui seda seisukohta aktsepteerida, sarnaneks ta paljudes olulistes aspektides vägagi omadusdualismiga, mida pooldan mina. Selle seisukoha järgi on teadvuse olemasolu samuti võimatu tuletada füüsikalistest teadmistest, ja seega ei ole teadvus reduktiivselt seletatav. Ja nagu enne, nii ka nüüd on tarvis mingeid ühendavaid algprintsiipe, et seletada fenomenaalse maailma superventsust füüsikalise maailma suhtes. Ainus erinevus nende kahe vaate vahel on selles, et asjaomased psühhofüüsikaprintsiibid ei ole loodusseadused, vaid “paratamatuse seadused”. Teooria konstrueerimisel jääme asja seletuslike aspektide mõttes samasse olukorda nagu omadusdualismiga; erinevus on vaid ühesainsas ontoloogilises postulaadis.

Selle vaate ainsaks tõeliseks motiiviks näib olevat püüd iga hinna eest päästa materialism — võib-olla selleks, et vältida näilisi raskusi dualismiga. Kuid selline materialism tundub palju mõistatuslikumana kui dualistlik alternatiiv. “Metafüüsiliselt paratamatute” algprintsiipide lisamine võimalike maailmade ruumi kitsendamiseks toob sisse elemendi, mis on palju problemaatilisem ja isegi palju vähem naturalistlik kui vaid uued loodusseadused, mida postuleerib omadusdualism. Lõppkokkuvõttes on uue paratamatusastme appivõtmine *ad hoc* eeldus, mis lahendab ühed probleemid, kuid seab meie ette sama palju teisi. Materialismi päästmise hind on, et jääb täielikuks mõistatuseks, kuidas teadvus saaks olla füüsiline.[[165]](#footnote-166)

#### Kognitiivsed kitsendused

Materialistil, kes peab sombimaailma mõeldavaks, kuid soovib siiski materialismi päästa, on üks viimane võimalus veel. Ülalkirjeldatud positsiooni korral tunnistab materialist sombi mõiste täiesti koherentselt kohaldatavaks isegi maksimaalselt ratsionaalsele olendile, kuid eitab samas ikkagi sombide metafüüsilist võimalikkust, jõudes seega “kahekihilise” pildini loogiliselt ​**[lk.139]**​ ja metafüüsiliselt võimalikest maailmadest. Kuid materialist võib asuda ka seisukohale, et näilise mõeldavuse allikaks on mingi suhtes puudulik ratsionaalsus: kui me oleksime vaid intelligentsemad, näeksime kohe, et maailma niisugune kirjeldus ei ole ikkagi koherentne. Selle seisukoha järgi ei ole sombimaailm juba loogiliseltki võimalik, meie aga usume ekslikult tema võimalikkusesse lihtsalt sellepärast, et meie kognitiivsetel võimetel on piirid. (Niisugune on üks McGinni (1989) vaadete võimalikke tõlgendusi.)

Sellele seisukohale võiks otsida tuge analoogiast keerukate matemaatiliste tõdedega, mis ületavad meie matemaatilise mõistmisvõime. Kui meie matemaatilised võimed on algoritmilised, peab niisuguseid tõdesid (Gödeli teoreemi järgi) leiduma, ent neid võib olla ka vastasel korral. (Näiteks võib tuua Goldbachi konjektuuri, võib-olla ka kontiinumihüpoteesi tõesuse või vääruse probleemi). Need on paratamatud tõed, kuigi meil ei ole võimalik neid ​*a priori* teada ja nad ei toetu ka ​*a priori* teada olevate ja empiiriliste tegurite kombinatsioonile nagu Kripke paratamatused​.[[166]](#footnote-167) Võib-olla ongi fenomenaalsete faktide implitseerumine füüsikalistest faktidest seesugune paratamatus, mis mingil moel ületab meie modaalse mõistmisvõime piirid?[[167]](#footnote-168)

See analoogia on aga ebatäielik. Matemaatika asjus jätab meie modaalne mõtlemine küsimuse lahtiseks; intuitsioon ei aita meil selles küsimuses valikut teha. Mingis nõrgas mõttes on ehk “mõeldav”, et mainitud väited on väärad — seda nad ju on, niipalju kui me teame —, kuid mitte selles mõttes, mis lubab pidada mõeldavaks maailma, kus nad ei kehtiks. Sombidega aga ei jää küsimus lahtiseks; tundub olevat ilmne, et me suudame kujutleda maailma, milles see järeldus on väär. Materialismi päästmiseks tuleb säärase maailma võimalikkus välistada, kuigi meie modaalsed võimed kõnelevad selle kasuks; kuid matemaatika puhul pole märkigi võimalusest leida näide, mille varal saaks sel viisil välistada pealtnäha võimaliku maailma. Ja igasugune lõhe mõeldavuse ja võimalikkuse vahel, mida materialist siin ära kasutada saaks, peab jällegi olema *sui generis* lünk, mida ei kinnita ükski asjakohane analoogia üheski teises valdkonnas.[[168]](#footnote-169)

Loomulikult võib materialist astudagi selle meeleheitliku sammu ja väita, et on olemas mingi *sui generis* kognitiivne puudulikkus. Selleks peaks ta tunnistama, et [3. peatükis](#ORG.3) esitatud argumendid on ekslikud viisil, mida me ei suuda mõista. Vähe sellest, et selle seisukoha järgi on meie intuitiivsed arusaamad selle kohta, mis on mõeldav ja mis mitte, rängalt ekslikud — lisaks eeldab ta ka veel seda, et Mary võimekam versioon võib füüsikalise informatsiooni põhjal tõepoolest teada saada, mis tunne on näha punast, ning et fenomenaalseid mõisteid on võimalik analüüsida niisugusel (võib-olla struktuursel või funktsionaalsel) viisil, mis kinnitaks fenomenaalsete faktide tuletatavust füüsikalistest faktidest, kuid mille õigsuse hindamiseks meil ei jätku võimeid.

Kuigi ei saa eitada, et iga filosoofiline argument võib kognitiivse puudulikkuse tõttu ekslik olla, tundub see vastuväide ​**[lk.140]**​ üsna meelevaldsena, kuna puudub igasugune tõsine alus selle uskumiseks. Nagu ennegi, nii ka siin on peamiseks motiiviks nähtavasti soov jääda iga hinna eest truuks materialismile. Niisugust varianti tuleks alati kaaluda kõige viimases järjekorras, alles siis, kui me oleme tunnistanud oma võimetust vastata tõsistele argumentidele, mis osutavad meie vigadele, ja leidnud, et kõik tõsised katsed materialismile alternatiivi leida on lootusetud. Niipea kui meil õnnestub leida tõsine ja rahuldav alternatiiv, kaob ka igasugune põhjus niisugusele seisukohale asumiseks.[[169]](#footnote-170)

### 3. Veel argumente dualismi kasuks\*

Ma pole esimene, kes kasutab materialismi vastu loogilise võimalikkuse argumenti.[[170]](#footnote-171) Ent kuigi oma ühel või teisel kujul ongi see minu arvates õigupoolest peamine materialismivastane argument kogu vaimufilosoofias, ei ole talle osutatud nii palju tähelepanu, nagu ta vääriks. Rohkem on tähele pandud kaht teist antimaterialistlikku argumenti — Jacksoni (1982) ja Kripke (1972) oma. Need argumendid on, nagu mulle tundub, asjalikud, kuid vahest mitte nii fundamentaalsed. Jacksoni argumenti võib pidada oluliseks kui sissejuhatust loogilise superventsuse argumendile, Kripke argumendi kõige veenvam osa aga sõltubki, nagu kohe näeme, loogilise superventsuse argumendist.

#### Jacksoni argument

Jacksoni argumenti ehk teadmiseargumenti vaatlesin ma juba [ülal](#ORG.3.1.4) seoses loogilise superventsuse puudumise tõestusega, kus tal oli abistav roll. Meenutagem: jutt on neuroteadlasest Maryst, kes on sündimisest saadik elanud mustvalge toas ja samas teab kõiki füüsikalisi fakte värvusinfo töötluse kohta ajus. Hiljem, kui ta näeb esimest korda punast objekti, saab ta teada mingeid uusi fakte. Muu hulgas saab ta teada, mis tunne on näha punast. Argument lõpeb järeldusega, et füüsikalised faktid ei ammenda kõiki fakte ning et materialism on väär.

Sellel argumendil on palju ühist sombiargumendiga ja invertspektrite argumendiga: mõlemad toetuvad asjaolule, et fenomenaalseid fakte ei saa tuletada füüsikalistest faktidest. Mingis mõttes on need ühe ja sama argumendi kaks vastaskülge. Otsese materialismivastase argumendina peetakse Jacksoni argumenti aga sageli kergesti rünnatavaks, kuna see kasutab teadmise intensionaalset mõistet. Paljud oponendid ründavadki seda argumenti intensionaalsuse küljelt — näiteks kui väidetakse, et üht ja sama fakti on võimalik teada kahel erineval viisil. Minu arvates ei taba need rünnakud märki, kuid kõige lihtsam viis seda näha on vaadelda otseselt superventsuse puudumist, väljendades seda mitte epistemoloogilistes, vaid metafüüsilistes mõistetes. Minu mudel aitab mõista, mis põhjusel nimelt mitmesugused sellele argumendile esitatud vastuväited ei tööta. Järgnevalt vaatlen mõningaid neist vastuväidetest.

​**[lk.141]**​ Kõigepealt on mitmed oponendid öelnud, et kuigi Mary saab punast nähes uue teadmise, ei vasta see teadmine ühelegi uuele faktile. Tänu teadmise intensionaalsusele saab ta lihtsalt teada juba tuntud fakti uuel viisil, uues “presentatsioonimooduses” (Church­land 1985; Hor­gan 1984b; Ly­can 1995; Mc­Mullen 1985; Pa­p­ineau 1993; Teller 1992; Tye 1986). Näiteks Tye ja Lycan osutavad intensionaalsele erinevusele otsustuste “see vedelik on vesi” ja “see vedelik on H₂O” vahel: need väljendavad mingit mõttes sama fakti, kuid on võimalik teada üht, aga mitte teist. Analoogiliselt kõneleb Churchland lüngast temperatuuri teadmise ja keskmise kineetilise energia teadmise vahel, Horgan arutleb, mille poolest erinevad omavahel teadmine Clark Kentist ja teadmine Supermanist, McMullen aga juhib tähelepanu Mark Twainile ja Samuel Clemensile.

Need lüngad tulenevad just nimelt primaarse ja sekundaarse intensiooni erinevusest. On võimalik teada midagi vee kohta, teadmata seda H₂O kohta, sest nende mõistete primaarsed intensioonid on erinevad — mõtetel veest ei ole mingit aprioorset seost mõtetega H₂O-st. Kuid sellegipoolest on mingis mõttes olemas ainult üks faktide kogum nende mõlema kohta: vee ja H₂O sekundaarsed intensioonid ühtivad, kuna need kaks mõistet on aposterioorselt identsed. (Ei ole ilma pikemata selge, et fakte tuleb tingimata individualiseerida nii, et vee faktid ja H₂O faktid oleksid samad faktid, kuid ma nõustun sellega praegu mõttekäigu piires.[[171]](#footnote-172)) Varem tarvitusele võetud terminoloogias võib öelda, et otsustus “kui miski on vesi, siis on ta H₂O” on loogiliselt kontingentne, kuid metafüüsiliselt paratamatu. Seega ei erine vaadeldav vastuväide sisuliselt [juba varem käsitletud](#ORG.4.2) vastuväitest, mis puudutas erinevust loogilise paratamatuse ja (Kripke) metafüüsilise paratamatuse vahel, ning vastuväite kummutamiseks piisab viitest [sealsele](#PARA.84-4) arutelule primaarsete ja sekundaarsete intensioonide kohta.

Seda mõtet võib sõnastada ka otsesemalt: alati, kui keegi teab mingit fakti ühes, aga mitte teises presentatsioonimooduses, on olemas veel üks fakt, millest tal teadmist ei ole — fakt, mis neid kaht presentatsioonimoodust ühendab.[[172]](#footnote-173) Kui inimene teab, et Hesperos on nähtav, kuid ei tea, et Phosphoros on nähtav (kuna ta ei tea, et Hesperos on Phosphoros), siis ei tea ta ka, et hommikutaeva kõige heledam täht ja õhtutaeva kõige heledam täht on üks ja sama objekt. See on omaette fakt, millest tal pole mingit teadmist. Või kui keegi teab, et Superman oskab lennata, kuid ei tea, et Clark Kent oskab lennata, siis ei tea ta ka seda, et on olemas inimene, kes on “Daily Planeti” esireporter ja kannab keepi. Kui inimene teab, et vesi on märg, kuid ei tea, et H₂O on märg, siis ei tea ta ka seda, et järvi täitev ollus koosneb H₂O molekulidest. Ja nii edasi.

Formaalsemalt: Oletame, et “$a$ on $G$” ja “$b$ on $G$” on selles mõttes üks ja sama fakt, kuid nende kahe fakti vahel ei ole aprioorset seost. Siis peab põhjus olema selles, et $a=b$ ning et mõistete $a$ ja $b$ sekundaarsed intensioonid ühtivad, kuid primaarsed intensioonid on erinevad: näiteks on võimalik, et $a$ on ekvivalentne $dthat\left(P\right)$-ga ja $b$ on ekvivalentne $dthat\left(Q\right)$-ga. Kui subjekt ​**[lk.142]**​ teab, et $a$ on $G$, kuid ei tea, et $b$ on $G$, siis puudub tal faktiline teadmine, et on olemas miski, mis on korraga nii $P$ kui ka $Q$. Üldisemalt: tal puudub faktiline teadmine, et miski on ühekorraga nii $P’$ kui ka $Q$, kus $P’$ ja $Q’$ on suvalised identifitseerivad kirjeldused, mille põhjal on võimalik teada, et $a$ on $P’$ ja et $b$ on $Q’$. Sellel faktil ei ole midagi ühist faktidega, mida subjekt esialgselt teadis. Isegi sekundaarsete intensioonide alusel tõlgendades võib kõnelda võimalikust maailmast, kus $a$ on $F$, kuid mitte miski ei ole ühtaegu nii $P$ kui ka $Q$ (või nii $P’$ kui ka $Q’$).

(Nagu ​2. jaos​, nii ka siin teeb asja keerulisemaks see, et P ja Q võivad olla indeksikaalsed omadused. Kuid põhimõtteliselt ei muutu sellest midagi. Selleks, et teha uudne, seni teadmata fakt rangelt mitteindeksikaalseks, tuleb vaid tuletada temast fakt “On olemas punkt (omadusega ​*X*​), kust vaadates *P* ja *Q* osutavad ühele ja samale asjale”. *X* on varuvariant juhuks, kui subjekt peaks teadma veel mõnd teist kohta, kust vaadates *P* ja *Q* osutavad samale asjale; sellisel juhul me lihtsalt määratleme omaduse *X* nii spetsiifiliselt, et ta kehtiks ainult selles ühes kohas, aga mitte mujal. Äärmuslik juht, kus inimesel puudub igasugune eristav teadmine iseendast, taandub puht-indeksikaalsele juhule, millest tuleb juttu allpool.)

Siit järeldub, et kui Mary omandab mingi faktilise teadmise, mida tal varem ei olnud — isegi kui see on vaid vana fakti teadmine teistsuguses presentatsioonimooduses —, siis pidi ta nüüd saama teada mingi päriselt uudse fakti. Vähemalt pidi ta sama teada mingi uue fakti selle presentatsioonimooduse kohta. Kuna ta kõiki füüsikalisi fakte juba teadis, siis järelikult on materialism väär, füüsikalised faktid ei ammenda mitte üheski mõttes kõiki fakte.

See vastus võib tunduda vähem sirgjoonelisena kui vastav vastus loogilise võimalikkuse argumendile. Superventsuse mõistele toetuv mudel elimineerib vähem selgepiirilise küsimuse üksikute teadmiste individualiseerimisest ning teeb seega arutelu vähem segaseks. Ent sellegipoolest näitab põhjalik analüüs, et analoogia vee ja H₂O-ga ning sellele toetuvad vastuväited osutuvad mõlemal juhul ühtviisi ekslikeks. Kuigi see on kahtlemata kõige populaarsem vastuväide teadmise argumendile, on see sama kahtlematult ka kõigist peamistest vastuväidetest kõige nõrgem. Ta lihtsalt ei kannata lähemat uurimist.

Teine, peenem vastuväide, mis kuulub Loarile (1990), eeldab samuti, et Mary saab vanadest faktidest tänu intensionaalsusele uusi teadmisi, kuid ei jää pidama tavaliste analoogiate juurde vee ja H₂O-ga. Loar tunnistab, et selletaolised tavapärased analoogiad ei saa materialisti aidata, sest nad möönavad niisuguste olukordade võimalikkust, kus (kui kasutada meie terminoloogiat) füüsikalistel ja fenomenaalsetel mõistetel on erinevad primaarsed intensioonid, ja antimaterialist võib lihtsalt kohaldada oma argumenti omadusele, mis vastab primaarsele intensioonile. Loar sõnastab seda nii: isegi kui “soojus” ja mingi statistilis-mehhaaniline predikaat tähistavad üht ja sama omadust (sekundaarne intensioon), on omadused, mida nad kirjeldavad (primaarne intensioon), sellegipoolest erinevad. Seepärast liigub ta oma argumendiga edasi ja väidab, et kaks predikaati võivad kirjeldada üht ja sama omadust — s.t. neil võib olla sama primaarne intensioon — isegi siis, kui seda samasust ei ole võimalik ​*a priori* teada. Kui see on nii, siis võib Mary teadmine fenomenaalsetest omadustest ​**[lk.143]**​ olla lihtsalt teadmine füüsikalistest/funktsionaalsetest omadustest, isegi kui ta ei osanud varem neid kaht kokku viia.

Aga kuidas saavad kaks primaarset intensiooni ühtida, kui meil ei ole võimalik sellest ​*a priori* teada? Ainult juhul, kui võimalike maailmade ruum on väiksem, kui me oleksime ​*a priori* arvanud. Me peame intensioone erinevateks, sest me peame mõeldavaks maailma, kus nende referendid on erinevad — näiteks sombimaailma. Loari seisukoht eeldab seega, et säärane maailm ei ole tegelikult võimalik, kuigi me ei saa seda võimalust kontseptuaalsetel põhjustel välistada ja kuigi Kripke aposterioorne paratamatus meid siin kuidagi ei aita. Seega ei erine see positsioon põhimõtteliselt mitte millegi poolest ülalkäsitletud “tugeva metafüüsilise paratamatuse” vastuväitest, ja tema kohta kehtib täpselt samasugune kriitika. Nagu too vastuväide, nii nõuab ka Loari seisukoht, et konditsionaal füüsikalistelt faktidelt fenomenaalsetele faktidele​[[173]](#footnote-174) oleks hoolimata oma loogilisest kontingentsusest metafüüsiliselt paratamatu, kusjuures seda lünka ei saa seletada primaarsete intensioonide erinevusega. Nagu too vastuväide, nii eeldab ka Loari positsioon võimalike maailmade ruumi drakoonilist ja meelevaldset kitsendamist, mida Loar mitte millegagi ei põhjenda[[174]](#footnote-175).[[175]](#footnote-176) [[176]](#footnote-177)

Võiks oodata, et siin saaks esitada ka kolmanda vastuväite, mis oleks analoogiline [2. jaos](#ORG.4.2.1) kirjeldatud vahepealse “alternatiivse strateegiaga”. See viiks välja väiteni, et Mary tegelikult ei tea kõiki füüsikalisi fakte. Ta teab kõiki füüsika mõistetes väljendatud fakte, kuid tal puuduvad teadmised füüsikaliste entiteetide vaadeldamatute (fenomenaalsete või protofenomenaalsete) olemuste kohta. Kui tal oleksid need teadmised, teaks ta seevaral ka fenomenaalseid fakte. Kuid sedagi seisukohta saab vaid suurivaevu materialismiks nimetada. Nagu minu enda positsioon, nii peab ka see tunnistama fenomenaalsed või protofenomenaalsed omadused fundamentaalseteks omadusteks.

Neljas vastuväide kõneleb seoses Mary olukorraga indeksikaalse teadmise puudumisest (Bigelow ja Par­get­ter 1990; Mc­Mullen 1985; Pa­p­ineau 1993) ning ütleb, et kuigi Mary omandab uusi teadmisi, ei ole tema juhtum mõistatuslikum kui teised juhtumid, kus inimene, kes teab kõiki relevantseid objektiivseid fakte, avastab midagi uut: näiteks kõiketeadev amneesik, kes avastab, et “Rudolf Lingens olen mina”, või unetu, kellel on laialdased teadmised, kuid kes ei tea, et praegusel hetkel on kell 3:49 (vt. Perry 1979 ja Lewis 1979). Kõigil neil juhtudel on füüsikalise teadmise ja indeksikaalse teadmise vahel samasugune lõhe nagu Maryl füüsikalise teadmise ja fenomenaalse teadmise vahel.

Seda seost võib otsida kahel moel. Esiteks võib vastuväite esitaja püüda taandada fenomenaalne teadmine indeksikaalsele teadmisele ning väita, et Maryl ei puudu midagi peale indeksikaalse teadmise. Teiseks võib otsida kahe juhtumi vahel analoogiat ning püüda näidata, et indeksikaalsel juhul ei too episteemiline lünk kaasa ontoloogilist lünka (indeksikaalsus ei falsifitseeri materialismi), ja seega ei peaks ontoloogilist lünka kaasa tooma ka fenomenaalne lünk.

​**[lk.144]**​ On ilmne, et reduktiivne strateegia ei tööta. Nagu nägime [2. peatükis](#PARA.60-1), ühendab indeksikaale teadvuselamuse faktidega loogilise superventsuse puudumine füüsikaliste faktide suhtes, kuid indeksikaalide puhul saab olukorra lahendada, lisades üheainsa väikese indeksikaalse fakti — kõnealuse agendi asukoha. Ent isegi kui me annaksime Maryle täielikud teadmised tema indeksikaalsetest suhetest kõigi objektidega füüsikalises maailmas, ei täineneks tema teadmine punaste elamuste kohta vähimalgi määral. Fenomenaalse teadmise puududes puudub tal palju enam kui sellel, kellel puuduks ainult indeksikaalne teadmine.

Analoogiaid rõhutava strateegia juht on huvitavam. Sellele saaks vastu väita, et ontoloogiline lünk on olemas isegi indeksikaalsel juhul (vt. näiteks Nagel 1983). Kuid leidub ka sirgjoonelisem vastus. Nimelt paneme tähele, et indeksikaalsel juhul on [1. jaos](#PARA.81-2) esitatud mõttekäiguga analoogiline argument juba ette võimatu: ei saa olla mõeldavat keskpunktita maailma, kus füüsikalised faktid oleksid identsed meie maailma omadega, kuid indeksikaalsed faktid oleksid erinevad. Keskpunktita maailmades ei ole indeksikaalsetele faktidele üldse kohta. Vastav mõeldav keskpunktiga maailm oleks muidugi olemas, kuid ontoloogilisele küsimusele on relevantsed just keskpunktita maailmad. (Vastasel juhul oleks ontoloogiline lünk olemas ka indeksikaalsel juhul, nii et oponendil ei ole argumenteerimist kusagilt alustadagi.[[177]](#footnote-178)) Nii et ainult sellel juhul on episteemilist lünka võimalik olematuks seletada, märkides, et episteemilisi seoseid määratlevad keskpunktiga primaarsed intensioonid, ontoloogilisi seoseid aga keskpunktita intensioonidele vastavad omadused. See kajastubki ainsas lüngas, mille me leidsime ​2. jaos kirjeldatud argumendis ja analoogilises argumendis, millest oli juttu käesolevas jaos​: asjaolus, et primaarsed intensioonid määratlevad ainult keskpunktiga seostuvaid omadusi. Sellest lüngast pääseb ontoloogiliste kadudeta läbi üksainus taandumatu indeksikaalne teadmine (teadmine keskpunktiga maailma keskpunkti asukohast), kuid ei miski muu. Niipea kui keskpunkti asukoht on kindlaks määratud, on ka lünk täidetud. Fenomenaalsete faktide juhul jääb lünk aga alles ka siis, kui keskpunkti asukoht on kindlaks tehtud, ja seega see argument teadvuselamuse puhul ei tööta[[178]](#footnote-179).[[179]](#footnote-180) [[180]](#footnote-181) [[181]](#footnote-182)

Materialist, kes tahab endale kindlaks jääda, peab eitama, et Mary maailma kohta üldse midagi uut avastab. Materialism eeldab loogilist superventsust, mis nõuab, et Mary ei saaks esmakordsest punasest elamusest ühtki uut faktilist teadmist. Nii väidavadki Lewis (1990) ja Nemirow (1990) viiendat strateegiat järgides, et parimal juhul omandab Mary vaid uue oskuse — näiteks oskuse kujutleda punaste esemete välimust ja neid nähes ära tunda kui punaseid. Ent see on ainult “kuidas”-teadmine, mitte “et”-teadmine. Esmakordsest punasest elamusest ei saa ta maailma kohta teada ühtki fakti.[[182]](#footnote-183)

Erinevalt eelmistest variantidest ei ole selles strateegias seesmisi vigu. Tema peamiseks veaks on äärmine ebatõepärasus. Kahtlemata omandab Mary esmakordsest punasest elamusest mingeid oskusi, samuti nagu siis, kui ta õpib jalgrattaga sõitma. Kuid ei saa vist eitada, et ta ​**[lk.145]**​ saab teada veel midagi: ta omandab fakte elamuse loomuse kohta. Tema seniste teadmiste järgi oleks punastest asjadest saadav elamus olla niisugune või naasugune või üldse mitte mingisugune. Nüüd ta aga teab, et see on just niisugune. Ta on tugevasti ahendanud episteemiliste võimaluste ruumi. Sellist uut teadmist ei teki kõikteadjal mehhaanikul jalgrattaga sõitma õppides (kui ehk mitte arvestada teadmisi jalgrattasõidu fenomenoloogiast). Seega ei taba kõnealune vastus seda, mis toimub, kui Mary saab teada, mis tunne on näha punast.

Seda, et Mary saab oma avastuse läbi mingi faktilise teadmise, võib näha ka kaudsemal teel. Näiteks Loar (1990) juhib tähelepanu sellele, et seesuguseid teadmisi saab võtta sisse tingimuslausesse: “Kui punaseid asju nähes on “nii” ja siniseid asju nähes “nii”, siis on lillade asjade nägemisel tõenäoliselt “nii” ”; “Kui koertel on punase värvuse elamus niisugune, siis järelikult …”, jne. Veel üks näide: nagu osutab Lycan (1995)[[183]](#footnote-184), võib see, mida me ette kujutame, osutuda õigeks või ekslikuks; seega võib Mary, nähes mingit värvust, kujutleda, mis tunne oleks näha mõnd teist värvust, ja tema kujutlus võib osutuda õigeks või ekslikuks. See tähendab, et teadmine, mis tunne on midagi näha, on mingi tõe teadmine maailma kohta, ja analüüs, mis ütleb, et Mary omandab vaid uue oskuse​, ei tööta.

Dennetti (1991) seisukoht on sarnane, kuid äärmuslikum: tema väidab, et Mary ei saagi mingit uut teadmist. Ta ütleb, et selleks, et punast objekti nähes ära tunda, et see on punane, võiks Mary kasutada oma neurofüsioloogilisi teadmisi, uurides objekti mõju oma reaktsioonidele, mis arvatavasti erineksid siniste objektide mõjust. (Kui eksperimentaatorid püüavad teda petta, näidates talle sinist õuna, ei tarvitseks ta alt minna.) Võib-olla ongi tal õigus, kuid sellest järeldub ainult niipalju, et vastupidiselt sellele, mida ütlevad Lewis ja Nemirow, olid Maryl mingid äratundmisoskused juba enne oma esmakordset punast elamust. Sellest ei tulene kuidagi, nagu oleks tal olnud see meie jaoks kriitiline teadmine: teadmine, mis tunne oleks näha punast. Niisuguse järelduse saaks teha ainult juhul, kui me oleksime juba tunnistanud õigeks analüüsi, mille järgi teadmine, “mis tunne on”, saadakse tänu mingile oskusele; ent kui me oleksime selle analüüsiga nõustunud, siis oleks materialismivastane argument juba kummutatud. Seega on Dennetti argument siinkohal lihtsalt puru silmaajamine.

Lõpuks jääb materialistile ikkagi üksainus strateegia — eitada, nagu omandaks Mary maailma kohta uusi teadmisi. Ja ainus viis, kuidas ta seda veenvalt teha saab, on analüüsida teadmist, “mis tunne on”, oskuse mõiste kaudu. See on ainus loogiliselt järjekindel positsioon, mida ei ole võimalik tehniliste vastuväidetega kummutada, samamoodi nagu materialisti viimaseks ja kõige koherentsemaks võimaluseks loogilise superventsuse argumendile vastu panna on analüütiline funktsionalism. Kuid väide, et Mary ei omanda maailma kohta mingeid uusi teadmisi, on nii väheusutav, et juba see asjaolu üksi kuulutab materialismile hukku.[[184]](#footnote-185)

Nägime, et modaalne argument (ehk loogilise võimalikkuse argument) ja teadmise argument on ühe ja sama mündi kaks külge. Ma arvan, et põhimõtteliselt töötaks kumbki ka eraldi, kuid praktikas toimivad nad kõige paremini üheskoos.[[185]](#footnote-186) Kui võtta teadmise argument omaette, siis: ehkki enamik materialiste leiab ​**[lk.146]**​ olevat raske eitada, et Mary saab maailma kohta teadmisi, ei ole nad sageli nõus sellega, et sellest järeldub materialismi väärus. Kui võtta modaalne argument omaette, siis ehkki enamikul materialistidest on raske eitada argumenti, mis tuletab sombide või invertspektrite mõeldavusest materialismi vääruse, eitavad nad sageli argumendi eeldust. Ent neid kaht argumenti koos vaadeldes selgub, et modaalne argument toetab teadmiste argumenti just seal, kus tarvis, ja ümberpöördult. Kõige tõhusam viis nende kahe argumendi kombineeriseks käiks vististi nii: kõigepealt tõestame teadmise argumendi abil veenvalt ära loogilise mittesuperventsuse, ja seejärel tuletame modaalargumendi abil mittesuperventsusest veenvalt materialismi vääruse.

#### Kripke argument

Kripke ründas oma argumendiga Place’i (1957) ja Smarti (1959) identsusteesi, kuid sama mõttekäiku saab rakendada ka laiemalt, argumendina materialismi vastu kõigis oma vormides. Ma võtan selle argumendi tugevad ja nõrgad küljed nüüd põhjalikuma vaatluse alla ja püüan leida tuge järeldusele, et tugevad küljed on just need, mis vastavad loogilise superventsuse argumendile.

Mõttekäik on umbkaudu järgmine. Identsustees​[[186]](#footnote-187) ütleb, et mingid mentaalsed seisundid, näiteks valu, on identsed mingite aju seisunditega, näiteks ​C-kiudude aktivatsiooniga​[[187]](#footnote-188), olgugi et väljendid “valu” ja “C-kiudude aktivatsioon” tähendavad eri asju. Esialgselt on eeldatud, et see identsus ei ole paratamatu, vaid kontingentne​[[188]](#footnote-189), nii nagu vee ja H₂O identsus on kontingentne. Kripke aga väidab, et identsus on alati paratamatu: kui ​*X* on ​*Y*​, siis ​*X* on ​*Y* paratamatult, senikaua kui ​*X* ja ​*Y* on *rigiidsed designaatorid*​[[189]](#footnote-190)​, s.t. senikaua kui nad osutavad kõigis maailmades ühele ja samale indiviidile või liigile​[[190]](#footnote-191). Ta väidab, et vesi on paratamatult H₂O või, teisiti öeldes, vesi on H₂O kõigis võimalikes maailmades. Võib tunduda, et see identsus on kontingentne — see tähendab, võib tunduda, et on olemas võimalik maailm, kus vesi ei ole H₂O, vaid XYZ —, kuid see on illusioon. Tegelikult ei sisalda see kujuteldav võimalik maailm üldse vett, vaid ainult mingit vesijat ollust — midagi, mis näeb välja ja käitub nagu vesi, kuid koosneb XYZ-st. Kinnitada, et see vesijas ollus on vesi, tähendab kirjeldada teda valesti.

Analoogiliselt väidab Kripke, et kui valu on identne C-kiudude aktivatsiooniga, siis peab see identsus olema paratamatu. Samas aga ei tundu nende identsus paratamatuna. Pealtnäha tundub, nagu saaks kujutada ette võimalikku maailma, kus valuga ei kaasne üldse mingeid ajuseisundeid (valu on kehatu), ja ka maailma, kus C-kiud aktiveeruvad ilma, et sellega kaasneks valu (näiteks sombil). Kusjuures, väidab Kripke, neid võimalikkusi ei saa olematuks seletada kui pelgalt näilisi võimalikkusi, sarnaselt sellega, kuidas me äsja seletasime olematuks veega, kuid ilma H₂O-ta maailma võimalikkuse. Sest see tähendaks, et me kirjeldasime “kehatu valu” maailma valesti kui maailma, kus on olemas valu, kuigi tegelikult on vaid “valujas ollus” (s.t. miski, mis tundub ​**[lk.147]**​ nagu valu), ja ütlesime sombi kohta valesti, et tal puudub valu, samas kui tegelikult puudub tal vaid valujas ollus. Kui asja niiviisi seletada, siis tuleb välja, et reaalne valu, see tähendab C-kiudude aktiveerumine, oleks sombil olemas; ta lihtsalt ei tunne seda nagu reaalset valu.​[[191]](#footnote-192)

Kripke järgi ei saa see aga nii olla: valu ei olegi mitte midagi muud kui see, mis on aistitav valuna. Valu ja valujas ollus ei ole kaks ise asja, nagu seda on vesi ja vesijas ollus. Miski võib tunduda nagu vesi, olemata vesi, kuid mitte miski ei saa tunduda nagu valu, olematagi valu. Valutunne on valu olemuslik omadus. Seega ei saa, erinevalt varasematest juhtudest, heita kõrvale võimalust, et valu võib olla olemas ilma ajuseisunditeta (ja vastupidi). Nood võimalikud maailmad on tõepoolest võimalikud, ja mentaalsed seisundid ei ole paratamatult identsed aju seisunditega. Ja seega ei saa nad aju seisunditega üldse identsed olla.

Kripke toob seda argumenti mängu kahel viisil: kõigepealt objekti​[[192]](#footnote-193)identsusteooriate vastu, ja siis tüübiidentsusteooriate vastu. Objektiidentsusteooriate järgi on valu üksikjuhud (näiteks valu, mida ma praegu tunnen) identsed konkreetsete ajuseisunditega (näiteks C-kiudude aktivatsiooniga minu peas praegusel hetkel). Kripke väidab äsjakirjeldatud vaimus, et valu konkreetne üksikjuht võib leida aset, ilma et sellega kaasneks mingi konkreetne ajuseisund, ja ka ümberpöördult, ning seega ei saa need kaks olla identsed. Tüübiidentsusteooriad kinnitavad, et psüühilised seisundid ja aju seisundid on tüüpidena identsed: näiteks valu võib olla tüübina identne C-kiudude aktivatsiooniga. Kripke leiab, et sellise väite kummutab otseselt asjaolu, et vaimuseisundi tüüp võib realiseeruda, ilma et realiseeruks ajuseisundi tüüp, ja ümberpöördult. Põhimõtteliselt saab siin kokku lugeda neli eraldi argumenti, jaotades need kahte liiki vastavalt teooriale, mida me parajasti ründame (objekti- või tüübiidentsusteooria), ja vastavalt fenomenile, mille võimalikkusest me kõneleme (kehatud aistingud või sombid).

Kripke argumendil on ilmseid erinevusi minu argumendist. Esiteks on Kripke argumendi aluseks ainuüksi identsuse mõiste, mina olen aga toetunud superventsuse mõistele. Teiseks kuulub Kripke argument lahutamatult kokku tema teoreetilise aparaadiga, kus oluline koht on rigiidsetel designaatoritel ja aposterioorsel paratamatusel, minu argumendis aga mängib see aparatuur vaid abistavat rolli mõnede vastuväidete kummutamisel. Kolmandaks toetub Kripke argument, nagu seda tavaliselt mõistetakse, mõningate seisundite osas teatavale essentsialismile​[[193]](#footnote-194), minu argument aga ühelegi niisugusele õpetusele ei viita. Neljandaks ei apelleeri minu argument, erinevalt Kripke omast, kusagil kehatute aistingute võimalikkusele. Siiski on ka ilmseid sarnasusi. Mõlemad argumendid on modaalsed — otsustavas rollis on paratamatuse ja võimalikkuse mõiste. Ja mõlemad apelleerivad sellele, et füüsiliste seisundite lahusus nendega kaasnevatest fenomenaalsetest seisunditest on loogiliselt võimalik.

Vaatame nüüd, missugused on Kripke argumentide tugevad ja nõrgad küljed. Alustame argumentidest objektiidentsuse vastu. Enamik autoreid peab neid väheveenvateks. Põhjus on suuresti selles, et need argumendid toetuvad intuitiivsele arusaamale selle kohta, mida tuleks ​**[lk.148]**​ kõigis võimalikes maailmades lugeda üheks ja samaks objektiks, ja kõik teavad, kui ebausaldusväärne on neis asjus intuitsioon. Kripke väide, et kellelgi võib olla just nii- ja niisugune valuseisund ilma just nii- ja niisuguse ajuseisundita, toetub väitele, et sellele valuseisundile on olemuslikult omane just nii- ja niisugune kaasnev aisting. Kuid taolisi väiteid indiviidide olemuslike omaduste kohta on raske põhjendada. Objektiidentsusteoreetik võib vastuseks väita, et sama hästi võib selle seisundi olemuslikuks omaduseks olla C-kiudude aktivatsioon. Muidugi ei tundu C-kiudude aktivatsioon olevat olemuslikult omane valule kui tüübile, aga kes ütleb, et ta ei ole omane just sellele konkreetsele valujuhule, eriti veel kui see juht on identne mingi ajuseisundiga? Kui see on nii, siis ei saaks seda konkreetset valu ilma selle konkreetse ajuseisundita olemas ollagi. (Analoogiliselt on arutlenud Feldman (1974), kes väidab, et valusus ei tarvitse mingile kindlale valule olemuslikult omane olla, ja McGinn (1977), kes sisuliselt ütleb, et ühele kindlale valule võivad olla olemuslikult omased nii valusus kui ka C-kiudude aktivatsioon.) Kui nii, siis ei kujutle see, kes kujutleb minu valu kehatut versiooni, mitte täpselt seda kindlalt valu, vaid numeeriliselt erinevat valu. Sama võib öelda ka kujutluse kohta minu C-kiudude aktiveerumisest, millega ei kaasne valu. Seega ei ole objekti­identsuse vastased argumendid veenvad, kuigi tüübiidentsuse vastased argumendid võivad jõusse jääda.

Edasi, argumendi seda kuju, mis kõneleb kehatust valust, ei saa pidada veenvaks vastuväiteks materialismile. Võib-olla saab tema abil kummutada tüübiidentsusteesi niisugusel kujul, nagu seda esitavad Place ja Smart, kuid materialismiks pole see tees tingimata nõutav.[[194]](#footnote-195) [[195]](#footnote-196) Nagu märgib Boyd (1980), ei tarvitse materialist tingimata arvata, et vaimu seisundid on kõigis võimalikes maailmades füüsilised seisundid: materialismiga ei ole vastuolus võimalus, et on maailmu, kus vaimu seisundeid konstitueerib mittefüüsiline aines, peaasi, et meie maailmas oleks nende alus füüsiline. Kehatute aistingute võimalikkusest tuleneb vaid dualismi võimalikkus, mitte tema tõesus.[[196]](#footnote-197) Selle mõtte illustreerimiseks võime tähendada, et pole kuigi palju neid, kes väidaksid, et mittefüüsilise elu võimalikkusest tuleneb bioloogiline dualism. Võib-olla sooviski Kripke rünnata ainult identsusteesi, kuid materialismi oma üldisemal kujul ta sellega ei kummuta.

Üle jääb argument, mis kõneleb füüsikaliste seisundite realiseerumise võimalikkusest ilma vastavate fenomenaalsete seisunditeta — sisuliselt sombide võimalikkusele toetuv argument. Kummalisel kombel on just see osa Kripke mõttekäigust pälvinud kriitikutelt kõige vähem tähelepanu — peamiselt on kommenteerijad arutlenud kehatute aistingute võimalikkuse üle. Jällegi ei tarvitse argument, mille annavad tugeva tüübiidentsusteesi vastu sombid, ka siin relevantne olla — kuna materialism säärast teesi ei nõua —, kuid tema varjust piilub välja üldisem argument. Kripke väidab (lk. 153–54), et valule vastavate füüsiliste seisundite realiseerumise võimalikkus, ilma et sellega kaasneks valu ise, tähendab, et isegi pärast seda, kui Jumal lõi kõik valuga kaasnevad füüsilised protsessid — näiteks aktiveerunud C-kiududega aju —, pidi ta tegema veel midagi, et see aktivatsioon oleks tajutav valuna. Sellest piisab tõestamaks, et materialism on väär.[[197]](#footnote-198) [[198]](#footnote-199)

​**[lk.149]**​ Kripke argumendi see kuju, mis kõneleb füüsikalistest seisunditest ilma kaasuvate fenomenaalsete seisunditeta, on otseses vastavuses minu materialismivastase argumendiga. Vastavus laieneb isegi edasistele manöövritele. Vastuväitele, et säärane olukord on üksnes mõeldav, kuid mitte päriselt võimalik, vastab Kripke, et seda mõeldavat olukorda ei saa olematuks seletada kui olukorda, kus puudub valutunne, aga ei puudu valu ise, kuivõrd olla valu tähendab kõigis võimalikes maailmades midagi, mis on tuntav kui valu. (Teisiti öeldes, sõnal “valu” ühtib sekundaarne intensioon primaarsega). Sellele võiksime lisada (koos Jacksoniga (1980)), et isegi kui selle ekvivalentsusega mitte nõustuda, töötab materialismivastane argument ikkagi, kui rakendada seda valu asemel valutundele. (Teisiti öeldes, isegi kui intensioonid oleksid erinevad, töötaks argument primaarse intensiooni alusel ikkagi.) Need vastused on isomorfsed vastustega, mis ma andsin sarnastele vastuväidetele siinses peatükis ülal[[199]](#footnote-200).

(Paneme tähele, et oma teesiga, mis ütleb, et pealtnäha mõeldavad, kuid võimatud olukorrad tuleb seletada olematuks kui episteemiliselt võimalikud olukorrad, mida on lihtsalt valesti kirjeldatud, toetab Kripke tegelikult aposterioorse paratamatuse “nõrka” tõlgitsust: mõeldavate maailmade ruum ühtib võimalike maailmade ruumiga, kuid nende korrektsele kirjeldusele seavad kitsendusi aposterioorsed tegurid.[[200]](#footnote-201) Mõistmaks, et see nii on, paneme tähele, et “tugeva” metafüüsilise paratamatuse pooldaja, kes peab võimalike maailmade ruumi mõeldavate maailmade ruumi pärisosahulgaks, seda teesi ei pooldaks. Sellise vaate kohaselt on episteemiliselt võimalikku olukorda võimalik õigesti kirjeldada, kuid olukord ise võib ikkagi olla (lihtviisil) metafüüsiliselt võimatu. See, et Kripke otsustab seevastu pidada neid kirjeldusi ekslikeks, osutab, et ta toetab kaudselt kahemõõtmelist mudelit: tõepoolest, kõiki tema näiteid ekslikest kirjeldustest võib vaadelda kui juhte, kus maailma kirjeldatakse primaarsete, mitte sekundaarsete intensioonide alusel.)

See mõttekäik füüsikalistest seisunditest ilma fenomenaalsete seisunditeta on minu arvates Kripke argumendi kõige veenvam osa. Kogu muu diskussiooni seas identsusteesi, kehatute aistingute ja muu üle jäetakse see sageli tähele panemata; ka Kripke ise ei sea seda aspekti oma analüüsis kesksele kohale. Minule aga tundub, et kogu Kripke argumentatsioni koormat kannab viimselt just see osa tema mõttekäigust.

Kokkuvõttes tundub mulle, et niivõrd kui Kripke argument materialismi vastu märki tabab, on 1) kehatute aistingute võimalikkus materialismivastase argumendina väheveenev, kuid pole praegusel juhul ka oluline; 2) identsuse mõiste abil sõnastatud argumendid on samuti ebaveenvad, kuid ebaolulised; 3) essentsialistlik metafüüsika on oluline ainult kuivõrd valu tajumine on olemuslikult omane valule kui tüübile — kuid see on vaid fakt termini “valu” tähenduse kohta; ja 4) Kripke rigiidsete designaatorite jms. aparaat on küll vajalik mõnede vastuväidete kummutamiseks, kuid ei mängi keskset rolli.[[201]](#footnote-202) Kuid tema argumendil on soliidne tuum, milleks on sisuliselt loogilise mittesuperventsuse argument.

### ​**[lk.150]**​ 4. Kas see on epifenomenalism?\*

Üks selle positsiooni raskusi, mille poolt ma olen siin sõna võtnud, seisneb selles, et kui teadvus on vaid looduslikult superventne füüsika suhtes, siis peaks tal vististi puuduma põhjuslik toimivus. Füüsikaline maailm on põhjuslikult enam-vähem kinnine, selles mõttes et igale füüsikalisele sündmusele tundub leiduvat füüsikaline seletus (kui mitte arvestada väikese kvantmääramatusega). See tähendab, et mittefüüsikalisele teadvusele ei jää iseseisvaks põhjuslikuks toimimiseks ruumi, nii et ta on justkui füüsikalise põhjuslikkuse mootori pelk ripats, epifenomen, millest füüsikalises maailmas midagi ei sõltu. Ta on küll olemas, kuid füüsikalise maailma poolest võiks teda sama hästi ka mitte olla. Säärast vaadet pooldas Huxley (1874), kuid paljud peavad seda intuitsioonivastaseks ja peletavaks. Sellest järeldusest on piisanud, et panna mõned autorid (näiteks Kirk (1979) ja Seager (1991)) kahtlema oma materialismivastastest argumentidest tulenevates järeldustes ja kaaluma võimalust, et teadvus on füüsika suhtes võib-olla ikkagi loogiliselt superventne.

Sedasama argumenti on teisiti, kuid mõneti sarnaselt formaliseerinud ka Kirk (1979), Horgan (1987) ja Seager (1991). Kui me eeldame, et füüsiline maailm on põhjuslikult kinnine ja et teadvus põhjustab mingeid füüsikalisi sündmusi, siis tuleneb mõningatest loomulikest eeldustest põhjuslikkuse kohta, et teadvus peab olema füüsika suhtes loogiliselt (või metafüüsiliselt) superventne.[[202]](#footnote-203) Kui nii, siis järeldub teadvuse pelgalt looduslikust superventsusest  — kui füüsiline maailm on põhjuslikult kinnine —, et teadvus on epifenomenaalne nähtus. Selle mõttekäigu üldine kuju on selge: kui on võimalik fenomenaalsed nähtused meie maailmast lahutada, nii et ülejääv maailm Z jääb põhjuslikult kinniseks, siis on kõigel, mis toimub maailmas Z, põhjuslik seletus, mis on sõltumatu kõigest fenomenaalsest, sest selles maailmas ei ole midagi fenomenaalset. Ent kuna kõik, mis toimub maailmas Z, toimub ka meie maailmas, siis kehtib põhjuslik seletus, mis kehtib maailmas Z, samamoodi ka siin. Seega on fenomenaalsus põhjuslikult irrelevantne. Isegi kui teadvuselamusi ei oleks olemas, oleksid käitumise põhjuslikud mehhanismid ikka täpselt samasugused.

Sellele vastamiseks rakendan ma kaheharulist strateegiat. Esiteks ei ole sugugi ilmne, et pelk looduslik superventsus peaks eeldama kõige tugevamal kujul epifenomenalismi: kuigi ei saa salata, et temaga seostuv pilt sarnaneb tõepoolest mõnevõrra epifenomenalismiga, on põhjuslikkuse olemus ise üsna salapärane ning pole võimatu, et kui põhjuslikkust hakatakse paremini mõistma, suudame me võib-olla mõista ka, kuidas teadvuselamus võib mingil raskesti tabataval moel siiski põhjuslikult relevantne olla. (Näiteks võib selguda, et ülaltoodud argumentide tausteeldused on väärad). Järgnevalt kirjeldan mõningaid võimalusi, kuidas selline analüüs võiks kulgeda. [Teises harus](#ORG.4.4.3) võtan vaatluse alla põhjused, miks epifenomenalismi võidakse vastuvõetamatuks pidada, ning analüüsin, kui suur võib olla nende jõud argumentidena. Kui neist intuitiivsetest aimustest ei õnnestu välja arendada veenvaid argumente, siis võib selguda, et intuitsioonivastasus on kõnealusest positsioonist tuleneva ​**[lk.151]**​ epifenomenalismi ainus viga ning et lõpuks võib epifenomenalismiga mingis annuses ka leppida.

#### Strateegiaid epifenomenalismi vältimiseks

On terve hulk võimalusi tunnistada loogilise mittesuperventsuse teesi ja samas vältida epifenomenalismi. Kõige ilmsem neist on füüsikalise maailma põhjusliku kinnisuse eitamine ning tugeva interaktsionistliku dualismi[[203]](#footnote-204) omaksvõtt niisuguses variandis, mille järgi vaimsed protsessid täidavad põhjuslikkuslünki füüsikalistes protsessides. Minu arvates tuleks seda strateegiat vältida — põhjustel, millest peagi juttu tuleb. Kuid olenevalt vaadetest metafüüsikale ja eriti põhjuslikkusele leidub ka vähem ilmseid võimalusi. Alljärgnevas võtan vaatluse alla neli sellist võimalust.

##### 1. Regulaarsusele toetuv põhjuslikkus.

Esimene võimalus on võtta omaks Hume’i radikaalne​[[204]](#footnote-205) arusaam põhjuslikkusest, mille järgi selleks, et sündmust ​*A* saaks pidada sündmuse ​*B* põhjuseks, piisab ühetaolisest regulaarsusest ​*A*​-tüüpi sündmuste ja ​*B*​-tüüpi sündmuste vahel. Selline vaade lubaks kõnelda fenomenaalse sfääri põhjuslikust rollist: juba ainuüksi asjaolu, et valutundele järgnevad üldjuhul hoidumisreaktsioonid, tähendaks, et valu on hoidumisreaktsioonide põhjus.

Hume’i arusaamast erinev, kuid sellele lähedane võimalus on pidada põhjuslikku seost samaväärseks suvalise *noomilise* (s.t. seaduspärase) seosega, isegi kui noomiline regulaarsus on midagi enamat kui ühetaoline regulaarsus​.[[205]](#footnote-206) Loodusliku superventsuse idee on täielikult ühitatav noomilise seose olemasoluga elamuse ja käitumise vahel. (Näiteks võib seaduspärane seos olla elamuse ja selle aluseks oleva ajuseisundi vahel, samuti selle ajuseisundi ja käitumise vahel). Võiks öelda, et põhjuslikkuseks sellest piisab. Seda arusaama võiks toetada ka tähelepaneku abil, et kontrafaktiline väide “Käitumine oleks olnud samasugune ka ilma elamuseta” on kõige loomulikuma tõlgenduse kohaselt väär: kui elamust poleks olnud, oleks aju seisund olnud teistsugune, ja teistsugune oleks olnud ka käitumine. (Kontrafaktuaali õigsuse hindamiseks vaatleme praegusel juhul looduslikult võimalikke, mitte loogiliselt võimalikke maailmu.)

Mina leian, et mõlemad seisukohad on ebaveenvad. Hume’i põhjuslikkusekäsitluse vastu esitasin ma argumente ​2. peatükis​, ja ka viimati kirjeldatud mittehume’liku vaate korral pole tõenäoline, et põhjuslikkuseks piisab suvalisest noomilisest seosest — võtame näiteks juuste värvuse korrelatsiooni identsetel kaksikutel. Siiski annavad selletaolised kaalutlused meile vähemalt mingi ettekujutuse sellest, miks meile tundub, nagu mängiks teadvus põhjuslikku rolli. Teadvuselamuste ja nendele järgnevate füüsiliste sündmuste vahel on kõikvõimalikke süstemaatilisi reeglipäraseid vastavusi, mis kõik kallutavad meid järeldama põhjusliku seose olemasolu. Selliste reeglipäradega silmitsi seistes võiks oodata, et inimesed toetuvad neist põhjuslikke suhteid tuletades argumentidele, millel on laias laastus hume’lik iseloom. Selle asjaoluga võib seega seletada olematuks mõned meie intuitiivsetest arusaamadest, mille järgi ​**[lk.152]**​ teadvusel on põhjuslik toime, ning anda sellega tuge minu strateegia teisele harule.

##### 2. Liiane põhjuslik determinatsioon

Võiksime ehk väita, et hilisema füüsilise seisundi põhjustajaks võib pidada nii füüsilist kui ka fenomenaalset seisundit, olgugi need täiesti erinevad. Kui füüsiline seisund ​*P*​₁ kaasneb fenomenaalse seisundiga ​*Q*​₁, siis võib olla tõsi nii see, et hilisema füüsilise seisundi ​*P*​₂ põhjustas ​*P*​₁, kui ka, et selle põhjustas ​*Q*​₁. See on aga intuitsioonivastane: kuna sündmus ​*P*​₁ on sündmusele ​*P*​₂ juba piisavaks põhjuseks, siis peaks ​*Q*​₁ olema põhjuslikult liiane. Kuid pole ilmne, et ka sündmusel *Q*​₁ ei võiks ikkagi olla põhjuslikku suhet sündmusega ​*P*​₂[[206]](#footnote-207). See võimalus tundub eriti tõepärasena juhul, kui tunnistada põhjuslikkuse mittereduktiivset käsitlust (nagu seda soovitab Tooley 1987). Võib-olla on siin kaks teineteisest sõltumatut mitteredutseeritavat põhjuslikku seost: üks kahe füüsilise seisundi vahel ning teine fenomenaalse seisundi ja vastava füüsilise seisundi vahel.

Sellisele sündmuste põhjuslikule liigdeterminatsioonile vaadatakse sageli kahtlustava pilguga, kuid pole kerge veenvalt näidata, et selles midagi valesti oleks. Me ei mõista põhjuslikkuse olemust veel küllalt hästi , et liigdeterminatsiooni välistada. Ma ei kavatse seda mõtet siin ise edasi arendada, kuid ta väärib siiski tõsist tähelepanu.

##### 3. Põhjuslikkuse mittesuperventsus.

Kolmas strateegia toetub otseselt põhjuslikkuse enda olemusele. Me nägime [2. peatükis](#ORG.2.5.0.2), et on olemas kaks konkreetsete füüsikaliste faktide suhtes loogiliselt mittesuperventsete faktide klassi: teadvuse faktid ja põhjuslikkuse faktid. On loomulik oletada, et nende kahe mittesuperventsuse vahel on tihe seos ning et teadvuse ja põhjuslikkuse vahel on mingi sügav metafüüsiline side. Mõlemad on ju üsna salapärased, ja kaks salapärast nähtust oleks ehk seadsam üheks kokku pakkida. Näiteks võib teadvuselamus ise olla mingitliiki põhjuslik seos; võib-olla realiseerub temas mingil moel põhjuslik suhe, mis on Hume’i järgi teadmisele kättesaamatu[[207]](#footnote-208); või siis on siin mingi keerukam suhe. Mingi selletaoline seos osutab võib-olla, et elamusel on põhjuslikes protsessides mingi tavalist liiki põhjuslikkusest keerukam roll, mis aga võimaldab siiski vältida epifenomenalismi kõige tugevamat vormi.

Sellise teooria on välja töötanud Rosenberg (1996), kes väidab, et paljudel teadvuse probleemidel on täpsed paralleelid põhjuslikkuse probleemistikus. Tema sõnul lubavad need paralleelid oletada, et aktuaalses maailmas realiseerib põhjuslikkust või põhjuslikkuse mõningaid aspekte teadvuselamus: põhjuslikkus vajab midagi, mis tema paljusid omadusi realiseeriks, ja elamus on sellele kohale loomulikuks kandidaadiks. Kui nii, siis võib olla, et just elamuse olemasolu võimaldabki põhjuslikel suhetel eksisteerida, nii et elamus mängib põhjuslikkuses mingis peenes mõttes relevantset rolli.

Muidugi põrkub see äärmiselt spekulatiivne mõte mõningatele raskustele. Kõigepealt tundub, et ta viib välja mingitliiki panpsühhismini, s.t. positsioonini,​**[lk.153]**​ mille järgi teadvus on kõigel, kuid mida paljud autorid peavad intuitsioonivastaseks. Endistviisi jääb alles ka sombimaailma probleem: tundub, et kõiki neid põhjuslikke protsesse võib kujutleda kulgevana ka ilma elamusteta, ja seega võib endiselt tunduda, et elamus on epifenomenaalne. Sellel võiks vastata, et põhjuslikkust peab ju miski realiseerima; sombimaailmas realiseerib teda miski muu, kuid meie maailmas on elamus ikkagi relevantne, kuna siin realiseerib ta põhjuslikkust. Mulle pole ilmne, et põhjuslikkust peab tingimata realiseerima miski, millel on veel muid omadusi; kui ei pea, siis võib põhjuslikkuse fenomenaalne olemus ikkagi liigseks osutuda. Kuid tähendame veel kord, et põhjuslikkuse metafüüsika ei ole veel kaugeltki selge, ja see vaateviis väärib kindlasti põhjalikumat uurimist.

##### 4. Füüsilise maailma seesmine olemus

Kõige enam köidab mind strateegia, mis toetub tähelepanekule, et füüsikaline teooria iseloomustab oma alusentiteete ainult *relatsiooniliselt* — põhjuslike jm. suhete kaudu teiste entiteetidega. Näiteks fundamentaalseid osakesi iseloomustatakse peamiselt nende soodumuse kaudu suhelda teiste osakestega. Nimetatakse küll nende täpne mass ja laeng, ent viimselt ei tähenda näiteks mass midagi enamat kui soodumust liikuda mingi kindla suurusega jõu toimel mingi kindla kiirendusega, jne. Iga entiteeti iseloomustab tema suhe teiste entiteetidega, neid omakorda nende suhe veel mingite entiteetidega, ja nii lõpmatult edasi (võib-olla kui mitte arvestada entiteete, mida iseloomustab nende suhe vaatlejaga). Pilt, mis sellest kõigest kokku tuleb, kujutab füüsilist maailma hiiglasliku põhjuslikkusvoona, kuid ei ütle meile midagi selle põhjuslikkuse kandja​[[208]](#footnote-209) kohta. Prootonist kõneldes mõeldakse midagi, mis põhjustab mingit kindlat liiki vastasmõjusid, võib kombineeruda mingil kindlal viisil teiste entiteetidega jne.; kuid mis see põhjustaja ja kombineeruja ise on? Nagu tähendab Russell (1927), selles küsimuses füüsikaline teooria vaikib.[[209]](#footnote-210)

Võiksime lasta end ahvatledagi vaatest maailmale kui puhtale põhjuslikkusvoole, kus põhjuslikkusel polegi kandjat. Kuid see annaks meile füüsikalisest maailmast kummaliselt sisutühja pildi[[210]](#footnote-211) [[211]](#footnote-212), mille järgi maailm sisaldaks ainult põhjuslikke ja noomilisi suhteid tühjade, ilma ühegi omaenda omaduseta kandjate vahel. Intuitiivselt tundub mõistlikum oletada, et alusentiteetidel, mille juurde kogu see põhjuslikkus kuulub, on mingi oma seesmine loomus, mingid seesmised omadused, nii et maailmal on ka mingi substants. Kuid füüsika saab viidata neile omadustele parimal juhul vaid väliste seoste kaudu; ta ei ütle meile otseselt midagi selle kohta, mis need omadused olla võiksid. Meil on nende omaduste kohta mõned ebamäärased intuitsioonid, mille aluseks on elamused, mida me saame nende makroskoopilistest analoogidest — näiteks see, mida ütleb intuitsioon meile massi enda “massiivsuse” kohta —, kuid neid intuitiivseid tundeid on raske sõnadesse panna, ja õigupoolest polegi kindel, et neis üldse midagi on.

On olemas ainult üks seesmiste, mitterelatsiooniliste omaduste klass, mida me otseselt tunneme, nimelt fenomenaalsete omaduste klass. ​**[lk.154]**​ On loomulik oletada, et füüsikaliste entiteetide määratlemata olemuslikel omadustel on mingi seos või isegi kattuvus teadvuselamuse olemuslike omadustega, mis on meile tuttavad. Võib-olla on õigus Russellil, kui ta oletas, et vähemalt mõned füüsikalise maailma olemuslikest omadustest on ise mingid fenomenaalsete omaduste erikujud?[[212]](#footnote-213) See mõte kõlab esialgu pööraselt, järele mõeldes aga mitte enam nii väga. Ei ole meil ju füüsikaliste entiteetide seesmistest omadustest mingit ettekujutust. Nende olemus on lahtine, ja fenomenaalsed omadused ei tundu selle koha täitjaks halvemini sobivat kui miski muu.

Siin hakkab meid muidugi ähvardama panpsühhismi oht. Ma ise ei arva, et see oleks nii halb perspektiiv — on ju loomulik oletada, et kui fenomenaalsed omadused on fundamentaalsed, siis võib neid kohata mujalgi —, kuid see ei ole paratamatu järeldus. Teine võimalus on, et asjaomasteks omadusteks on protofenomenaalsed omadused. Sellisel juhul ei too niisuguse omaduse realiseerumine iseenesest veel kaasa teadvuselamust, kuid seda võib teha paljude sääraste omaduste realiseerumine üheskoos. On raske ette kujutada, kuidas see võiks tegelikult käia (me teame, et tavaliste füüsikaliste omadustega midagi niisugust toimuda ei saa), kuid säärased seesmised omadused jäävad meie arusaamadest üsna kaugele. Kuid seda võimalust ei saa *a priori* välistada.

Nii või teisiti osutab selline sügav seos võimalusele, et fenomenaalsusel on mingi põhjuslik roll. Kui füüsikalisel maailmal on seesmised omadused, siis kuulub füüsikaline põhjuslikkus viimselt selle juurde, mis neid omadusi realiseerib. Kui need on fenomenaalsed omadused, siis on olemas ka fenomenaalne põhjuslikkus; kui need on aga protofenomenaalsed omadused, siis omandavad fenomenaalsed omadused põhjusliku relevantsuse oma superventsuse läbi pärandusena nendelt, just samuti nagu piljardikuulid omandavad oma põhjusliku relevantsuse pärandusena molekulidelt. Mõlemal juhul võib inimestest agentide elamuste fenomenoloogia saada põhjusliku relevantsuse pärandiks füüsikaliste entiteetide seesmiste omaduste põhjuslikult rollilt.

Loomulikult ei toimiks säärane põhjuslik relevantsus sama sirgjooneliselt nagu tavalistel juhtudel. Näiteks kuigi me võime ka nüüd mõtteliselt kaotada kõik fenomenaalsed omadused, jättes põhjuslikkusvoo kuju endiseks, võib sellele​[[213]](#footnote-214) nüüd vastata, et niisugust stsenaariumi ette kujutades muudame me tegelikult füüsikaliste entiteetide seesmisi omadusi ja asendame nad millegi muuga. (See on muidugi raske, kuna meil puudub täielikult harjumus kujutleda füüsikaliste entiteetide seesmisi omadusi.) Seega oleme me kujutlenud lihtsalt maailma, kus põhjuslikkusel on mingi muu alus. Kui saaks olla olemas maailm, kus pole mitte midagi peale põhjuslikkusvoo, ei peaks see argument paika, kuid sellist maailma võib pidada loogiliselt võimatuks, kuna seal puuduks põhjuslikkusel kandja.

See positsioon on üsna sarnane ​[2. jaos](#ORG.4.2.1) kirjeldatud alternatiivse​[[214]](#footnote-215) seisukohaga, mis ütleb, et elektronidel on mingi tabamatu seesmine olemus, millele füüsikalised kirjeldused fikseerivad vaid viite. Ma arvan, et sealsamas toodud põhjustel ei tohiks neid olemuslikke omadusi samastada füüsikaliste omadustega, nagu seda on näiteks mass. On vist mõistlik öelda, et ka sombimaailmas on mass olemas, kuigi tema seesmine olemus on teistsugune. See tähendaks, et mass on väline omadus, mida võivad erinevates maailmades “realiseerida” erinevad seesmised omadused. Ent ​**[lk.155]**​ ükskõik, kumbapidi me selle semantilise otsuse ka langetaksime, jääb selle positsiooni puhul ikka alles omaduste duaalsus: ühelt poolt omadused, milledega füüsika otseselt tegeleb, ja teiselt poolt tabamatud seesmised omadused, mis konstitueerivad fenomenoloogia.

Mingis mõttes võib seda vaadet pidada pigem monismiks kui dualismiks. Kuid see ei ole materialistlik monism. Erinevalt füsikalismist peab see vaade mingeid fenomenaalseid või protofenomenaalseid omadusi fundamentaalseteks. Lõpuks jõuab ta välja seesmiste omaduste võrgustikuni, milledest vähemalt mõned on fenomenaalsed või protofenomenaalsed ja mida seovad omavahel mingid kindlad põhjuslikud/dünaamilised seadused. Neis omadustes “realiseeruvad” välised füüsikalised omadused, neid omavahel siduvad seadused aga realiseerivad füüsikalisi seadusi. Piirjuhul, kui kõik seesmised omadused on fenomenaalsed, on seda seisukohta kõige sobivam pidada mingitliiki idealismiks. See on aga hoopis teistsugune kui Berkeley idealism: siin ei ole maailm vaatleja teadvuse suhtes superventne, vaid maailma moodustab tohutu fenomenaalsete omaduste põhjuslike suhete süsteem, ja see süsteem ongi aluseks füüsikaseadustele, mida postuleerib teadus. Vähem äärmuslikku juhtumit, kus seesmised omadused on protofenomenaalsed või kus mõned neist ei ole ei fenomenaalsed ega protofenomenaalsed, on ehk kõige parem pidada Russelli neutraalse monismi teisendiks. Siin ei ole maailma alusomadused ei füüsikalised ega fenomenaalsed, vaid nii füüsikalised kui ka fenomenaalsed omadused alles konstrueeruvad neist. Nende seesmiste olemuste kogusummast konstrueeruvad fenomenaalsed omadused, nende välistest suhetest füüsilised.

Selle arusaama järgi on kõige fundamentaalsemad seadused need, mis ühendavad seesmisi alusomadusi. Tavalised füüsikaseadused kirjeldavad nende seaduste relatsioonilist kuju, abstraheerides neist välja olemuslikud omadused. Psühhofüüsika seadused võib ümber tõlgendada seadusteks, mis seovad olemuslikke omadusi (või neist konstrueeruvaid omadusi) nende relatsiooniliste profiilidega (või keerukate relatsiooniliste struktuuridega). Seega ei ole need seadused ontoloogilises mõttes füüsikaseaduste “ripatsid”, vaid nii ühed kui teised tulenevad seadustest, mis on tõeliselt fundamentaalsed. Kuid epistemoloogiline järjekord erineb ontoloogilisest: me leiame eest kõigepealt põhjuslike suhete süsteemi relatsioonilise struktuuri ning liigume alles sealt pikkamööda edasi olemuslike omaduste juurde, millele see struktuur toetub. Igapäevaste seletuste tarbeks on seepärast kõige kasulikum mõtestada seda vaateviisi endistviisi füüsikaseaduste võrgustiku kaudu ning siduda füüsikalised omadused fenomenaalsetega alles lisaprintsiipide abil.

.

Ehkki kõigi nende metafüüsiliste spekulatsioonidega tuleb olla parasjagu ettevaatlik, ilmneb neist siiski, et epifenomenalismi küsimus ei ole triviaalne. Põhjuslikkusega ja teadvuselamuse olemusega seoses kerkib hulk peeni probleeme, mida tuleb paremini mõista, enne kui kindlalt öelda, kas elamus on epifenomenaalne või mitte. Igatahes jätan ma nüüd metafüüsilised spekulatsioonid kõrvale ja tulen tagasi maisemale tasandile. (Võtan aga mõningad neist küsimustest uuesti vaatluse alla ​8. peatükis​).

​**[lk.156]**​ Nii või teisiti on looduslikul superventsusel epifenomenalistlikku hõngu. Võiksime öelda, et see vaade on epifenomenalistlik esimeses lähenduses: kui ta võimaldabki omistada elamusele mingit põhjuslikku relevantsust, siis pole selge, kuidas ta seda teeb. Ma arvan, et me tabame selle esimese lähenduse mõtet, kui ütleme, et see vaade teeb elamuse seletuslikult irrelevantseks. Käitumist saab seletada puht füüsikaliselt või informaatiliselt mõistetes, mis ei eelda fenomenoloogiat ega viita fenomenoloogiale. Kui elamusel on mingi raskesti tabatav osa põhjuslikkuses, siis on see niisugune, millest need seletused saavad abstraktsioonide abil mööda minna. Meid võib ärevaks teha isegi seletuslik irrelevantsus; sellest tuleb palju põhjalikumalt juttu [järgmises peatükis](#ORG.5.4).

Mõnel autoril on olnud kiusatus teha epifenomenalismi vältimiseks suur hüpe ja võtta omaks [käesoleva peatüki 2. jaos](#ORG.4.2.2) vaadeldud “tugeva metafüüsilise paratamatuse” positsioon. Kui elamus ei ole füüsika suhtes loogiliselt superventne, siis ei ole tema põhjusliku rolli säilitamiseks mõnede arvamuste järgi muud viisi kui lugeda ta lihtviisil identseks ühe või mitme füüsikalise omadusega, või siis nende suhtes metafüüsiliselt superventseks. Lisaks juba mainitud raskustele on ka sellel seisukohal ikka veel tõsiseid raskusi seletusliku irrelevantsusega. Juba sombide mõeldavus näitab, et selle vaate kohaselt on käitumist võimalik seletada mõistetes, mis ei viita elamusele ega eelda selle olemasolu. Seletuslikud suhted on kontseptuaalsed suhted, mistõttu tugev metafüüsiline paratamatus ei puutu siin asjasse. See vaade jätab käitumise ikkagi tugevas mõttes elamusest sõltumatuks ja peab seisma silmitsi enamikuga samadest raskustest, millele põrkub ka omadusdualism. Seepärast ei ole meil niisugusele positsioonile asudes kuigi palju võita.

#### Interaktsionistlik dualism?

Need, kes peavad dualismi argumente veenvateks, kuid usuvad, et fenomenaalne teadvus peab mängima olulist põhjuslikku rolli, võivad lasta end ahvatleda dualismi interaktsionistlikul versioonil, mille järgi elamus täidab põhjuslikkuslünki füüsikalistes protsessides. Sellele ahvatlusele järele andes ei jää raskusi aga vähemaks, vaid tuleb veelgi juurde. Esiteks tähendab see mitte sugugi paljulubava panuse tegemist füüsika tulevikule: vähemalt praegu tundub küll, et füüsikalisi sündmusi on alati võimalik seletada ainuüksi teiste füüsikaliste sündmuste kaudu. Teiseks nõuab see vaade olulise panuse tegemist ka kognitiivteaduste tulevikule, kuna ta eeldab, et tavalistest füüsikalistest/funktsionaalsetest mudelitest käitumise seletamiseks ei piisa. Kuid kõige olulisem puudus on, et see vaade ei pruugi  — põhjustel, mida ma [peagi](#PARA.97-5) arutama hakkan — epifenomenalismi raskustega paremini toime tulla kui vaade, mis eeldab põhjuslikku kinnisust.

Ainus interaktsionistliku dualismi vorm, mis tänapäevaste ettekujutuste valguses vähegi vastuvõetavana tundub, kasutab ära mõningaid kvantmehhaanika iseärasusi. Selleks on kaks võimalust. Esiteks on mõned autorid (näiteks Eccles 1986) apelleerinud kvantmääramatusele ja pakkunud välja oletuse ​**[lk.157]**​, et võib-olla ongi mittefüüsikaline teadvus see, mis täidab sellest määramatusest tulenevad põhjuslikkuslüngad, determineerides füüsikaliste suuruste tegelikud väärtused näiliselt “tõenäosusliku” jaotuse piires. Kuigi nende otsuste otsene mõju oleks üliväike, pole ehk võimatu, et mittelineaarne dünaamika võimendab need tibatillukesed fluktuatsioonid märgatavaks makroskoopiliseks mõjuks käitumisele.

See julge ja huvitav mõte toob aga kaasa terve hulga raskusi. Esiteks on ta vastuolus kvantmehhaanika postulaadiga, mille järgi need mikroskoopilised “otsused” on täiesti juhuslikud, ning ütleb põhimõtteliselt, et neis peaks olema mingi tuvastatav muster. See on kontrollitav hüpotees. Teiseks, selleks, et see teooria lubaks öelda, et teadvus mängib mingit huvitavat põhjuslikku rolli, peab käitumine, mis sünnib mingite kindlate mikroskoopiliste otsuste tagajärjel, olema mingis suhtes teistsugune kui käitumine, mida põhjustaks enamik teisi, puhtalt juhusliku protsessi läbi sündinud otsuseid. Eeldatavasti peaks see käitumine olema ratsionaalsem, kui ta oleks muidu olnud, ja tooma kaasa ütlusi — näiteks “ma näen praegu punast” —, mis ei oleks olnud võimalikud, kui protsessid oleksid olnud juhuslikud. Ka seda järeldust on põhimõtteliselt võimalik testida: selleks tuleks käivitada ajusimulatsioon, kus neid otsuseid determineeriksid päriselt juhuslikud protsessid. Muidugi ei tea me kindlalt, millise tulemuse see test annaks, kuid väita, et käitumine, mis sünniks juhuslike protsesside tagajärjel, oleks oluliselt halvem, tähendaks sõlmida kihlvedu väga väikese võiduvõimalusega.

Teine võimalus kvantmehhaanika seostamiseks põhjusliku kinnisuse küsimusega tuleneb asjaolust, et mõnes kvantformalismi tõlgenduses mängib teadvus ise esmatähtsat põhjuslikku rolli, olles tarvilik nn. *lainefunktsiooni kollapsiks*​. Kollaps toimub eeldatavasti iga mõõtmisakti käigus; ja ühe tõlgenduse järgi saab mõõtmist eristada kõigest, mis ei ole mõõtmine, ainuüksi teadvuse kohalolu põhjal. See ei ole kindlasti üldtunnustatud teooria (alustades juba sellest, et selle järgi ei ole teadvus ise füüsiline, see aga on teadagi vastuolus füüsikute enamiku seas valitsevate vaadetega), ja ma ei tunnista seda ka ise, kuid igal juhul tundub, et põhjuslik mõju, mida teadvus selle järgi avaldab, on hoopis teist liiki kui mõju, mis on vajalik, et teadvus mängiks rolli käitumise suunamisel.[[215]](#footnote-216) Pole selge, kuidas väliselt tajutud objektides toimuv kollaps võimaldab teadvusel mõjutada füüsilisi protsesse ajus; säärased teooriad ei ütle tavaliselt midagi selle kohta, mis toimub kollapsi ajal ajus. Ja isegi kui teadvusel õnnestub kuidagi kutsuda esile ajuseisundi kollaps, jäävad ikkagi kehtima kõik ülaltoodud märkused näiliselt juhuslike protsesside kohta ja nende seose kohta käitumisega.

Igal juhul takerdub interaktsionistlik dualism igas oma versioonis mõnesse kontseptuaalsesse raskusse, ja see viib mõttele, et see vaade ei suuda epifenomenalismi sugugi nii edukalt vältida, nagu tunduda võiks, või vähemalt, et ta ei ole üheski oma vormis parem kui vaade, mida olen kaitsnud mina. Kui nende seisukohtadega nõustudagi, tuleb ikkagi välja, et fenomenaalne maailm on mingis mõttes irrelevantne. Igast seletusest ​**[lk.158]**​ saab fenomenaalse komponendi alati lahutada, nii et järele jääb ainuüksi puhtalt põhjuslik komponent. Kujutlegem (koos Ecclesiga), et mittefüüsilises vaimus on “psühhonid”, mis ajus füüsikalisi protsesse siia-sinna lükkavad ja mis ongi elamuse kandjad. Me võime pikalt-laialt pajatada põhjuslikest suhetest psühhonite ja füüsikaliste protsesside vahel ning psühhonite omavahelisest põhjuslikust dünaamikast, kordagi mainimata asjaolu, et neil on fenomenaalsed omadused. Nagu füüsikaliste protsesside puhul, nii ka siin võib psühhonite fenomenaalsed omadused pildist mõtteliselt välja jätta, nii et olukorra põhjuslik dünaamika jääb senisega isomorfseks. Sellest järeldub, et kui psühhonid ongi elamuse asupaik, siis ei mängi see asjaolu põhjuslikes seletustes mingit olulist rolli, ja et isegi selle vaate järgi on elamus seletuslikult irrelevantne.

Sellele võidakse vastu väita, et psühhoneid (või ektoplasmat või mida veel) konstitueerivad ammendavalt nende fenomenaalsed omadused. Ent isegi niisugusel juhul on nende fenomenaalsed omadused käitumise seletamisel mingis mõttes irrelevantsed; põhjusliku dünaamika kirjeldamisel tulevad arvesse ainult nende relatsioonilised omadused. Kui keegi kinnitab jonnakalt, et psühhonitel on ikkagi ka muid seesmisi omadusi, mis on põhjuslikult relevantsed, satume olukorda, mis sarnaneb [ülal](#PARA.96-3)kirjeldatud olukorraga, kus oli juttu füüsikalistele entiteetidele seesmiselt omastest fenomenaalsetest omadustest. Mõlemal juhul on jutt mingit laadi põhjuslikust relevantsusest ja sellega kaasnevast seletuslikust irrelevantsusest. Seega ei võida me füüsikalise maailma põhjuslikust kinnisusest loobudes tõepoolest midagi erilist. Suurem põhjuslikult kinnine suhete süsteem jääb endiselt alles ning sellesse kuuluvate entiteetide fenomenaalne loomus jääb seletuslikus mõttes endiselt tarbetuks.

Võime isegi kujutleda, et kui interaktsionism on õige, siis on teadvuselamusest täiesti sõltumatutel põhjustel viimaks ikkagi hädatarvilik postuleerida käitumise seletamiseks, täheldatud põhjuslikkuslünkade täitmiseks ja olemasolevate andmete tõlgendamiseks psühhonite olemasolu. Niisugusel juhul saaksid psühhonitest teoreetilised entiteedid, millede staatus oleks samasugune nagu teoreetilistel entiteetidel füüsikas. Kuid kogu selles loos ei oleks midagi, mis viitaks elamusele või eeldaks elamust, ja seega oleks elamus seletuslikult samavõrd ülearune nagu tavalisel juhul; ka psühhonite olemasolu ei tee sombisid võimatuks. Asjaolu, et neil psühhonitel oleksid lisaks muudele omadustele ka fenomenaalsed omadused, ei kaitseks meid epifenomenalismi eest paremini ega halvemini kui analoogiline tähepanek, et füüsikalistel entiteetidel (võib-olla alusentiteetidel, võib-olla mingitel väga keerukatel objektidel) võib lisaks välistele omadustele olla ka fenomenaalseid omadusi. Seega ei aita füüsikalise maailma põhjusliku kinnisuse eitamine epifenomenalismi vältimisele kuigivõrd kaasa.[[216]](#footnote-217).

#### Epifenomenalismi raskused

Igal vaatel, mis võtab teadvust tõsiselt, tuleb, vähemasti mingis nõrgas vormis, seista silmitsi epifenomenalismiga. Juba asjaolu, et elamust saab koherentsust minetamata lahutada suvalisest põhjuslikust kirjeldusest, tähendab, et elamus on ​**[lk.159]**​ käitumise seletamisel liigne, olenemata sellest, kas ta on mingil tabamatul moel põhjuslikult relevantne või mitte. Võimalik, et elamus osutub põhjuslikult irrelevantseks ka mingis tugevamas mõttes; see küsimus jääb esialgu lahtiseks. Seepärast peame proovima meie strateegia teist haru ja selgitama välja, missuguseid raskusi tuleneb elamuse põhjuslikust või seletuslikust irrelevantsusest ja kas need raskused on kokkuvõttes ületamatud. Ma arutlen nende küsimuste üle palju põhjalikumalt [5. peatükis](#ORG.5.4), kuid siinkohal võib teha olukorrast väikese ülevaate.

Kõige sagedamini heidetakse epifenomenalismile ette lihtsalt seda, et ta on intuitsioonivastane või koguni “talumatu”. Mingi järelduse pidamine intuitsioonivastaseks või peletavaks ei ole siiski piisav põhjus tema eitamiseks, eriti kui ta soliidse mõttekäigu lõppjäreldus. Epifenomenalism võib olla intuitsioonivastane, kuid ta ei ole ilmselgelt väär, ja see tähendab, et korrektne mõttekäik, mis viib meid epifenomenalismini, tuleb tunnistada õigeks. Muidugi võib intuitsioonivastane järeldus õhutada meid oma mõttekäiku üle kontrollima, kuid me peame ikkagi leidma selles mingi vea, mis ei sõltuks tema intuitsioonivastasusest. Kui selgub, et loogilise superventsuse puudumine viib välja epifenomenalismini, siis võib küll soovida, et loogiline superventsus kehtiks, kuid me ei saa seda ilma pikemata väita, nagu oleks see lihtsalt fakt. Loogilise superventsuse tunnistamiseks on tarvis kuidagigi seletada, kuidas võiksid füüsikalistest faktidest tuleneda teadvuse faktid, aga, nagu ma olen loodetavasti näidanud, just seda me ei suudagi.

Konkreetsemad etteheited epifenomenalismile jagunevad kolme rühma: ühed puudutavad elamuse suhet tavalise käitumisega, teised elamuse suhet otsustustega elamuse kohta, ja kolmandad üldist maailmapilti, milleni me epifenomenalismi omaks võttes jõuame.

Alustame esimesest rühmast. Paljud peavad lihtsalt ilmselgeks, et just valutunne on põhjus, miks nad käe tulest välja tõmbavad, või et minu peavaluelamus ei saa olla irrelevantne seletusele, miks ma võtan tablette. Ei saa eitada, et need arvamused on intuitsiooniga väga heas kooskõlas. Teisest küljest on selle intuitsiooni allikas hõlpsasti oletamatuks seletatav nende sündmuste vaheliste süstemaatiliste reeglipäraste vastavustega. Me oleme elamusest ja käitumisest palju otsesemalt teadlikud kui neid kandvatest ajuseisunditest; täheldades süstemaatilisi reeglipäraseid vastavusi elamuse ja käitumise vahel, on loomulik järeldada neist tugeva põhjusliku seose olemasolu. Isegi kui see seos oleks ainult kaudne ja noomiline, tulenedes suhtetest elamust kandvate ajuseisunditega, on säärane järeldus ikkagi ootuspärane. Seega on võimalik sellest intuitiivsest tundest mööda minna. Kuid olgu sellega kuidas on, epifenomenalismi see vastuväide kummutada ei saa, sest sellest tundest ei saa otseselt tuletada mingit argumenti. Seega ei ole epifenomenalismil peale intuitsioonivastasuse midagi viga.

Teist liiki vastuväited teevad rohkem muret. Mõte, et meie elamused ei mängi mingit rolli näiteks seletamisel, miks me elamustest kõneleme, või ehk isegi meie seesmistes otsustustes elamuste kohta, tundub väga kummaline, palju kummalisem kui lihtsalt minu valu irrelevantsus seletusele, miks ma käe tulest välja tõmban. Mõned autorid kinnitavad ​**[lk.160]**​, et see väidetav irrelevantsus ei ole mitte lihtsalt intuitsioonivastane, vaid epifenomenalismile hukutav, võib-olla näiteks kuna see on vastuolus meie teadmisega oma elamustest või meie võimega neile viidata. Ma usun, et kui need argumendid selgesti formuleerida, siis neist niisugust järeldust tuletada ei saa, kuid need küsimused on kahtlemata siiski tõsised. Ma pühendan neile [5. peatüki](#ORG.5).

Kolmandat liiki vastuväited puudutavad epifenomenalistliku maailmapildi üldist struktuuri. Üks neist ütleb, et see pilt on inetu ja väheveenev: elamus kõlgub füüsikalise maailma küljes “nomoloogilisi ripatseid” pidi, millel pole mingit kokkupuudet teiste loodusseadustega. Ma arvan, et sellest etteheitest võib üle saada, töötades välja teooria, mis annaks terviklikuma pildi. Nimetus “epifenomenalism” kipub vihjama ettekujutusele, nagu hõljuks elamus “kusagil ülal”, sõltumatuna kõigist protsessidest; parem ja loodusliku superventsusega ikkagi ühilduv pilt kujutaks elamust istuvana “all”, põhjuslikkuse mõrades. Me võime vähemasti püüda teha psühhofüüsika seadused võimalikult lihtsaks ja elegantseks. Sellesse vastuväidete liiki kuulub ka mure teadvuse evolutsiooniloo üle epifenomenalistlikus käsitluses. Kuid pole raske näha, et see ei valmista minu positsioonile mingeid raskusi. Võtan selle küsimuse lähemalt arutluse alla [käesoleva peatüki lõpul](#PARA.101-14).

Uurimine näitab, et argumente, mis võiksid epifenomenalismi tõsiselt kõigutada, pole palju. Kõige tõsisemad argumendid on need, mis puudutavad otsustusi elamuse kohta; ma käsitlen neid [järgmises peatükis](#ORG.5.2). Kui argumendid kõrvale jätta, siis mõnedele autoritele ütleb intuitsioon, et epifenomenalism peab olema väär, ent arvestades veenvaid argumente epifenomenalismi kasuks, ei piisa tema kummutamiseks intuitsioonist.

Mina ei pea oma positsiooni epifenomenalistlikuks. Elamuse põhjusliku relevantsuse küsimus jääb lahtiseks ja selle lahendamiseks on tarvis detailsemat teooriat nii põhjuslikkuse kui ka elamuse kohta. Kuid vähemalt mingis nõrgas vormis epifenomenalismini viib minu vaade siiski ja võib viimaks viia välja ka mingi tugevama vormini. Isegi kui see nii on, arvan ma siiski, et argumendid loodusliku superventsuse kasuks on piisavalt kaalukad, et nendega nõustuda. Epifenomenalism on intuitsioonivastane, kuid alternatiivid on veelgi intuitsioonivastasemad. Nad on lihtsalt väärad, nagu me juba nägime ja allpool veel näha saame. Kogu selle probleemistiku üldine moraal on, et kui argumendid kõnelevad loodusliku superventsuse kasuks, siis tuleb õppida loodusliku superventsusega elama.

Mõned leiavad siiski, et epifenomenalistlik iseloom on minu vaate parandamatu viga. See arvamus on mulle mõneti sümpaatnegi; seda võib vaadelda kui *teadvuse paradoksi* ilmingut: kui jutt läheb teadvusele, siis võib tunduda, et kõik alternatiivid on halvad. Kuid ma arvan, et kõigi teiste positsioonide raskused on veel palju tugevamas mõttes lootusetud kui minu positsiooni intuitsioonivastasus. Arvestades, et loogilises ruumis peab leiduma õige variant, tundub minu vaade ainsa mõistliku kandidaadina.

### ​**[lk.161]**​ 5. Probleemide loogiline geograafia

Argument minu seisukoha kasuks seisneb üldjoontes järelduses neljast eeldusest:

1. Teadvuselamus on olemas.
2. Teadvuselamus ei ole füüsika suhtes loogiliselt superventne.
3. Kui leidub fenomene, mis ei ole füüsika suhtes loogiliselt superventsed, siis on materialism väär.
4. Füüsikaline maailm on põhjuslikult kinnine.

Eeldustest (1), (2) ja (3) tuleneb selgelt materialismi väärus. Ühenduses eeldusega (4) ja tõepärase eeldusega, et füüsiliselt identsetel olenditel on identsed teadvuselamused, annab see kõik kokku seisukoha, mida ma olen nimetanud looduslikuks superventsuseks: teadvuselamused kasvavad kooskõlas mingite loodusseadusega välja füüsikalisest maailmast, kuid ei ole ise füüsikalised. Mitmesugused alternatiivsed seisukohad liigituvad vastavalt sellele, kas nad eitavad eeldust (1), (2), (3) või (4). Muidugi on mõningaid neist eeldustest võimalik eitada rohkem kui ühel viisil.

.

1. eelduse eitus:

i. *Eliminativism.* Selle vaate järgi ei leidu teadvuselamuste kohta positiivseid fakte. Mitte kellelgi ei ole teadvust selle sõna fenomenaalses mõttes.

2. eelduse eitused:

Eeldust (2) võib eitada mitmel viisil, olenevalt sellest, kuidas kulgeb vastav arutluskäik, s.t. millised on need kesksed füüsikalised omadused, mida peetakse teadvuse aluseks. Ma nimetan kõiki neid seisukohti “reduktiivmaterialistlikeks”, sest nad kõik peavad võimalikuks teadvuse mõiste analüüsimist viisil, mis on ühitatav reduktiivse seletamisega.

ii. *Reduktiivne funktsionalism.* Selle vaate järgi on teadvus kontseptuaalselt tuletatav füüsikalistest omadustest funktsionaalsete või soodumuslike omaduste kaudu. See, et mingi seisund on teadvuslik, tähendab selle arusaama järgi seda, et ta mängib mingit põhjuslikku rolli. Meie maailmaga füüsiliselt identses maailmas toimiksid ka kõik vastavad põhjuslikud rollid samamoodi, ja seega oleksid ka kõik teadvusseisundid samasugused. Sombimaailm on seega loogiliselt võimatu.

iii. *Mittefunktsionalistlik reduktiivne materialism.* Selle vaate järgi on teadvuse faktid kontseptuaalselt tuletatavad füüsikalistest faktidest mingi mittefunktsionaalse omaduse kaudu. Võimalikeks kandidaatideks on biokeemilised ja kvantfüüsikalised omadused, võib-olla ka mingid seni veel määratlemata omadused.

iv. *Uue füüsika materialism.* Selle vaate kohaselt ei ole meil praegu mingit ettekujutust, kuidas füüsikalistest faktidest saaks tuletada teadvuse seletuse, kuid ​**[lk.162]**​ nii on see ainult sellepärast, et meie praegune arusaam füüsikalistest faktidest on liiga kitsas. Kui keegi väidab, et sombimaailm on loogiliselt võimalik, siis väidab ta tegelikult, et kõik ajalis-ruumilises muutkonnas omavahel vastasmõjustuvad väljad ja osakesed, mida postuleerib praegune füüsika, võiksid eksisteerida ka ilma teadvuseta. Kuid uues füüsikas ei tarvitse see nii olla. Võidakse leida mingi radikaalselt teistsugune teoreetiline mudel, mille entiteetidest piisab teadvuse tuletamiseks ja seletamiseks.

3. eelduse eitus:

v. *Mittereduktiivne materialism.* See seisukoht ütleb, et kuigi teadvuse faktid ei pruugi olla loogiliselt tuletatavad füüsikalistest faktidest ja seega ei pruugi teadvusele leiduda reduktiivset seletust, on teadvus ikkagi füüsiline. Teadvuse faktide tulenevus füüsikalistest faktidest on “metafüüsiliselt paratamatu”. Kuigi sombimaailma idees pole mingit vastuolu, on selline maailm metafüüsiliselt võimatu.

4. eelduse eitus:

vi. *Interaktsionistlik dualism.* See vaade tunnistab, et teadvus on mittefüüsiline, ent füüsiline maailm ei ole põhjuslikult kinnine, ja peab seega võimalikuks, et teadvus mängib iseseisvat põhjuslikku rolli.

Edasi tuleb minu seisukoht, mis tunnistab kõiki nelja eeldust:

vii. *Naturalistlik dualism.* Teadvus on füüsika suhtes looduslikult, kuid mitte loogiliselt või “metafüüsiliselt” superventne.

Ja viimaks võib sageli kohata ka kaheksandat seisukohta, mida kirjeldatakse enamasti üsna ebamääraselt:

viii. ​*Aimutu materialism* [​*don’t have-a-clue materialism*​]​*.* “Mul ei ole õrna aimugi, mis teadvus olla võiks. Teadvus on minu jaoks täielik müsteerium. Aga ta peab olema füüsiline, sest materialism peab olema tõsi.” Selline seisukoht on laialt levinud, kuigi trükisõnas kohtab sääraseid mõtteavaldusi harva (kuigi vt. Fodor 1992).

Teeme olukorrast, nagu ma seda näen, kiire kokkuvõtte. Variant (i) tundub olevat silmanähtavalt väär; (ii) ja (iii) toetuvad teadvuse mõiste ebaõigele analüüsile ja lähevad seega teemast mööda​[[217]](#footnote-218); (iv) ja (vi) rajavad suuri ja vähetõotavaid lootusi füüsika edasisele arengule ning takerduvad ka ületamatutesse kontseptuaalsetesse raskustesse; ja (v) kas apelleerib põhjendamatult Kripke aposterioorsele paratamatusele või siis toetub mingile ekstravagantsele metafüüsikale. (viii) on mulle mõneti sümpaatne, kuid temast tuleks viimaks välja jõuda mõnele konkreetsemale positsioonile, nendest aga ei tundu ükski paika pidavat. Ainsa vastuvõetava võimalusena jääb seega üle (vii).

Võtame nüüd aeglasemas tempos. Alustame variantidest (iv) ja (vi). Variant (vi) — interaktsionistlik dualism — nõuab, et uus füüsika avastaks lünki, mida saaks täita ​**[lk.163]**​ mittefüüsilise vaimu toimega. Seniste andmete valguses on see ebatõenäoline. Variant (iv) nõuab, et füüsika pale muutuks nii radikaalselt, et teadvuselamuse faktid saaks tuletada füüsikast; kuid mitte keegi ei oska arvata, kuidas ükskõik missugune füüsika sellega iial hakkama saada võiks. Tõepoolest, kuna füüsika tegeleb viimselt struktuursete ja dünaamiliste omadustega, siis pole näha, kuidas füüsika võiks iial anda meile midagi muud kui vaid veel rohkem struktuuri ja dünaamikat, millest ei hakka (kui mitte võtta omaks mõnd teist reduktiivset varianti) kunagi olema võimalik tuletada elamuse olemasolu.

Kõige sügavam põhjus variantide (iv) ja (vi) kõrvaleheitmiseks seisneb selles, et viimaks põrkuvad nad ikkagi samale raskusele, millega seisab silmitsi tavapärasem füüsika: fenomenaalset komponenti on vastuoludesse langemata võimalik lahutada põhjuslikust komponendist. Interaktsionistliku vaate (vi)[[218]](#footnote-219) üle arutledes nägime, et kui mittefüüsikalistel entiteetidel ongi fenomenaalne aspekt, siis on ikkagi võimalik fenomenaalne komponent neist mõtteliselt lahutada, nii et alles jääb vastasmõjude ja käitumise puhtalt põhjuslik​/​dünaamiline kirjeldus, milles pole ainsatki vastuolu. Uue füüsika vaate (iv) kohta võib öelda, et isegi kui mingi tulevane füüsika võtakski oma maailmapilti otsesõnu sisse ka fenomenaalsed omadused, ei saa nende omaduste fenomenaalsus põhjuslik​/​dünaamilises kirjelduses otsustavat rolli mängida; füüsika jääb koherentseks ka pärast selle aspekti mõttelist kaotamist. Kummalgi juhul ei ole põhjuslike vastasmõjude seletamiseks tarvis midagi peale dünaamika, ja ühestki dünaamika faktide kogumist ei ole võimalik tuletada ainsatki fenomenoloogilist fakti. Seega on endiselt võimalik pajatada lugusid sombidest.

Sellele vastamiseks võib ette võtta mitmesuguseid manöövreid, ent kõiki neid saab rakendada ka tavapärasele füüsikalisele pildile. Näiteks vaatab mingi kujuteldav​[[219]](#footnote-220) abstraktne dünaamika võib-olla mööda faktist, et interaktsionistliku vaate järgi on mittefüüsiline aines olemuslikult fenomenaalne, nii et fenomenaalsetel omadustel on põhjuslike suhete süsteemis oluline roll. Aga sama hästi võib ka olla, et praeguse füüsika abstraktne dünaamika vaatab mööda faktist, et tema alusentiteedid on olemuslikult fenomenaalsed (kirjeldab ju füüsika neid vaid väliselt). Tulemus on mõlemal juhul sama: ühtviisi jõuame välja olemuslike fenomenaalsete omaduste seletusliku irrelevantsuseni põhjuslik/dünaamilises kirjeldus. Seega ei päästa interaktsionismi või uue füüsika appivõtmine omadusdualismi, mida ma siin kaitsen, ühestki tema seesmisest raskusest. Kokkuvõttes võib neis näha lihtsalt sama positsiooni keerukamaid teisendeid.

Mis puudutab varianti (iii) (mittefunktsionalistlik reduktiivne materialism), siis selle kõige ahvatlevam versioon otsib tuletatavuse võtit mingitest tundmatutest omadustest, mida me pole seni märganud. Ent kokkuvõttes vaatab meile ka siin vastu sama raskus: füüsika annab meile ainult struktuuri ja dünaamika, struktuurist ja dünaamikast aga ei saa mitte kuidagi fenomenoloogiat. Tundub, et ainsad omadused, mida meil võtta on, on need, mis iseloomustavad füüsikalist struktuuri või funktsiooni või on neist kahest konstrueeritud. Kuid struktuursed omadused on elamuse mõiste analüüsimiseks ilmselgelt ebasobivad, ja palju parem pole lugu ka funktsionaalsete omadustega (kuigi nendest teen ma [allpool](#fn-4-36) veel juttu). Iga selline vaade vaatab kokkuvõttes teemast mööda[[220]](#footnote-221).

​**[lk.164]**​ Üle jäävad variandid (i), (ii), (v) ja (vii), mis vastavad positsioonidele, mida on meieaegses kirjanduses kõige tõsisemalt käsitletud: eliminativismile, reduktiivsele funktsionalismile, mittereduktiivsele materialismile ja omadusdualismile. Variandi (i) heidan ma kohe kõrvale, kuna ta on vastuolus ilmsete faktidega. Võib-olla annaks mingi tõepoolest väga eriskummalise argumendi varal näidata, et teadvuselamusi ei ole olemas, kuid mina ei ole kunagi näinud argumenti, mis niisuguse väite tõestusele ligilähedalegi jõuaks. Niisuguse argumendi puudumisel tähendab valik (i) lihtsalt probleemist möödahiilimist fenomeni eitamise teel.

Variant (v) (mittereduktiivne materialism) ahvatleb paljusid, kes tahaksid võtta teadvust tõsiselt ja ühtlasi jääda materialistiks. Kuid see, nagu ma olen näidanud, lihtsalt ei tööta. Mittereduktiivne materialism, nagu seda propageerib Searle, põrkub seesmistele raskustele ja taandub viimaks ühele neist teistest vaadetest (kõige sagedamini omadusdualismile). Teised selle vaate pooldajad toetuvad Kripke aposterioorsele paratamatusele; kuid sedaliiki aposterioorne paratamatus, millest kõneleb Kripke, ei suuda materialismi päästa. On üksainus viis, kuidas võtta vastuoludesse sattumata omaks variant (v): apelleerida tugevale aposterioorsele paratamatusele, mis läheb Kripke omast palju kaugemale, ja seada “metafüüsiliselt võimalike” maailmade ruumile drakoonilisi kitsendusi. Me oleme näinud, et sellistesse kitsendustesse või mingisse maailmade võimalikkuse kolmandasse, vahepealsesse astmesse ei ole mingit põhjust uskuda. See metafüüsika ei saa kinnitust üheltki muult fenomenilt ja pole selge, kuidas need saaksidki teda kinnitada.

Isegi kui see paratamatuse metafüüsika omaks võtta, sarnaneb see enamiku seletusülesannete lahendamisel vaatega, mida kaitsen mina. Sellest metafüüsikast tuleneb, et teadvus ei ole reduktiivselt seletatav; et teadvuselamus on füüsika valdkonnas seletuslikult irrelevantne; ning et teadvuseteooria peab füüsikalise ja fenomenaalse maailma ühendamiseks võtma appi mingid sildprintsiibid, mis ise ei tulene füüsikaseadustest. Kõnealune positsioon nimetab neid printsiipe “metafüüsiliselt paratamatuteks”, kuid igas praktilistes suhtes on tulemus sama. Selline teooria ühtib oma kontuuridelt dualistlike teooriatega, mida kaitsen mina, ja peaaegu kõik, mida ma ütlen järgnevates peatükkides, kus ma hakkan arendama mittereduktiivset teooriat, kehtib samavõrra ka siin.

Variant (ii) — reduktiivne funktsionalism — on kõigist materialistlikest vaadetest kõige tõsiseltvõetavam. Kui materialism on tõene, siis on teadvus — kui mitmesugused pöörased kontseptsioonid kõrvale jätta — loogiliselt superventne, ja ainus vähegi mõistlik vahend tema loogilise superventsuse näitamiseks on funktsionaalne analüüs. Sellelt positsioonilt vaadates tähendab teadvuselamus vaid teatavat põhjuslikku rolli mingis süsteemis. Fenomenaalseid omadusi käsitleb see vaade täpselt samamoodi nagu psühholoogilisi omadusi, näiteks õppimist või kategoriseerimist.

Selle positsiooni häda on muidugi selles, et ta kujutab teadvuselamuse ja teadvuslikkuse tähendust vääriti. Kui mind huvitab, kas ühel või teisel olendil on teadvus, siis ei mõtle ma tema võimetele või seesmistele mehhanismidele, millest ma võib-olla juba kõike tean; mind ​**[lk.165]**​ huvitab, kas on olemas tunne olla tema. Selle väite põhjendamiseks on mitu teadaolevat võimalust. Üks võimalus on panna tähele, et isegi kui me oleme mitmesugustele funktsionaalsetele võimetele seletuse leidnud, ei tarvitse see ikkagi lahendada elamuse seletamise probleemi. Võib ka toetuda tähelepanekule, et me oskame suvalise funktsionaalse rolli täitmist ette kujutada ka ilma, et sellega kaasneks teadvuselamusi. Kolmas võimalus on kasutada ära asjaolu, et funktsionaalsete rollide tundmine ei anna meile automaatselt teadmist teadvuse kohta. Võib mainida ka juba ülal esitatud vastuväidet, mis ütleb, et funktsionalistlik analüüs ei suuda seletada, miks saab igaüks nii hõlpsasti aru sellest, kui me ütleme kellegi kohta, et tal on teadvus​[[221]](#footnote-222), ja et ta kaotab kontseptuaalse erinevuse teadvuse ja teadlikkuse vahelt.

Kokkuvõttes ei erine reduktiivne funktsionalism kuigivõrd eliminativismist. Mõlemad möönavad, et on olemas eristamine, kategoriseerimine, informatsiooni kättesaadavus, raporteerimisvõime jms., ja mõlemad ütlevad, et kõige selle kõrval ei ole olemas enam midagi muud, mis seletust nõuaks. Nende peamine erinevus seisneb selles, et reduktsionist kinnitab, et mõned neist *explanandum*​itest väärivad elamuse nimetust, eliminativist aga, et seda ei vääri neist ükski. Kui seda terminoloogilist nüanssi mitte arvestada, siis langevad need vaated sisuliselt enam-vähem ühte. Sageli tähendatakse, et piir, mis lahutab reduktsionismi eliminativismist, on hägune, ning et sedamööda kuidas meil tuleb asjaomaseid mõisteid reduktsiooni sooritamiseks teisendada, sulavad nad teineteisega aina enam ühte. Soostudes tunnistama teadvuse olemasolu ainult kuivõrd seda saab määratleda mingi funktsionaalse võimena, teeb reduktiivfunktsionalistlik vaade teadvuse mõistele vist piisavalt liiga, et näha temas pigem üht eliminativismi erikuju. Kumbki neist kahest vaatest ei võta teadvust tõsiselt.

Nii jääbki ainukeseks vastuvõetavaks variandiks (vii) — omadusdualism, mida esindan mina. Küllap ei saa salata, et eeldused, millele ta toetub, tunduvad olevat soliidsed. Mõnes mõttes on see positsioon intuitsioonivastane, kuid ta on ainus vaade, millel pole ühtki hukatuslikku viga. Mõnele on selle vaate dualistlik iseloom vastuvõetamatu; kuid allpool püüan ma peagi näidata, et selline dualism ei ole nii ebamõistlik, nagu paljud arvavad, ja et talle ei ole kuigi palju tõsiseid vastuväiteid. Selle vaateviisi kõige suurem häda on selles, et temast tuleneb fenomenaalsete omaduste mõnetine irrelevantsus käitumise seletamisel ja ta võib, kuigi mitte automaatselt, viia epifenomenalismi. [Järgmises peatükis](#ORG.5.4) püüan ma siiski näidata, et seletusliku irelevantsuse järelmid ei ole hukatuslikud. Kokkuvõttes annab see vaade meile sidusa, naturalistliku, müsteeriumivaba arusaama teadvusest ja selle kohast looduse korras.

#### A-tüüp, B-tüüp ja C-tüüp.

Seda loogilist geograafiat laiemalt vaadates võib öelda, et vaated teadvuselamusele jagunevad kolme suurde klassi. *A-tüüpi*​ vaated ütlevad — mitmesugustel funktsionalistlikel või eliminativistlikel põhjendustel —, et teadvus, kuivõrd ta olemas on, on füüsika suhtes loogiliselt superventne. *B-tüüpi* seisukohad tunnistavad, et teadvus ei ole loogiliselt superventne ning et füüsilisest maailmast pole võimalik ​*a priori*​ tuletada fenomenaalset maailma, kuid peavad endid sellegipoolest materialistideks. *C-tüüpi* seisukohad eitavad nii loogilist superventsust kui ka materialismi.

A-tüüpi seisukohtadel on palju variante — eliminativism, biheiviorism, mitmesugused reduktiivse funktsionalismi vormid —, kuid neil kõigil on midagi ühist. A-tüüpi teoreetikud ütlevad, et (1) füüsilised ja funktsionaalsed teisikud, kellel pole samasuguseid elamusi nagu meil, on mõeldamatud; (2) Mary ei saa esimest korda punast värvust nähes maailma kohta teada midagi uut (parimal juhul omandab ta uue võime); ja (3) kõike, mida teadvuse kohta seletada on, saab seletada, seletades ära mõnede funktsioonide talitlused. Arhetüüpsete A-tüüpi teoreetikute hulka kuuluvad Arm­strong (1968), Den­nett (1991), Lewis (1966) ja Ryle (1949). Arvatavasti võib A-tüüpi teoreetikuteks pidada ka Drets­ket (1995), Reyd (1982), Rosentha­li (1996), Smarti (1959), White’i (1986) ja Wilkes’i (1984).

Enamik B-tüüpi seisukohtadest, s.t. materialismi mittereduktiivsetest vormidest, ägab seesmiste raskuste all. Ainus B-tüüpi vaade, mis näib olevat vähemalt seesmiselt koherentne, on niisugune, kus otsustav roll on tugeval metafüüsilisel paratamatusel. Seda vaadet tunnistav B-tüüpi teoreetik peab uskuma, et (1) sombid ja invertspektrid on mõeldavad, kuid metafüüsiliselt võimatud; (2) Mary saab punast värvust nähes teada midagi uut, kuid selle teadmise saab olematuks seletada analüüsiga Loari vaimus; ja (3) teadvus ei ole reduktiivselt seletatav, kuid ta on sellegipoolest füüsikaline nähtus. B-tüüpi seisukoha põhisisu ei ole keegi selgelt sõnastanud; kõige enam on täpsele formuleeringule lähenenud Levine (1983, 1993) ja Loar (1990). Autorite hulka, kes pooldavad füsikalismi ilma loogilise superventsuseta, kuuluvad veel Byrne (1993), Flana­gan (1992), Hill (1991), Hor­gan (1984b), Ly­can (1995), Pa­p­ineau (1993), Tye (1995) ja van Gulick (1992).

C-tüüpi vaadete hulka kuuluvad mitmesugused omadusdualismi vormid, mis peavad materialismi ekslikuks ja ühtesid või teisi fenomenaalseid või protofenomenaalseid omadusi mitteredutseeritavateks. Selline vaade ütleb, et (1) sombid ja invertspektrid on loogiliselt ja metafüüsiliselt võimalikud; (2) Mary saab teada midagi uut, ja see uus teadmine puudutab mittefüüsikalisi fakte; ja (3) teadvust ei saa seletada reduktiivselt, küll aga võib-olla mingite uute loodusseaduste abil mittereduktiivselt. C-tüüpi seisukohti pooldavad Camp­bell (1970), Hon­de­rich (1981), Jack­son (1982), H. Ro­bin­son (1982), W. Ro­bin­son (1988), Sprig­ge (1994) ja käesoleva raamatu autor.

Võib-olla tasub eraldi mainida ülal​[[222]](#footnote-223) juba vaadeldud seisukohta, mis samastab​[[223]](#footnote-224) fenomenaalsed omadused füüsikaliste entiteetide seesmiste omadustega. Sellist seisukohta pooldavad Feigl (1958), Lock­wood (1989), Max­well (1978) ja Rus­sell (1926), ja ka mina tunnen selle vastu mõningat sümpaatiat. Ma loen seda C-tüüpi vaadete erikujuks, sest ta peab fenomenaalseid või protofenomenaalseid omadusi fundamentaalseteks. Kuid samas on tal eriomane metafüüsiline kuju. Näiteks kaldub see vaade C-tüüpi ​**[lk.167]**​ vaadete loomuliku interpretatsiooniga võrreldes rohkem monismi poole. Teda annaks ehk nimetada ka C′-tüüpi positsiooniks, kuid tavaliselt liigitan ma ta C-tüübi alla.

A-, B- ja C-tüüpi vaadete vahelises valikus on kaks hargnemispunkti. Esiteks, kas teadvus on loogiliselt superventne (A-tüüp või üks ülejäänutest)? Teiseks, kas füsikalism on tõene (B-tüüp või C-tüüp)? Alustan teisest küsimusest. Mul pole kuigi raske kõrvale heita B-tüüpi: kuigi tal on see hea omadus, et ta võtab teadvust tõsiselt, toetub ta vastuolulisele või ebaselgele metafüüsikale ilma ühegi motiivita peale püüde iga hinna eest vältida dualismi. Kokkuvõttes ei erine see vaade oma seletuslikult kujult C-tüübist, lisades sellele vaid omajagu metafüüsilist salapära. C-tüüp on B-tüübiga võrreldes palju sirgjoonelisem.

Peamiseks hargnemispunktiks on valik A-tüübi ja teiste tüüpide vahel. Minu silmis on reduktiivne funktsionalism ja eliminativism nii ilmselt väärad, et mul on raske mõista autoreid, kes peavad õigeks mõnd A-tüüpi seisukohta. Mulle tundub, et sellist vaadet saab aktsepteerida ainult see, kes usub, et teadvus ei valmista üldse mingeid olulisi probleeme. Siiski näitab kogemus, et peaaegu kolmandik inimesi on valmis kõhklematult aktsepteerima mõnd A-tüüpi seisukohta. Siin avaldubki eessõnas mainitud Suur Vastasseis: vastasseis vaadete vahel, mis võtavad teadvust tõsiselt ja mis seda ei tee.

A-tüübi erinevus teistest tüüpidest on sügavam kui B-tüübi ja C-tüübi vaheline erinevus. Kaks viimast on lahkarvamusel suhteliselt peentes metafüüsilistes küsimustes, A-tüübi erinevus teistest puudutab aga mõningaid väga fundamentaalseid intuitiivseid ettekujutusi. Kuigi B- ja A-tüüpi vaated on mõlemad “materialistlikud”, on B-tüüp oma vaimult palju lähedasem C-tüübile: erinevalt A-tüübist tunnistavad mõlemad teadvuse probleemi sügavust.

Selle küsimuse lõplikuks lahendamiseks ei piisa argumentidest. Kui keegi kinnitab visalt, et kui on ära seletatud ligipääs ja raporteerimisvõime, siis ongi kõik seletatud; et Mary ei saa oma esimesest punasest elamusest ühtki uut teadmist maailma kohta; et funktsionaalne isomorf, mis erineks originaalist teadvuselamuste olemasolu poolest, on mõeldamatu — siis võin ma ainult järeldada, et me elame elamuse küsimuses täiesti eri tasanditel. Võib-olla meil ongi täiesti erinev siseelu. Võib-olla on üks meist teise arusaamadele “kognitiivselt kinnine”. Tõenäolisem on aga, et üks meist ajab midagi segi või on mingi dogma küüsis. Igatahes niipea kui dialektika jõuab sellesse punkti, on meie vahel kuristik, millest ei vii üle ükski argument. Me oleme jõudnud toore intuitsioonide konfliktini, nagu seda sügavate filosoofiliste küsimuste arutamisel sageli juhtub. Argumentide sõnaselge väljaütlemine võib aidata meie konflikti välja joonistada ja kirjeldada, kuid mitte lahendada.

Ütlesin raamatu algul, et minu vaateviis eeldab tõsist hoiakut teadvusele. Nüüd näeme, mida see täpsemalt tähendab. Teadvuse tõsiselt võtmine tähendab tunnistamist, et teadvuses on midagi huvitavat, mis vajab seletamist, midagi enamat kui ühtede või teiste ​**[lk.168]**​ funktsioonide talitlus.[[224]](#footnote-225) See on *prima facie* eeldus, mille kummutamiseks oleks tarvis erakordselt tugevat argumenti. Ükski argument, mida mina iial näinud olen, ei jõua selle eelduse kummutamisele ligilähedalegi. Tõepoolest, A-tüüpi teoreetikud selle eelduse vastu üldiselt argumente ei esitagi, vaid nad lihtsalt eitavad teda. Ja ümberpöördult, mingist punktist alates on peaaegu võimatu selle eelduse kasuks argumenteerida, samamoodi nagu pole võimalik tõestada teadvuselamuse olemasolu. Parimal juhul võib püüda neid küsimusi selgitada, lootuses, et see aitab jõuda paremale mõistmisele.

Need küsimused selgesti formuleeritud, võib lugeja ise otsustada, kas ta soovib teadvust tõsiselt võtta. Mina väidan ainult niipalju, et kui me võtame teadvust tõsiselt, siis ei ole meil ühtki teist mõistlikku valikut peale omadusdualismi. Kui me keeldume nõustumast reduktiivse funktsionalismiga ja eliminativismiga, siis peame vältimatult tunnistama ka, et teadvus ei ole füüsika suhtes loogiliselt superventne. Ja kui me ei nõustu loogilise superventsusega, viib meie tee otsejoones omadusdualismini. B-tüüpi vaated on populaarsed, kuid tundub, et nad ei kannata tõsist filosoofilist uurimist. Peamine metafüüsiline valik, mis jääb veel teha, on, kas valida standardne C-tüüpi vaade või C′-tüüpi vaade. Seda ei tule otsustada kohemaid — mul ei ole ka endal selle kohta veel kindlat arvamust —, kuid mõlemal juhul jääb õigeks, et kui me tahame teadvust tõsiselt võtta, peame tunnistama fenomenaalsete või protofenomenaalsete omaduste fundamentaalsust.

Filosoofilises kirjanduses võib kohata ka seisukohti, mis minu mudelisse nii selgelt ei mahu. Ent selle mudeli abil on neid juba kerge paika panna ja nende raskusi analüüsida. Käsitlen märkustes lühidalt üheksat sellist seisukohta: need on bioloogiline materialism[[225]](#footnote-226), füsikalistlik funktsionalism[[226]](#footnote-227), psühhofunktsionalism[[227]](#footnote-228), anomaalne monism[[228]](#footnote-229) [[229]](#footnote-230), representatsionalism[[230]](#footnote-231), teadvus kui kõrgemat järku mõte[[231]](#footnote-232), reduktiivne teleofunktsionalism[[232]](#footnote-233), emergentne põhjuslikkus[[233]](#footnote-234) ja müsterianism[[234]](#footnote-235) [[235]](#footnote-236).

### 6 Mõtteid naturalistlikust dualismist

Paljud, sealhulgas omal ajal ka mina, on arvanud, et teadvust saab võtta tõsiselt, jäädes materialistiks. Olen püüdnud käesolevas peatükis näidata, et see ei ole võimalik, ja seda väga selgetel põhjustel. Moraal on, et see, kes tahab teadvuse fenomeniga toime tulla, peab omaks võtma mingis vormis dualismi. Võiks öelda: materialism ja teadvus ei mahu ühte lauta.

Sellest hoolimata otsivad paljud minu siinsele positsioonile alternatiivi, sest nad ei suuda leppida tema dualistliku iseloomuga. See reaktsioon on loomulik, arvestades mitmesuguseid dualismiga seostuvaid negatiivseid assotsiatsioone, kuid ma kahtlustan, et peale meieaegsete dogmade puudub säärasel hoiakul igasugune soliidne alus. See saab selgemaks, kui kaaluda mitmesuguseid põhjendusi, mida võidakse ​**[lk.169]**​ tuua dualismist loobumise ja materialismi omaksvõtu kasuks, ja hinnata nende põhjenduste jõudu tegelikku olukorda arvestades.

Esimene põhjus materialismi eelistamiseks on lihtsus. See on väga hea põhjus: muudel võrdsetel tingimustel tuleb eelistada lihtsamat teooriat ontoloogiliselt pillavamale. Occami habemenuga ütleb, et entiteete ei tohi ilma vajaduseta rohkendada. Kuid muud tingimused ei ole võrdsed, ja meie juhul on vajadus täiesti olemas. Me nägime, et materialism ei suuda seletust nõudvaid fenomene seletada. Nagu Maxwell, kes ohverdas lihtsa mehhanistliku maailmapildi, postuleerides mõnede looduslike fenomenide seletamiseks elektromagnetvälja, peame ka meie ohverdama lihtsa materialistliku maailmapildi, et seletada teadvust. Me oleme osutanud Occamile täit lugupidamist, mida ta väärib, tunnistades, et materialismist loobumiseks on tarvis häid argumente. Ent kui need argumendid on olemas, ei suuda Occami habemenuga materialismi päästa.

Teine ja võib-olla kõige mõjuvam põhjus materialismi kasuks on induktiivne: igal pool mujal on materialism alati töötanud. Teiste fenomenide, näiteks elu, kognitsiooni ja ilma kohta on materialistlik seletus kas juba olemas või on meil tõsist alust arvata, et selleni pole palju maad minna. Miks peaks teadvusega olema teisiti?

Kuid sellest argumendist on kerge jagu saada. Oleme juba näinud​, et materialistlike kirjelduste edukus mitmesugustes välistes valdkondades on hõlpsasti seletatav. Õppimise, elu ja ilma puhul nõuavad seletust ainult struktuurid ja funktsioonid. Füüsikalise maailma põhjuslikku kinnisust arvestades on ootuspärane, et nende struktuuride ja funktsioonide kirjeldused on füüsikalised. Kuid teadvus on unikaalne juhtum: siin on tarvis seletada midagi enamat kui struktuure ja funktsioone, mistõttu pole alust oodata, et tema seletus oleks samalaadne.

Tõepoolest, ​2. peatükis nägime​[[236]](#footnote-237), et kuna meie ligipääs välistele fenomenidele on niisugune, nagu ta on, ongi ootuspärane, et kõigile neile fenomenidele leidub materialistlik seletus. Me saame oma teadmised nendest fenomenidest füüsiliste nähtuste — valguse, heli ja teiste tajumeediumide vahendusel. Füüsilise maailma põhjuslikku kinnisust arvestades on ootuspärane, et fenomenid, mida me nende meediumide kaudu täheldame, on füüsika suhtes loogiliselt superventsed — muidu me ei teakski neist midagi. Kuid meie episteemiline ligipääs teadvuselamustele on sootuks teist liiki. Teadvus on otse meie episteemilise universumi keskmes, meie ligipääsu teadvusele ei vahenda tajud. Seepärast ei kehti teadvuse puhul samasugused põhjused nagu need, mille alusel me võisime oodata materialistlikku seletust välistele fenomenidele, ja igasugune neist lähtuv induktsioon on vähemasti kahtlane.

Kolmandaks on paljud eelistanud materialismi sellepärast, et nad soovivad võtta tõsiselt teadust. Nende arvates seab dualistlik positsioon teaduse juba iseenesest kahtluse alla. Churchlandi (1988) sõnul on “dualism vastuolus ​**[lk.170]**​ evolutsioonibioloogiaga ning tänapäeva füüsika ja keemiaga”. Kuid see ei ole sugugi nii. Dualistlikus positsioonis, mida ma siin kaitsen, ei ole midagi, mis nõuaks, et seaksime kahtluse alla selle, mida ütlevad meile füüsikalist maailma kirjeldavad teadused. Füüsilise maailma põhjuslik kinnisus ei kao kuhugi; füüsika, keemia, neuroteadus ja kognitiivteadused võivad käia oma tavalist rada. Omaendi valdkondadega saavad füüsikalised teadused suurepäraselt hakkama. Füüsikalisi fenomene seletavad nad imetlusväärselt, nad ei suuda seletada vaid teadvuselamusi.

Churchland osutab, et dualismi hülgamiseks on ka teisi põhjusi: (1) vaimsete fenomenide süstemaatiline sõltuvus neurobioloogilistest fenomenidest; (2) uuemaaegse informaatika saavutused, mis näitavad, et keerukaid tulemusi on võimalik saavutada ka ilma mittefüüsiliste homunkulusteta; ja (3) tõenduste puudumine dualismi kasuks, samuti selgituste ja metoodika puudumine. Kaks esimest põhjust ei õõnesta minu seisukohta mitte kuidagi. Mis puutub kolmandasse, siis oma argumendid dualismi kasuks olen ma  [juba](#ORG.4.1.2) esitanud, dualistliku seletusviisi ja metoodika kohta aga toon ma näiteid käesoleva raamatu ülejäänud osas.

Neljanda motiivi dualismist hoidumiseks annavad paljudele kõikvõimalikud selle vaatega seostuvad spiritistlikud, usulised, supranaturalistlikud jt. ebateaduslikud konnotatsioonid. Kuid need ei puutu sugugi asjasse. Vaade, mida ma siin kaitsen, ütleb, et teadvust reguleerivad loodusseadused ning et pole võimatu, et teadvuse kohta luuakse kord mingi mõistlik teaduslik teooria. Mitte ükski aprioorne printsiip ei ütle, et kõik loodusseadused on füüsikaseadused; materialismi eitamine ei ole naturalismi eitamine. Naturalistlik dualism avardab meie maailmapilti ega too sellesse mingeid pimedusjõude.

Sarnases vaimus on paljud kartnud, et dualismi aktsepteerimine tähendaks seletamisest loobumist. Dennett (1991) väljendab seda kartust sõnadega: “Arvestades dualismi kalduvust müsteeriumides püherdada, tähendab dualismi aktsepteerimine allaandmist” (lk. 37). Võib-olla ongi mõnel dualismi vormil sääraseid kalduvusi, kuid need kalduvused ei ole kaugeltki dualismi automaatne järelm, nagu käesoleva raamatu edasistes osades loodetavasti selgeks saab.

Aeg-ajalt võib dualismi vastu kuulda ka viiendat vastuväidet: see vaade ei suutvat seletada füüsilise maailma vastasmõjusid mittefüüsilise vallaga. Kuid loodusliku superventsuse mudelis leidub sellele lihtne vastus: nende kahe valla omavaheline vastasmõjustumine toimub psühhofüüsikaseaduste vahendusel. On olemas seaduste süsteem, mis tagab, et nii- ja niisuguse füüsikalise konfiguratsiooniga kaasneb nii- ja niisugune elamus, samuti nagu on olemas seadused, mis dikteerivad, et üks või teine füüsikaline objekt mõjutab teisi objekte mingil kindlal viisil gravitatsiooniliselt.

Võidakse vastu vaielda, et see ei ütle meile, missuguse ühenduslüli kaudu see vastasmõju toimib või kuidas sünnib füüsilisest konfiguratsioonist elamus. Kuid sellise ühenduslüli otsimine on eksitus. Me ei leia ka ühenduslüli, mille kaudu toimivad fundamentaalsed füüsikaseadused. Kõik toimub lihtsalt seaduste järgi, ja kusagil tuleb punkt, kust alates “kuidas”-küsimusi enam esitada ei saa. Nagu näitas Hume, ei kanna sääraste viimsete ühenduslülide otsimine vilja. Kui sellised ühendused ongi tõepoolest olemas, siis on nad nii füüsikalisel kui ka psühhofüüsikalisel juhul täiesti tabamatud, nii et mingit omaette probleemi viimasega ei kerki.

​**[lk.171]**​ Siinkohal väärib märkimist, et juba Newtonile esitasid tema gravitatsiooniteooria oponendid[[237]](#footnote-238) sarnase vastuväite: Mil viisil saab üks keha mõjutada teist üle kauge vahemaa? Kuid ajapikku kaotas see küsimus oma jõu. Me oleme õppinud leppima sellega, et on asju, mida tuleb pidada fundamentaalseteks.

Aeg-ajalt juhitakse tähelepanu veel ühele ärevusttekitavale küsimusele: mida ütleb dualistlik mudel teadvuse evolutsioonilise tekke kohta — kas loodusse ilmus korraga uus element otsekui võluväel? Kuid ka selle pärast ei tarvitse muret tunda. Nagu füüsika fundamentaalseadused, nii on ka psühhofüüsikaseadused igavesed, eksisteerinud aegade algusest saadik. Muidugi võib olla, et universumi algsetes staadiumides ei olnud objekte, mis oleksid rahuldanud nende seaduste füüsikalisi eeldusi, mistõttu ei saanud olla ka teadvust — kuigi see sõltub nende seaduste iseloomust. Kuid nii või teisiti tekkisid universumi arenedes füüsikalised süsteemid, mis asjaomaseid tingimusi rahuldasid. Ja kui nad viimaks tekkisid, siis tekkisid koos nendega kooskõlas kõnealuste seadustega automaatselt ka teadvuselamused. Kui tunnistada koos naturalistliku dualismiga, et psühhofüüsika seadused on olemas ja et nad on ajatud, siis ei valmista teadvuse evolutsioon mingit omaette raskust.

Lühidalt öeldes riivavad vaid väga vähesed tavapärased põhjused dualismi kõrvaleheitmiseks vaadet, mida ma siin kaitsen. Peamisteks motiivideks, mis dualismi eitamiseks veel üle jäävad, on vististi lihtsalt selle nimetuse negatiivsed konnotatsioonid ning asjaolu, et dualism käib risti vastu kõigile uskumustele, mille keskel on üles kasvanud paljud meist. Niipea kui me need assotsiatsioonid läbi näeme, et jää enam mingit põhjust, miks ei võiks dualism olla mõistlik ja vastuvõetav positsioon. Ja ma arvangi, et seisukoha, mida mina olen siin visandanud, võivad pikapeale omaks võtta ja seda isegi hindama hakata kõik, kes peavad endid materialistideks, kuid tahavad ühtlasi võtta teadvuselamust tõsiselt.

Tõepoolest, paljud, kes nimetavad endid materialistideks, pooldavadki võib-olla sisimas juba minu seisukohta. Siin olen ma vaid tõstnud esile ühe loomuliku vaate ontoloogilisi järelmeid — näiteks, et teadvus “kasvab välja” füüsikalisest nähtustest. Mõni dualist peab minu seisukohta oma maitse jaoks võib-olla isegi liiga materialistlikuks, aga sinna pole midagi parata. Ideaalis annab minu positsioon sellele, kes ta omaks võtab, nii materialismi kui dualismi eelised ilma kummagi puudusteta.

Seega ei nõua see dualism meilt loobumist kuigi paljudest meie praeguse teadusliku maailmapildi olulistest elementidest, vaid ainult ühe dogma kõrvaleheitmist. Igas muus suhtes toob ta teaduslikule maailmapildile vaid täiendust — ühe tarviliku laienduse, mida on tarvis, et tuua pildile ka teadvus. Meie kreedo on: Kui see on dualism, siis tuleb õppida dualismi armastama.

## ​**[lk.172]**​ 5. peatükk. Fenomenaalse otsustuse paradoks

### 1. Teadvus ja kognitsioon

Siiani oleme rõhutanud eelkõige teadvuse ja kognitsiooni erinevusi ja nende vahelisi piire. Teadvus on salapärane, kognitsioon mitte. Teadvus on ontoloogiline uudisasi; kognitsioon on ontoloogilises mõttes tasuta lõunasöök. Kognitsiooni on võimalik seletada funktsionaalselt; teadvus avaldab sellisele seletusviisile vastupanu. Kognitsiooni reguleerivad ainuüksi füüsikaseadused; teadvus järgib mingis osas sõltumatuid psühhofüüsikaseadusi.

Kuigi paljude teadvuselamusi ümbritsevate peente metafüüsiliste ja seletuslike küsimustega tegeldes on tulnud neid erinevusi rõhutada, võib see viia eksitava pildini vaimust — niisuguseni, kus teadvus ja kognitsioon on teineteisest täiesti lahus ja kumbki elab omaenda sõltumatut elu. Võib jääda mulje, et teadvuseteoorial ja kognitsiooniteoorial ei ole teineteisega kuigi palju asja.

See mulje on eksitav: meie vaimses elus ei ole säärast enesestvõõrdumist, nagu sellise pildi põhjal arvata võiks. Teadvust ja kognitsiooni ühendavad sügavad ja fundamentaalsed seosed. Ühest küljest kuulub meie teadvuselamuste sisu kindlalt ühte meie kognitiivsete seisundite sisuga. Iga kord, kui inimesel on rohelise värvuse aisting omaenda fenomenaalse individuatsiooniga​, on tal ka vastav rohelise värvuse taju omaenda psühholoogilise individuatsiooniga​. Teisest küljest võib suur osa kognitiivsest tegevusest koonduda teadvuselamusele. Meil on oma elamuste kohta teadmised ja me moodustame nende kohta otsustusi; neid ridu kirjutades tegeleb suur osa minu mõtetest teadvusega. Need teadvuse ja kognitsiooni vahelised suhted ei ole juhuslikud ega meelevaldsed, vaid süstemaatilised.

​**[lk.173]**​ Selle süstemaatilise seose analüüs võib anda teadvuseteooriale palju lähtematerjali. Niiviisi toimides näeme, et kognitsioonil ei puudu seos teadvusega, vaid tal on teadvuse seletamises, vastupidi, keskne koht. Loomulikult ei saa kognitsiooniteooria omaenda jõududega kogu seletamistööd ära teha, kuid ta võib siiski mängida olulist rolli. On ju teadvus meile kättesaadav esmajoones just kognitsiooni kaudu. Teadvuse ja kognitsiooni vaheliste seoste põhjalik uurimine võib anda meile teadvuseteooria mõttekaks piiritlemiseks tarviliku toetuspunkti ning viia meid pikapeale niisuguse teadvusekäsitluseni, mis seda fenomeni ühest küljest ei müstifitseeri ja teisest küljest ei lamesta.

Käesolevas peatükis teen ma ära eeltöö teadvuse ja kognitsiooni suhete uurimiseks. Järgnevas jaos tulevad tutvustamisele mõned mõisted, mis on neis suhetes kesksel kohal. Peatüki ülejäänud osa on suures osas apologeetiline, käsitledes mitmesuguseid takistusi, mida teadvuse ja kognitsiooni vaheline seos, võib-olla näiliselt, mittereduktiivse vaate teele seab. Järgmises peatükis asun konstrueerima positiivset teooriat, mis käsitleb teadvuse ja kognitsiooni vahelisi suhteid süstemaatiliselt, eesmärgiga tuua need kaks kokku ja moodustada seepõhjal ühtne pilt vaimust.

#### Fenomenaalsed otsustused

Teadvuse ja kognitsiooni vahelise suhte peamiseks ühenduslüliks on fenomenaalsed otsustused. Meie teadvuselamus ei seisa kõigest muust lahus, ta ei hõlju fenomenaalses vaakumis. Me oleme oma elamusest ja tema sisust teadlikud, me moodustame tema kohta otsustusi ja esitame nende põhjal väiteid. Kui mul on punane värvuselamus, moodustan ma mõnikord uskumuse, et mul on punane elamus, ja see võib väljenduda keelelises ütluses. Me moodustame teadvuse kohta otsustusi ka abstraktsemal tasandil, näiteks juureldes teadvuse mõistatuste üle, nagu mina seda kogu siinses raamatus teinud olen. Konkreetsemal tasemel moodustame me sageli otsustusi oma teadvuselamuste objektide kohta (kasvõi nende, mis meid parajasti ümbritsevad), näiteks kui me mõtleme: “Siin on midagi punast”. Neid teadvusega külgnevaid mitutliiki otsustusi nimetan ma *fenomenaalseteks otsustusteks* — mitte sellepärast, et nad oleksid ise fenomenaalsed seisundid, vaid sellepärast, et nende sisuks on fenomenoloogia või selle objektid.

Fenomenaalsed otsustused avalduvad sageli ütlustes teadvuse kohta, mis on nende otsustuste verbaalsed väljendused. Inimeste ütlused teadvuse kohta võivad olla eri puhkudel väga erinevad: “Mul on praegu torkiv valu”, “LSD põhjustab mul veidraid värvusaistinguid”, või lõpuks ka “Teadvuse probleem teeb mind täiesti nõutuks”. Need ütlused ja otsustused on tihedas seoses meie fenomenoloogiaga, viimselt aga kuuluvad nad meie psühholoogiasse. Keelelised ütlused on käitumisaktid ja seega funktsionaalselt ​**[lk.174]**​ seletatavad. Analoogiliselt on ka fenomenaalsed otsustused ise kognitiivsed aktid ja kuuluvad psühholoogia valdkonda.

Sageli arvatakse, et uskumusi tuleb mõista kui funktsionaalseid seisundeid, mida iseloomustavad nende põhjuslikud seosed käitumisega, keskkonnaga ja teiste uskumustega. Kuid sellist vaadet ei tunnista kõik. Mõned leiavad, et uskumust või uskumuse sisu võib mingis osas konstitueerida ka fenomenaalne elamus. Eriti vaieldav tundub funktsionaalse vaate rakendamine uskumustele teadvuse kohta: kui on olemas uskumusi, mis sõltuvad teadvuselamusest, siis kõige tõenäolisemalt mängivad seda rolli uskumused teadvuse kohta. Seepärast võtan ma kõnealuste funktsionaalsete seisundite jaoks tarvitusele vähem vaieldava nimetuse “otsustus” ja jätan lahtiseks küsimuse, kas teadvust puudutav otsustus ammendab teadvust puudutava uskumuse. Me võime mõelda otsustusest kui sellest, mis jääb üle, kui lahutada uskumusest kõik fenomenaalsed kvaliteedid.

Selle üle, et on olemas puhtalt psühholoogilised seisundid, mida saab kvalifitseerida säärasteks otsustusteks, ei tohiks vaidlust olla. Kõigepealt, juba soodumus esitada mingis kindlas vormis sõnalisi teateid on psühholoogiline seisund, ja vähemalt seda soodumust võib nimetada otsustuseks. Peale selle kaasneb iga uskumuse kujunemisega oma teadvuselamuse kohta, nagu uskumuste kujunemisega üldse, kõiksuguseid funktsionaalseid protsesse. Need protsessid on aluseks soodumusele anda sõnalisi ütlusi ning ka mitmesugustele muudele soodumustele. Kui inimene usub, et LSD kutsub esile kummalisi värvusaistinguid, siis võivad selle uskumusega kaasnevad protsessid saada aluseks soodumusele LSD-d võtta või sellest hoiduda, jne. Me võime kasutada terminit “otsustus” üldhõlmava tähisena kõigile seisunditele või protsessidele, mis mängivad kõnealust põhjuslikku rolli. Esimeses lähenduses võib öelda, et otsustuse, et väide on tõene, on süsteem moodustanud siis, kui ta kaldub vastama sellekohasele küsimusele jaatavalt, käituma (tema teisi uskumusi ja soove arvestades) vastaval viisil jne.

Võib vist öelda, et otsustused on see, mis minul on ühine minu sombist teisikuga. Temal ei ole mingeid teadvuselamusi, kuid tema enda sõnul on nad tal olemas; igatahes annab ta kõigis üksikasjades samasuguseid sõnalisi ütlusi nagu mina. Nii nagu mina terminit “otsustus” tarvitan, tundub mulle loomulik öelda, et minu sombist teisiku otsustuse järgi on tal teadvuselamused ja et tema otsustused selle kohta on üksüheses vastavuses minu omadega.

Peatüki lõpus püüan näidata, et mingi osa minu fenomenaalsete uskumuste semantilisest sisust konstitueerib teadvuselamus ise. (Näiteks punase värvuse aistingutel võib olla mingi osa mõnede punase värvuse aistinguid puudutavate uskumuste sisu konstitueerimises.) Kui see nii on, siis ei ole mõned sombi otsustused nii sisukad kui minu vastavad uskumused. Sellegipoolest on nad vähemalt üksüheses vastavuses minu otsustustega, neil on samasugune vorm ja nad funktsioneerivad käitumise suunamisel samamoodi nagu minu otsustused. Nii et kui ma räägin sombi otsustusest, mille järgi tal on punane aisting, siis räägin ma millestki huvipakkuvast tema psühholoogias: vähemasti ​**[lk.175]**​ võib minu sõnu võtta deflatiivse​[[238]](#footnote-239) viitena otsustusele, mida ta väljendab sõnadega “mul on punane aisting” (või sellekõlaliste sõnadega!). Sarnasel viisil kõnelen ma edaspidi ka ütlustest, jättes need asja sisulist külge puudutavad peened küsimused kõrvale.

Rangelt võttes tulekski kõiki fenomenaalsete ütluste ja otsustuste sisukirjeldusi (näiteks viiteid kellegi otsustusele, et tal on punane aisting) lugeda säärasel deflatiivsel viisil. Kogu sisu, mida me neile omistame, kuulub kahtlemata subjekti fenomenaalsetele uskumustele, kuid otsustuse sisu küsimus ei ole nii selge just sellepärast, et ei ole selge, millist rolli mängib fenomenaalse uskumuse sisu konstitueerimisel teadvus. Ma ei pane sellele vahele käesolevas peatükis liiga suurt rõhku, sest ma tahan mõned raskused, mida fenomenaalsed otsustused minu vaate ette seavad, võimalikult selgesti sõnastada. [Peatüki lõpus](#ORG.5.8) käsitlen üksikasjalikumalt ka neid väidete ja otsustuste sisu küsimusi.

#### Fenomenaalsete otsustuste kolm liiki

Teadvuselamust puudutavad otsustused jagunevad vähemalt kolme rühma. Ma nimetan neid *esimest järku*​​, *teist järku*​ ja *kolmandat järku* fenomenaalseteks otsustusteks. Tavaliselt jätan ma piiritleva täiendi ära, nii et kõneldes “esimest järku otsustustest” jne., pean ma alati silmas fenomenaalseid otsustusi.

Esimest järku otsustused on otsustused, mis kaasnevad teadvuselamustega ja ei puuduta elamusi endid, vaid elamuste objekte. Kui mul on punane aisting — näiteks kui ma vaatan punast raamatut —, siis on käib sellega tavaliselt kaasas sõnaline või sõnatu otsustus: “On midagi punast.” Kui ma kuulen muusikalist heli ja mul on sellele vastav elamus, käib sellega kaasas kuuldud noodiga seostuv psühholoogiline seisund. Tundub olevat õige öelda, et igal objektil, millest meil on teadvuslik elamus, on ka vastav kognitiivne representatsioon, kuigi see pole veel kõik. Iga teadvuselamusega seondub vastavat sisu kandev kognitiivne seisund. Seda kognitiivset seisundit nimetan ma esimest järku otsustuseks. (Võidakse väita, et see seisund erineb mingis mõttes uskumusest või otsustusest, näiteks sellepärast, et refleksioon ei tarvitse seda kinnitada. Võtan selle küsimuse pikema arutuse alla [järgmises peatükis](#PARA.123-3), kuid praegu nimetan ma seda seisundit esimeses lähenduses “otsustuseks”).

Me võime vaadata nende esimest järku otsustuste sisule kui materjalile, millest moodustub *teadlikkuse* sisu, kui mõelda teadlikkuse all teadvuse psühholoogilist vastet, millest oli juttu [1. peatükis](#ORG.1.5.2): informatsioon, millest me oleme teadlikud, tähendab enam-vähem informatsiooni, mis on kognitiivsele süsteemile kättesaadav, mida saab kasutada verbaalseteks raportiteks jne. Need otsustused ei ole rangelt võttes otsustused teadvuse kohta. Pigem kulgevad nad teadvusega rööbiti ning on enamasti otsustused objektide ja omaduste kohta inimese keskkonnas või isegi tema peas. Õigupoolest ongi ​**[lk.176]**​ mõistlik öelda, et esimest järku otsustus on otsustus vastava elamuse sisu kohta. Kui mul on punase raamatu elamus, siis vastav esimest järku otsustus on otsustus punase raamatu kohta. Mingis mõttes võime seega öelda, et elamusel ja esimest järku otsustustel — ning seega ka teadvusel ja teadlikkusel — on ühine sisu. [Järgmises peatükis](#ORG.6.2.1) selgitan nende kahe vahelist suhet üksikasjalikumalt.

Käesolevas peatükis pakuvad mulle kõige rohkem huvi *teist järku* otsustused. Need on otsustused otsesemalt just teadvuselamuste endi kohta. Kui mul on punane aisting, siis ma mõnikord märkan, et mul on punane aisting. Ma moodustan otsustuse, et ma tunnen valu, et ma elan läbi mingeid emotsionaalseid kvaliteete, jne. Üldiselt võib vist öelda, et iga teadvuselamuse korral on inimene, kui tal on vastavad kontseptuaalsed ressursid olemas, vähemalt võimeline moodustama otsustuse, et tal see elamus on.

Teadvuselamuste kohta saab moodustada ka üksikasjalikumaid otsustusi. Võib märgata, et purpurpunane värvitoon minu elamuses on eriliselt ere, või et valu läbistab kogu minu keha, või et roheline järelkujutis on mul täna juba kolmas selline. Lisaks otsustustele konkreetsete teadvuselamuste kohta kuuluvad teist järku otsustuste hulka ka otsustused teadvuselamuse eri liikide kohta, näiteks kui inimene märkab, et mõni farmakon teeb aistingud eriti intensiivseks, või et aevastust ennustav kõditunne ninas on kuidagi eriliselt meeldiv.

​*Kolmandat järku* otsustusteks nimetan ma otsustusi teadvuselamuse kui tüübi kohta. Need otsustused ei piirdu konkreetsete elamustega. Me moodustame kolmandat järku otsustusi, kui me mõtiskleme selle üle, et meil teadvuselamused üldse on, või kui me juurdleme nende olemuse üle. Kogu käesolevas raamatus olen ma esitanud kolmandat järku otsustusi. Tüüpiline kolmandat järku otsustus võib olla näiteks: “Teadvus teeb mind nõutuks, ma ei kujuta ette, kuidas teadvust saaks reduktiivselt seletada.” Niisugused otsustused on näiteks veel “teadvuselamused on tabamatud” ja isegi “teadvuselamusi ei ole olemas”.

Kolmandat järku otsustusi moodustavad eriti sageli filosoofid, ja filosoofidest eriti need, kellel on soodumus mõtiskleda eksistentsi saladuste üle. On võimalik, et paljud inimesed elavad oma elu ära, moodustamata ainsatki kolmandat järku otsustust. Siiski moodustavad selliseid otsustusi küllalt paljud inimesed. Seletust nõuab juba asjaolu, et nad seda teevad.

Nende eristuste meelespidamiseks võib kujutada otsustusi teadvuselamuse kohta järgmiselt:

* esimest järku otsustus: /See on punane!/
* teist järku otsustus: *Ma aistin praegu punast värvust.*​[[239]](#footnote-240)
* kolmandat järku otsustus: /Aistingutes on midagi salapärast./

### ​**[lk.177]**​ 2. Fenomenaalse otsustuse paradoks

Fenomenaalsete otsustuste olemasolu toob nähtavale mittereduktiivse teadvuseteooria keskse raskuse. See seisneb järgnevas. Me nägime, et teadvus ise ei ole reduktiivselt seletatav. Kuid fenomenaalsed otsustused kuuluvad psühholoogia valdkonda ja peaksid põhimõtteliselt olema reduktiivselt seletatavad kognitiivteaduste tavapäraste meetoditega. Peab näiteks leiduma mingi füüsikaline või funktsionaalne seletus, miks on meil soodumus anda teadvuse kohta neid ütlusi, mida me anname, ning seletus, kuidas me moodustame oma otsustusi teadvuselamuste kohta. Sellest järeldub, et meie ütlusi ja otsustusi teadvuse kohta saab seletada teadvusest täiesti sõltumatult. Veel enam: näib, et teadvus on *seletuslikult irrelevantne*​ meie ütlustele ja otsustustele teadvuse kohta. Ma nimetan seda järeldust *fenomenaalse otsustuse paradoksiks*​.

Fenomenaalse otsustuse paradoks ei ole nähtavasti äratanud kuigi palju tähelepanu, kuid teda on ilmekalt kasutanud füüsik Avshalom Elitzur (1989) argumendina seisukohtade vastu, mille järgi teadvus on “passiivne”; tema pakub asemele interaktsionistlikku dualismi.[[240]](#footnote-241)  Paradoksist on kõnelnud ka psühholoog Roger Shepard (1993), kelle arvates meil tuleb sellega leppida:

Lühidalt öeldes tundub, et me seisame endistviisi dilemma ees: tundub, et mitte mingisugune ajus (või arvutis) kulgevate puhtalt füüsikaliste protsesside analüüs ei suuda tabada neile protsessidele vastava subjektiivse elamuse erilist kvaliteeti. Samas peaks mingi sedalaadi analüüs kindlasti suutma põhjuslikult seletada, kuidas indiviid mingi äsjasega sarnase mõtteni jõuab. Võib-olla peame leppima sellega, et kuigi nii teadvuselamuste olemasolul kui ka neile vastavate kvaalide sarnasussuhetel on füüsiline kehastus — füüsikalised põhjused ja tagajärjed —, ei ole teadvuselamusi ja kvaale endid võimalik kirjeldada füüsikaliste sündmustena ega edastada ühest füüsilisest süsteemist teise. (lk. 242)

Nagu eelmises peatükis nägime, on küsimus, kas teadvus on käitumisele põhjuslikult irrelevantne, keeruline metafüüsiline probleem, mis on parem jätta lahtiseks. Seevastu teadvuse seletusliku irrelevantsusega on asi selgem ja raskused suuresti samad, mis tekiksid põhjusliku irrelevantsuse korral. Olenemata sellest, missuguseks osutub põhjuslikkuse metafüüsika, tundub olevat kaunis selge, et käitumisele saab anda füüsikalise seletuse, mis ei eelda teadvuse olemasolu ega apelleeri sellele.

Kui ma ütlen jutu sees: “Teadvus on kõige salapärasem asi, mis üldse olemas on”, siis see on käitumisakt. Kui ma ühes varasemas peatükis kirjutasin: “Teadvusele ei leidu reduktiivset seletust”, siis see oli käitumisakt. Kui ma teen märkuse eriti ereda purpurpunase kvaali kohta oma elamuses, siis see on käitumisakt. Need on, nagu kõik käitumisaktid, põhimõtteliselt ​**[lk.178]**​ seletatavad minu kognitiivse süsteemi seesmise põhjusliku organisatsiooni kaudu. Võib kanda ette jutustuse neuronite aktivatsioonimustritest, mis seletab, miks need aktid aset leidsid; kõrgemal tasandil saab arvatavasti jutustada kognitiivsetest representatsioonidest ja nende suhetest kõrgtasemel, ja see lugu annab meile kõik, mida seletamiseks tarvis. Praegu me muidugi ei tea selle seletuse üksikasju, ent kui füüsikaline maailm on põhjuslikult kinnine, siis peab leiduma mingi reduktiivne seletus füüsikalistes või funktsionaalsetes mõistetes.

Oma väiteid niiviisi füüsikalistes või funktsionaalsetes mõistetes sõnastades ei tule meil kunagi viidata teadvuselamuse enda olemasolule. Füüsikaline või funktsionaalne seletus on sellest sõltumatu ja kehtib ühtviisi nii sombide kui ka tõeliste teadvuslike läbielamistega subjektide puhul. Seega näib, et teadvuselamus on fenomenaalsete ütluste seletamisel irrelevantne ning et ta on seda ka fenomenaalsete otsustuste seletamisel, kuigi need ütlused ja otsustused puudutavad kõige otsesemalt teadvuselamust!

Sellele järeldusele võidakse vastu väita, et kogu minu fenomenaalsete ütluste ja uskumuste sisu ei ole ikkagi reduktiivselt seletatav, sest selle sisu konstitueerimises on teadvusel oma roll. Võiks näiteks püüda näidata, et sombi puhul ei ole need ütlused ja uskumused needsamad (kuigi nad näevad välja ja kõlavad täpselt samamoodi!), sest sombil puudub täisväärtuslik ettekujutus teadvusest. Ent ikkagi jääb vähemasti kummaliseks, et teadvus ei puutu kuidagi häälikutesse, mida me teadvusest kõneldes kuuldavale toome, minu sõrmeliigutustesse praegusel hetkel jne.; seega jätab see vastus meid endiselt mõnevõrra nõutuks. Seepärast jätan selle vaate esialgu kõrvale ning mõtlen ütluste ja otsustuste üle endistviisi “deflatiivsel” moel, mis võimaldab neid reduktiivselt seletada.

Sellele järeldusele vastupanemiseks on ka teine viis: võidakse öelda, et iga kõrgtaseme omaduse puhul, mida võib pidada seletamisel relevantseks, leidub alati madaltaseme seletus, mis selle omaduse olemasolu ei nõua. Võidakse väita, et ka psühholoogilised omadused, nagu näiteks mälu, on seletuslikult irrelevantsed, sest meie tegudele saab anda neurofüsioloogilise seletuse, mis ei maini kordagi mälu; võiks öelda, et füüsikas on isegi temperatuur seletuslikult irrelevantne, kuna seletustes võib viited temperatuurile põhimõtteliselt asendada kirjeldustega molekulaartasemel. (Kim [1989] nimetab seda mõttekäiku *seletusliku välistamise probleemiks*​.) See võiks viia mõttele, et teadvus ei ole seletusliku irrelevantsuse mõttes halvemas olukorras kui teised kõrgtaseme omadused. Kui teadvus on ühel pulgal mälu ja temperatuuriga, siis pole ta sugugi halvas seltskonnas.

Me olema aga juba näinud, et kõik kõrgtaseme omadused, nagu näiteks temperatuur ja mälu, on füüsika suhtes loogiliselt superventsed. Sellest järeldub, et kui anda mingile tegevusele seletus neurofüsioloogia mõistetes, siis ei tee see mälu veel seletuslikult irrelevantseks. Mälu võib omandada ​**[lk.179]**​ seletusliku relevantsuse oma loogilise superventsuse varal “pärandina”. See, et mehe igatsust naissoost seltsilise järele võib seletada sellega, et ta on meessoost ja vallaline, ei tee asjaolu, et ta on poissmees, seletuslikult irrelevantseks! Üldine põhimõte on siin, et kui kahe omaduste kogumi vahel on kontseptuaalne seos, siis ei tee seletatavus ühe kogumi põhjal teist kogumit seletuslikult irrelevantseks. Mingis mõttes võib üht seletust tänu asjaomaste mõistete vahelisele kontseptuaalsele seosele võtta kui teise seletuse ümberjutustust.

Lugu mälestuste omavahelistest vastasmõjudest on mingis mõttes füüsikalise loo ümberjutustus kõrgemal abstraktsioonitasemel. Sel kõrgemal tasemel jääb välja palju füüsilise loo üksikasju, mistõttu seletus on sageli palju rahuldavam (kõik need üksikasjad võisid olla irrelevantne müra), kuid lool on siiski loogiline seos madalama taseme looga. Nii on ka temperatuuriga. Madaltaseme seletuse olemasolu ei tee sääraseid kõrgtaseme omadusi seletuslikult irrelevantseteks, samamoodi nagu molekulaarsete protsesside olemasolu palli sees ei tee seletuslikult irrelevantseks piljardipalli kiirust. Üldiselt annavad kõnealused kõrgtaseme omadused ökonoomsema ümberkirjelduse sellele, mida kirjeldab madalama taseme seletus. Võiks isegi öelda, et kõrgtaseme omaduste loogilisest superventsusest tulenevalt sisaldab madaltaseme kirjeldus neid sageli vaikivalt, ka siis, kui ta neile selge sõnaga ei viita. Kus on loogiline superventsus, seal seletusliku irrelevantsuse probleemi ei teki.

Teadvusega on raskused palju tõsisemad. Teadvus ei ole füüsika suhtes loogiliselt superventne ja seega ei saa me väita, et füüsikaline või funktsionaalne seletus hõlmab vaikivalt ka teadvust või et tänu loogilisele superventsusele saab teadvus seletusliku relevantsuse pärandiks seletuses figureerivatelt omadustelt. Käitumise füüsikalised ja funktsionaalsed seletused on teadvusest sõltumatud palju tugevamas mõttes. Neid saab esitada mõistetes, mis ei eelda teadvuselamuse olemasolugi. Teadvus on kontseptuaalselt sõltumatu mõistetest, mida kasutatakse, seletades meie ütlusi ja otsustusi teadvuse kohta.

See ei tähenda, et käitumise seletamisel ei saa mitte kunagi apelleerida teadvuselamustele. Kui inimene tõmbab käe tulest välja, siis on täiesti mõistlik seletada seda tema valuelamusega. Tunnistab ju ka mittereduktiivne vaade seaduspäraseid vastavusi elamuse ja elamusele järgneva käitumise vahel. Need seaduspärasused sõltuvad aga viimselt füüsilise tasandi seaduspärasustest. Igale käitumise seletusele, mis apelleerib valuaistingule, leidub samaväärne fundamentaalsem seletus puhtalt füüsikalistes/funktsionaalsetes mõistetes, mis ei viita mitte mingitele elamuse omadustele ega eeldagi neid; valuelamuse asemel võib ta kasutada näiteks psühholoogilise valu või valutaju mõistet. Elamus omandab mingi kaudse seletusliku relevantsuse tänu ​**[lk.180]**​ oma noomilisele seosele nende füüsiliste ja funktsionaalsete protsessidega, kuid seletused alusmõistetes teda ei vaja.

Et heita probleemile veelgi heledamat valgust, mõelgem minu sombist teisikule naaberuniversumis. Ta räägib teadvuselamusest vahetpidamata — see paistab tal lausa kinnisidee olevat. Ta veedab naeruväärselt palju aega arvuti taga küürutades, kirjutades peatükk peatüki järel teadvuse saladustest. Ta kirjeldab sageli naudingut, mida mõned meelelised kvaalid talle pakuvad, ja tunnistab oma erilist armastust sügavate roheliste ja violetsete värvitoonide vastu. Ta satub sageli poleemikasse sombidest materialistidega ja heidab neile ette, et nende vaateviis teeb teadvuselamuse tegelikele seikadele liiga.

Ja ometi ei ole tal mitte mingeid teadvuselamusi! Tema universumis on materialistidel õigus ja tema eksib. Enamik tema ütlusi teadvuselamuste kohta on täiesti paikapidamatud. Kuid kindlasti on võimalik füüsikaliselt või funktsionaalselt seletada, miks ta räägib seda, mida ta räägib. Toimub ju tema universumis kõik vastavalt seadustele ja seal ei toimu imepäraseid sündmusi, seega peab tema ütlustele leiduma mingisugunegi seletus. Kuid säärane seletus tuleb anda viimselt füüsikaliste protsesside ja seaduste mõistetes, kuna mingeid muid protsesse ja seadusi tema universumis ei ole.

(Siin võib, nagu ennegi, veenvalt väita, et kui sombi ütleb “teadvus”, siis ei pea ta silmas teadvust selle sõna täies mõttes. Esialgu tuleb kogu seda juttu sombi ütlustest ja otsustustest teadvuse kohta mõista ülalmainitud deflatiivses mõttes. Ent kuigi tal puudub täistähenduslik teadvuse mõiste, on ta moodustanud endas otsustuse, et tal on peale lisaks struktuurilistele ja funktsionaalsetele omadustele veel mingi omadus, mida ta nimetab teadvuseks, ja selles vormis ei ole probleem vähem akuutne.)

Muidugi on minu sombist teisik ainult loogiliselt, kuid mitte empiiriliselt võimalik ja me ei tohiks liialt muretseda veidruste pärast, mida võib ette tulla loogiliselt võimalikes maailmades. Ometi on meil põhjust selles punktis ärevust tunda. Iga seletus minu teisiku käitumise kohta on ju samavõrra rakendatav ka minu käitumisele, sest protsessid, mis kulgevad tema kehas, vastavad täpselt protsessidele, mis kulgevad minu kehas. Kuna tema maailmas ei ole teadvust, siis on selge, et tema ütluste seletus ei sõltu teadvuse olemasolust. Sellest järeldub, et ka minu ütluste seletus ei sõltu teadvuse olemasolust.

Et paradoksaalsust veelgi selgemini esile tuua, paneme tähele, et täpselt samamoodi arutleb ka minu sombist teisik. Me teame, kuidas ta kurdab omaenda sombist teisiku saatuse üle, kes veedab kogu oma aja teadvuse üle juureldes, kuigi tal seda ei ole. Ta muretseb selle pärast, mida ütleb see teisik teadvuse seletusliku irrelevantsuse kohta omaenda universumis. Siiski on ta täiesti kindel, et teadvus on olemas ja ei ole reduktiivselt seletatav. Kuid tema puhul on see kõik üks üüratu illusioon. Tema universumis teadvust ei ole, ja tema maailmas on eliminativistidel alati ​**[lk.181]**​ õigus olnud. Hoolimata sellest, et tema kognitiivsed mehhanismid toimivad samamoodi nagu minu omad, on tema otsustused teadvuse kohta täiesti ekslikud.

See paradoksaalne olukord äratab ühtaegu vaimustust ja rahutust. Mittereduktiivset positsiooni see veel ilmselgelt ei kummuta, kuid igatahes on siin midagi, mis nõuab tähelepanu. See on kindlasti suurim raskus, millega mittereduktiivne teooria silmitsi seisab, ja mitte ükski teooria, mis sellele probleemile vähemalt näkku ei vaata, ei saa olla lõpuni rahuldav. Me peame hoolikalt uurima selle olukorra kõiki järelmeid ja tegema vahet, millised neist on vaid intuitsioonivastased ja millised ohustavad otseselt teadvuse mittereduktiivse käsitluse põhialuseid.

Nietzsche ütles: “Mis meid ei tapa, teeb meid tugevamaks”. Kui me suudame selle paradoksiga toime tulla, võib see anda meile väärtuslikke pilguheite teadvuse suhetele kognitsiooniga. Pühendan peatüki ülejäänud osa paradoksi lahendamisele; küsimuste juurde teadvuse ja kognitsiooni vahelise seose kohta tuleme tagasi järgmistes peatükkides. Nii saame rajada teadvuseteooria palju kindlamale alusele.

(Võiks arvata, et paradoksist võib üle saada, kui võtta omaks seisukoht, mida ma [olen nimetanud](#ORG.4.5.1) B-tüüpi positsiooniks ja mille järgi teadvuse superventsus on metafüüsiliselt, kuid mitte kontseptuaalselt paratamatu, või mingi C′-tüüpi positsiooni, kus fenomenaalsed omadused on füüsilise maailma olemuse konstituendid. Kuid paradoks kerkib ka nende seisukohtade puhul peaaegu sama suure jõuga. Isegi kui nad jätavad teadvusele mingit laadi põhjusliku relevantsuse, viivad nad ikkagi seletusliku irrelevantsuseni, sest seletuslik relevantsus peab toetuma kontseptuaalsetele seostele. Isegi nende seisukohtade puhul on võimalik anda reduktiivne seletus fenomenaalsetele otsustustele, kuid mitte teadvusele endale, mistap teadvus jääb nende otsustuste suhtes seletuslikult irrelevantseks. Neile otsustustele saab niisuguste positsioonide korral anda protsessuaalse seletuse, mis ei kasuta ega eelda üheski oma etapis elamuse olemasolu; kui on veel mingeid “metafüüsiliselt paratamatuid” seoseid või seesmisi fenomenaalseid omadusi, siis on nende olemasolu kontseptuaalselt täiesti sõltumatu kõigest, mida vajab käitumise seletus.

Seda võib mõista ka teisiti, pannes tähele, et ka nende vaadete järgi on sombid ikkagi mõeldavad ja nende käitumine täiesti seletatav. Seletuses, mis kehtib sombide puhul, ei saa teadvuse olemasolu mängida mingit olulist rolli. Kuid see, mis toimub sombidega, toimub ka meiega, ja seega kehtib sama seletus ka meie kohta. Nii et isegi nimetatud vaadete korral leiduvad fenomenaalsetele otsustustele seletused, mille seisukohalt teadvus on täiesti üleliigne.)

#### Vaatame paradoksile näkku

Meie käitumise seletamisel on teadvuse seletuslik irrelevantsus enamikul juhtudest võib-olla intuitsioonivastane, kuid ​**[lk.182]**​ mitte ka liiga paradoksaalne. Seletamaks liigutust, millega ma sirutan käe raamatu järele enda ees, ei ole tarvis viidata fenomenaalsele aistingule, mida raamat minus esile kutsub; piisab, kui viidata selle asemel tajule. Võiks arvata, et kui melomaan kontserdil mõnd eriti head pala kuulates õndsalt õhkab, siis on tema käitumise seletuses kesksel kohal heliaistingute kvaliteedid, kuid selgub, et kogu seletuse võib anda ainuüksi kuulmistaju ja funktsionaalsete reaktsioonide mõistetes. Isegi seletamaks, miks ma käe tulest välja tõmban, piisab funktsionaalsest seletusest psühholoogilise valumõiste alusel.

Üldiselt selgubki, et kuigi võiks arvata, et käitumise seletamiseks tuleb viidata fenomenaalsetele omadustele, võib selle asemel tavaliselt piirduda viitega psühholoogilistele omadustele. Me nägime 1. peatükis​, et iga fenomenaalse seisundi aluseks on mingi psühholoogiline seisund. Aistingu asemel võib seletuses viidata registreeringule tajus, emotsiooni fenomenaalse kvaliteedi asemel vastavale funktsionaalsele seisundile, mõtte asemel lihtsalt mõtte sisule. Just fenomenaalsete ja psühholoogiliste omaduste omavahelise vastavuse tõttu ei valmista fenomenaalsete omaduste seletuslik irrelevantsus enamasti tõsiseid raskusi. Ta on esimesel pilgul intuitsioonivastane, ent mitte midagi hullemat. Vähemalt niisuguse käitumise puhul, kus inimene ei tegele otseselt teadvuselamusega, ei näi olevat tungivat vajadust võtta seletamisel appi fenomenaalseid omadusi.

Muret hakkab teadvuselamuse seletuslik irrelevantsus tegema siis, kui jutuks tulevad meie ütlused ja otsustused teadvuse kohta. Tõsi, esimest järku fenomenaalsete otsustuste puhul, nagu “See on mingi punane asi”, ei pruugi teadvuse seletuslik irrelevantsus meid eriti ärevaks teha. Tundub mõistlik arvata, et neid tulekski seletada puhtalt tajumise ja teiste psühholoogiliste protsesside kaudu; ei puuduta taolised otsustused ju otseselt teadvuselamusi, vaid maailma seisundit. Teist ja kolmandat järku fenomenaalsete otsustuste puhul tundub aga tõesti, et seletuslik irelevantsus põhjustab tõsiseid raskusi. Need otsustused puudutavad teadvuselamust ennast, ja just nemad on põhjus, miks me oma aistingutest kõneleme ja miks filosoofid teadvuse saladuste üle pead murravad. Üks asi on tunnistada, et teadvus on irrelevantne seletamisel, kuidas ma toas ringi käin; teine asi on tunnistada teadvuse irrelevantsust seletamisel, miks ma teadvusest kõnelen. Me kalduksime kindlasti arvama, et asjaolu, et mul on teadvus, peaks kuuluma ühe osana seletusse, miks ma ütlen, et mul on teadvus, või miks ma moodustan otsustuse, et mul on teadvus; ja ometi ei paista see nii olevat.

Ühe osa seletusest, miks me ütleme ja otsustame, et vesi on olemas, moodustab ju asjaolu, et vesi on tõepoolest olemas. Samuti on tähtede ja planeetide olemasolu peaaegu kindlasti seletuslikult ​**[lk.183]**​ relevantsed meie otsustusele, et tähed ja planeedid on olemas. Ja üldse, kui me otsustame tõeselt ja usaldusväärselt, et ​*P*​, siis mängib otsustuse *P* tõesus seletamisel, miks me selle otsustuse moodustame, enamasti keskset rolli. On otsustusi, millede objektid on otsustuste endi suhtes seletuslikult irrelevantsed. Niisugused on näiteks religioossed uskumused ja uskumused UFO-de kohta, mida saab väidetavalt seletada ilma ühelegi jumalale või UFO-le viitamata. Kuid need on kõik üsna tõenäoliselt ekslikud uskumused, mida ei saa liigitada ilmselgete teadmiste alla. Ent teadvusega on vastupidi: me teame, et meil on teadvus.

Oleme keerulises olukorras: kuidas lepitada teadmist teadvuse kohta teadvuse seletusliku irrelevantsusega fenomenaalsetele otsustustele? Kui fenomenaalsed otsustused tekivad põhjustel, mis ei sõltu teadvusest, kas see ei tähenda, et nad on alusetud? See ongi keskne raskus, millega fenomenaalse otsustuse paradoks meid silmitsi seab, ja ma käsitlen seda pikemalt [käesolevas peatükis allpool](#ORG.5.5).

See paradoks tuleneb järgmistest faktidest: (1) füüsiline maailm on põhjuslikult kinnine; (2) otsustused teadvuse kohta on füüsika suhtes loogiliselt superventsed; (3) teadvus ei ole füüsika suhtes loogiliselt superventne; ja (4) me teame, et meil on teadvus. Eeldustest (1) ja (2) järeldub, et otsustused teadvuse kohta on reduktiivselt seletatavad. Koos eeldusega (3) tähendab see, et teadvus on meie otsustustele seletuslikult irrelevantne. Kuid see ei sobi kokku eeldusega (4). Nii tulebki välja paradoks. Seda võib püüda vältida, eitades üht neist eeldustest. Vaatleme lühidalt kõiki neid pääseteid.

Mõned dualistid eitavad eeldust (1). Traditsiooniliselt on kartesiaanliku interaktsionistliku dualismi ajendiks olnud mõte, et ainult nii saab anda teadvusele tegevuse suhtes relevantsuse, mida ta väärib. Tõepoolest, Elitzur (1989) tuletab teadvust puudutavate ütluste olemasolust otseselt järelduse, et füüsikaseadused ei saa olla kõikeammendavad ning et teadvus mängib füüsikaliste protsesside juhtimisel aktiivset rolli. (Ta peab võimalikuks, et termodünaamika teine seadus on väär.) Kuid, nagu ma olen [juba](#ORG.4.4.2) püüdnud näidata, ei ole interaktsionistlikust dualismist seletusliku irrelevantsuse probleemi vältimisel kuigi palju abi.

Mõnel on võib-olla kiusatus eitada eeldust (2). Kuid tuletagem meelde, et meie definitsiooni järgi on otsustused funktsionaalsed seisundid, mis on füüsika suhtes loogiliselt superventsed. Võidakse muidugi püüda tõestada, et niisugust funktsionaalset seisundit, mis oleks ligilähedaltki sarnane sellega, mida me mõtleme otsustuse all, ei ole olemaski; kuid isegi niisugusel juhul jääb meile lihtne taganemistee: me asendame otsustused teadvuse kohta ütlustega, mis on käitumisaktid ja mille loogiline superventsus on seega otsesem. Kuid ka nii satume peaaegu sama suurtesse raskustesse. Isegi kui keegi teataks, et käitumisaktid ei ole ainuüksi füüsilised (öeldes selle põhjenduseks näiteks, et selleks, et miski ei oleks lihtsalt müra, vaid ütlus, või et miski oleks ütlus teadvuse kohta, on nõutav teadvuselamus), on ikkagi üllatav, et teadvus ei ole relevantne helide seletamisel, mida me kuuldavale toome, või märkide seletamisel, mida me kirja paneme, kuigi kõiki ​**[lk.184]**​ neid saab süstemaatiliselt pidada teadvust puudutavateks. Seega kerkivad analoogilised raskused igal juhul, olenemata sellest, kuidas me asjaomaseid seisundeid defineerime. Siiski võib sellele kaalutlusele jääda meie paradoksi käsitlemisel vähemalt abistav roll, sest teadmisega piirnevad nähtavasti kõige tihedamalt just uskumused, mitte ütlused, ja võib ka olla, et mõne fenomenaalse uskumuse sisu konstitueeribki elamus ise. [Käesolevas peatükis allpool](#ORG.5.8) tulen ma selle küsimuse juurde veel tagasi.

Reduktsionistid ja eliminativistid eitavad muidugi eeldust (3) või (4). Punkti (3) kasuks olen ma ammendavad argumendid [juba](#PARA.100-11) esitanud ega hakka neid siinkohal enam kordama. Analoogiliselt viib eelduse (4) eitamine eliminativismini, mille ma olen juba kõrvale heitnud. Peagi võtame siiski vaatluse alla ühe võimaluse, kuidas reduktsionist võiks fenomenaalse otsustuse paradoksi ära kasutada.

Mulle tundub, et kõige mõistlikum hoiak on tunnistada, et kõik neli eeldust on arvatavasti tõesed, ja uurida, kuidas neid saaks omavahel ühitada. Me teame, et teadvuselamus on olemas; et füüsikaline maailm on peaaegu kindlasti põhjuslikult kinnine; ja me oleme [ülal](#ORG.3.1) juba kindlaks teinud, et teadvus ei ole füüsika suhtes loogiliselt superventne. Tuleb leida nõks, kuidas selle kombinatsiooniga elada.

### 3. Fenomenaalsete otsustuste seletamisest

Seda kõike arvestades võib seletuse otsimist meie ütlustele teadvuse kohta pidada mõistlikuks ja huvitavaks kognitiivteaduslikuks projektiks. Need ütlused on käitumisaktid ja peaksid seega olema samavõrd seletatavad nagu kõik teisedki käitumisaktid. Ja iga kognitiivteadlast, kes sellele teele asub, peaks ootama rikkalik saak. Meie ütlused ja otsustused teadvuse kohta ei pruugi olla hõlpsasti seletatavad, kuid neid pole nii raske seletada kui teadvust ennast. Muidugi ei anna loodetav seletus automaatselt seletust ka teadvusele, kuid ta võib näidata meile kätte õige suuna.

Me ei tarvitse piirduda aktsepteerimaga säärase seletuse võimalikkust kui füüsika põhjuslikust kinnisusest ja käitumise loogilisest superventsusest tulenevat intellektuaalset järeldust. On ka sõltumatuid põhjusi arvata, et fenomenaalsed otsustused on mingit tüüpi kognitiivsete protsesside loomulik kaasnähtus ning et järele mõeldes selgub, et teatava kujundusplaaniga mõistuslikult süsteemilt tulebki niisuguseid otsustusi oodata. Kui see on nii, siis ei pruugi sääraste ütluste ja otsustuste seletamine olla nii raske, nagu arvata võiks; võib selguda, et nad tulenevad otseselt mõnest kognitiivse organitsatsiooni alusprintsiibist.

Annan esialgu vaid väga lühida vastuse küsimusele, miks keegi võiks nii arvata; võtan küsimuse põhjalikumalt ette [8. peatükis](#ORG.8.4). Et olukorda veidi piltlikumaks teha, kujutlegem, et me oleme loonud arvutiintellekti — autonoomse agendi, mis tajub oma keskkonda ja on suuteline ​**[lk.185]**​ ratsionaalselt mõtlema selle üle, mida ta tajub. Mida me võime selle süsteemi kohta öelda? Kas tal oleks mingi arusaam teadvusest või millestki selletaolisest?

Et mõista, et see on võimalik, paneme tähele, et kõige loomulikuma kavanduse korral oleks sel süsteemil kindlasti mingi arusaamine iseendast, näiteks oskus eristada end ülejäänud maailmast ja teistest omataolistest entiteetidest. Samuti tundub mõistlik arvata, et sellisel süsteemil oleks palju otsesem ligipääs omaenda kui teiste süsteemide kognitiivsele sisule. Kui tal oleks mõtlemisvõime, siis oleks tal arvatavasti ka mingi otsene teadmine oma mõtete sisust, ja ta oskaks selle asjaolu üle arutleda. Väga loomulik on ka, et niisugusel süsteemil oleks, nagu meie endigi kognitiivsüsteemil, otsene ligipääs tajuinformatsioonile.

Mida vastaks niisugune süsteem, kui temalt küsida, missugused on tema tajud? Kas ta ütleks: “Ei mingisugused”? Kas ta võiks öelda: “Nojah, ma tean, et siin on punane kolme rattaga jalgratas, aga mul pole aimugi, kuidas ma seda tean. See informatsioon lihtsalt ilmus äsja minu andmebaasi”? Võib-olla, aga pigem siiski mitte. Niisuguse ülesehitusega süsteem oleks üpris ebatõhus ja ebaloomulik; tema juurdepääs oma tajude sisule oleks kentsakalt kaudne. Palju usutavam on, et ta ütleks: “Ma tean, et siin on punane kolme rattaga jalgratas, sest ma näen seda.” Kui me temalt seepeale küsime, kuidas ta teab, et ta jalgratast näeb, siis oleks vastus arvatavasti midagi niisugust nagu “Ma lihtsalt näen.”

Süsteem, mis vastaks: “Ma tean, et ma näen jalgratast, sest minu sensorid 78–84 on nii- ja niisuguses aktivatsiooniseisundis”, oleks kaunis kummaline. Nagu osutab Hofstadter (1979), ei ole mingit vajadust anda süsteemile nii üksikasjalikku ligipääsu oma madaltaseme osistele. Isegi Winogradi (1972) programmil SHRDLU ei olnud teadmisi koodist, milles ta kirjutatud oli, kuigi ta oli võimeline oma virtuaalmaailma tajuma, selle kohta järeldusi tegema ja mõningal vähesel määral oma teadmisi isegi põhjendama. Tundub, et taolised lisateadmised on täiesti tarbetud ning teeksid vaid keerulisemaks protsesse, mille varal süsteem teadlikuna püsib ja järeldusi teeb.

Tundub, et niisugusel süsteemil oleksid oma tajude sisuga umbes samasugused suhted nagu meil — tema teadmine neist oleks otsene ja vahetu, vähemalt tema enda seisukohalt. Küsimusele, kuidas ta teab, et ta näeb punast kolme rattaga jalgratast, vastaks otstarbekalt konstrueeritud süsteem: “Ma lihtsalt näen!”. Küsimusele, kuidas ta teab, et jalgratas on punane, vastaks ta nagu meiegi: “Ma lihtsalt näen, et ta on punane.” Kui sellisel süsteemil oleks mõtlemisvõime, hakkaks ta võib-olla juurdlema, kuidas see küll nii on, et mõni asi paistab talle punasena, ja miks on punane värvus just niisugune ja sinine just naasugune. Süsteemi seisukohalt on see, et punane värvus paistab nii- ja sinine naasugusena, lihtsalt toores fakt. Meie muidugi teame omaendi vaatepunktist, et selle taga pole midagi muud, kui et punane värvus viib süsteemi ühte seisundisse ja sinine teise. Aga masina seisukohalt see ei loe.

Oma mõtiskluste käigus hakkaks ta võib-olla juurdlema fakti üle, et tal paistab olevat mingi ligipääs oma mõtetele ning et tal on minatunne. ​**[lk.186]**​ Mõtlemisvõimeline masin, millel oleks tänu oma ehitusele otsene ligipääs oma tajude ja mõtete sisule, hakkaks võib-olla peagi juurdlema teadvuse saladuste üle (selle üle arutleb pikalt Hofstadter (1985a)): “Miks kaasneb niisuguse soojusega just niisugune tunne?”; “Miks ma olen just mina, aga mitte keegi teine?”; “Ma tean, et minu protsessid on lihtsalt elektrivoolu liikumine kiipides, aga kuidas see seletab minu mõtlemis- ja tajumiselamusi?”

Loomulikult ei maksa minu äsjaseid spekulatsioone liiga tõsiselt võtta, kuid nad aitavad mõista, kui loomulik on, et me otsustame ja ütleme endil olevat teadvuse, kui meie ehitusplaan on mõistlik. Kognitiivne süsteem, mis meie küsimuse peale, mis tunne on olla tema, ei saaks aru, millest me räägime, oleks kummaline küll. Asjaolu, et me mõtleme ja kõneleme teadvusest, võib olla meie konstruktsiooni väga loomulik tagajärg, ja nii on see ka nendega. Ning seletamaks, miks need süsteemid mõtlevad ja räägivad nii, nagu nad mõtlevad ja räägivad, pole kindlasti kordagi tarvis viidata teadvusele selle mõiste täies tähenduses. Võib-olla on neil süsteemidel tõesti teadvus, võib-olla mitte; meie seletus töötab mõlemal juhul ühtviisi. Seletusi nende süsteemide talitluse kohta saab anda ainult informaatilistes mõistetes. On ilmne, et säärasel juhul ei jää “vaimule masinas” enam kohta rolli mängimiseks seletusteks.

Kõik see tähendabki (et arendada edasi üht [1. peatükis](#PARA.23-12) esitatud väidet), et teadvuse olemasolu on üllatav, kuid ütlustes teadvuse kohta ei ole midagi üllatavat. Kuigi teadvus on maailma omadus, mida me füüsikaliste faktide põhjal ennustada ei suudaks, on meie ütlused teadvuse kohta lihtlabane kognitiivne fenomen. Kognitiivset struktuuri küllalt hästi tundes oleks võimalik ilma pikemata ennustada, kui suure tõenäosusega võib süsteemilt oodata ütlusi nagu “Ma tunnen endas teadvust nii, nagu ükski füüsiline objekt seda tunda ei saa”, või isegi Descartes’i “cogito, ergo sum”. Põhimõtteliselt peaks neile ütlustele leiduma reduktiivne seletus seesmiste protsesside kaudu, mille valguses nad ei oleks üllatavamad kui ükski teine käitumise aspekt. Olen [ülal](#ORG.2.2) osutanud, mis suunas tuleks sellise seletuseni jõudmiseks liikuda, ja võtan selle küsimuse üksikasjalikuma uurimise alla ühes järgnevatest peatükkidest​.

​Allpool näeme, et töötava seletuse üksikasjadest võib teadvuseteooria lähtepunktina väga palju kasu olla. Kuid seos fenomenaalsete otsustuste seletamise ja teadvuse seletamise vahel on siiski üpris peenekoeline. Enne kui jätkata, kaalume üht lihtsamat talitusviisi selles olukorras.

#### Kas otsustuste seletamisest piisab?

Küllap on paljudele lugejatele, eriti nendele, kes kalduvad reduktsionismi poole, tulnud pähe loomulik mõte: kui on seletatud, miks me ütleme, et meil on teadvus, ja miks me väljendame otsustusi, et meil on teadvus — kas pole siis seletatud kõik, mis seletada on? Miks mitte lihtsalt loobuda teadvuseteooria otsimisest ja tunnistada teadvus üldse kimääriks? ​**[lk.187]**​ Või veel parem, miks mitte kuulutada oma teooria selle kohta, miks me otsustame, et meil on teadvus, juba täiesti korralikuks teadvuseteooriaks? On vägagi võimalik arvata, et teooria, mis seletab ära meie otsustused, annabki meile kõik, mida teadvuseteooriaks tarvis.

Mõnevõrra toetavad seda positsiooni tähelepanekud otsustuste kohta teistes valdkondades. Võiks arvata, et laialt levinud usk jumalatesse, mida leidub väga mitmesugustes kultuurides, annab meile suurepärase põhjuse uskuda, et jumalad on olemas. Kuid sellele uskumusele leidub alternatiivne, sotsiaalsetest ja psühholoogilistest teguritest lähtuv seletus. Selgitamaks, miks religioossete uskumuste lai levik on inimese natuuri ja olusid arvestades peaaegu vältimatu, võib ateist viidata tema psühholoogilisele ebakindlusele kosmose palge ees, vajadusele hingekosutava vaimse või emotsionaalse eneseväljenduse järele ning mõningate ideesüsteemide olemuslikule kalduvusele end levitada. Võib isegi viidata mõnedele väga veenvatele, kuid ekslikele argumentidele mingi jumala olemasolu kasuks, nagu näiteks otstarbekuse argument ja kosmoloogilised argumendid. Kuigi need argumendid on vigased, ei ole nad vigased ilmselgelt (eriti ei ole seda otstarbekuse argument, mida oli enne Darwini aega täiesti mõistlik veenvaks pidada), ja pole raske mõista, miks nad andsid ühtekokku tuge mõttele religioosse usu loomulikkusest.

Ettekujutust, et religioosse usu laia levikut võib seletada taolisel viisil, ühegi jumala olemasolu eeldamata, peetakse tavaliselt veel üheks tõenduseks selle kohta, et jumalaid faktiliselt ei eksisteeri. Selle tõlgenduse kohaselt suudab ateistlik hüpotees seletada looduse keerulist struktuuri sama hästi kui teistlik, ning lisaks sellele suudab ta seletada ka teistliku hüpoteesi populaarsust. See on veenev viis vastupidise vaate vundamendi õõnestamiseks. Religioosse usu puhul tundub see argument olevat väga tugev. Sama argument on ahvatlev ka siis, kui jutt on teadvusest.

Pole kahtlust, et see on kõige mõjuvam argument teadvuse reduktiivse või eliminativistliku käsitluse kasuks. Kuid mitte piisav. Analoogia ei pea paika. Ka sellega, kui me seletame ära meie otsustused teadvuse kohta, ei hajuta me teadvuse müsteeriume veel ligilähedaltki. Miks? Sest teadvus on ise *explanandum*​. Jumala olemasolu hüpotees püstitati nähtavasti suuresti selleks, et seletada kõikvõimalikke silmanähtavaid fakte maailma kohta, näiteks tema korrapärasust ja näilist otstarbekust. Kui selgub, et alternatiivne hüpotees seletab neid andmeid sama hästi, langeb vajadus Jumala hüpoteesi järele ära. Ei ole olemas mingit eraldi “Jumala fenomeni”, millele saaks osutada ja öelda: siin on ikkagi midagi, mis nõuab seletust. Tõendid on parimal juhul kaudsed[[241]](#footnote-242). Analoogiliselt postuleeritakse sageli UFO-de olemasolu, seletamaks kummalisi sündmusi taevas, kujundeid maapinnal​[[242]](#footnote-243), kadumisi Bermuda kolmnurgas, UFO-dega kohtumise “üle elanute” kinnitusi jne. Kui selgub, et neid tõendeid saab seletada ilma UFO-de olemasolu eeldamata, ei jää meil UFO-desse uskumiseks enam põhjust.

​**[lk.188]**​ Kuid teadvus ei ole inimeste käitumise või mingite maailmas toimuvate sündmuste seletamiseks loodud konstrukt, vaid toores *explanandum*​, seletust nõudev omaette fenomen. See tähendab, et kui peakski selguma, et teadvusel ei ole teiste fenomenide seletamisel mingit rolli, ei muuda see midagi. Meie tõendid teadvuse kohta ei olegi kunagi sõltunud neist teistest fenomenidest. Isegi kui meie otsustusi teadvuse kohta õnnestubki reduktiivselt seletatada, ei näita see ikkagi midagi enamat kui seda, et meie otsustusi on võimalik reduktiivselt seletada. Keha–vaimu probleemi mõte ei ole seletada meie otsustusi teadvuse kohta — vastasel korral oleks see probleem suhteliselt triviaalne. Keha–vaimu probleemi puhul on asi teadvuse enda seletamises. Kui otsustusi on võimalik seletada, seletamata teadvust, siis on see huvitav ja võib-olla üllatavgi, ent keha–vaimu probleem sellest veel ei kao.

Seisukohta, mille järgi piisab, kui seletada ära meie otsustused teadvuse kohta (samamoodi nagu piisab, kui seletada ära meie otsustused Jumala kohta), on kõige loomulikum mõista kui eliminativistlikku vaadet teadvusele (analoogilist eliminativistliku vaatega Jumalale). Sellisena seisab ta silmitsi kõigi eliminativismile loomuomaste raskustega, sealhulgas sellega, et ta eitab seda, mida tunnistab meie endi kogemus. Midagi niisugust võib teha ainult filosoof — või intellektuaalsetes pundardes sipleja. Meie punase värvuse elamused säärasest eitusest ei kao. Tunne olla meie on endiselt olemas ja nõuab endiselt seletust. Heita fenomenaalse otsustuse paradoksi tõttu kõrvale teadvus ise tähendaks visata koos vanniveega välja ka laps.

Seisukohal, mille järgi piisab fenomenaalsete otsustuste seletamisest, ei puudu intellektuaalne köitvus. Ta mõjub julge sammuna, mis lahendab puhtalt kõik probleemid ja paiskab meie segadiku põrmu, kõigile nähtavaks. Ometi veenab seesugune “lahendus” ainult umbes pooleks minutiks. Seisatades ja järele mõteldes mõistame, et me oleme seletanud vaid mõningaid meie käitumise aspekte. Me oleme seletanud, miks me teatud asju ütleme ja miks meil on soodumus neid öelda, kuid ei ole veel riivanudki keskset probleemi — teadvuselamust ennast. Kolmekümne sekundi möödudes märkame, et silmitseme punast roosi, hingame sisse selle aroomi ja mõtleme: “Miks elan ma seda läbi just nii?” — ning mõistame, et meie seletus ei ütle selle kohta midagi.

Selle asemel et võtta seda positsiooni kui üht eliminativismi vormi, võib teda tõlgendada ka funktsionalistliku vaatena, mis mõistab teadvust kui “seda, mis on aluseks otsustustele teadvuse kohta”. Kuid see on sama ebapiisav nagu teisedki teadvuse funktsionaalsed definitsioonid. Olenemata sellest, kas teadvus on faktiliselt aluseks otsustustele teadvuse kohta või mitte, ei paista see olevat kontseptuaalne tõde. On ju vähemasti ​**[lk.189]**​ loogiliselt võimalik — olgu see usutav või mitte —, et me seletame need otsustused ära, kuid jätame teadvuse seletamata. Ja sellest piisab näitamaks, et teadvuse niisugune tõlgendus on väär.

Sellel argumenteerimisliinil on ka teisi variante. Näiteks võidakse öelda, et minu arvamusele, et teadvus ei ole reduktiivselt seletatav — või et ta ei ole loogiliselt superventne, või et ta ei ole funktsionaalselt defineeritav —, on võimalik anda puhtalt reduktiivne seletus. Võiks isegi püüda reduktiivselt seletada, miks ma arvan, et teadvuselamus on *explanandum*​. Võiks arvata, et see võtaks minu varasemates osades esitatud väitedetelt pinna jalge alt ja avaks tee teadvuse reduktiivsele käsitlusele. Kuid jällegi võib see vaade meid rahuldada vaid intellektuaalse poleemika tuhinas. Viimaks tuleb ju ikkagi seletada, miks on olemas tunne olla teadvusega agent. Seletada käitumist või mingit põhjuslikku rolli tähendab seletada vale asja. See jutt võib tunduda jäärapäisena, kuid tema aluseks on lihtne printsiip: meie teooriad peavad seletama seda, mis kisendab seletuse järele.

See argumentatsiooniliin on võib-olla kõige huvitavam kõigist, mis reduktsionistil või eliminativistil valida on — kui ma oleksin reduktsionist, siis oleksin ma just niisugune reduktsionist —, kuid viimselt on tal ikkagi sama viga mis kõigil sellistel seisukohadel: ta ei seleta seda, mis nõuab seletust. Nii ahvatlev kui see seisukoht ka poleks, ei võta ta probleemi viimaks ikkagi tõsiselt. Nii lihtsate vahenditega teadvuse mõistatusest lahti ei saa.[[243]](#footnote-244).

### Dennett fenomenaalsetest otsustustest

Teiste seas pooldab vaadet, mille järgi seletust nõuavad ainult meie otsustused teadvuse kohta, Daniel Dennett. 1979. aastal ütles ta ühes artiklis:

Ma kaitsen​[[244]](#footnote-245) vaadet, et sellised otsustused ammendavad meie vahetu teadvuse, et meie individuaalne teadvusvoog ei koosne muust kui sellistest propositsionaalsetest episoodidest, või paremini öeldes: et teadvusvood, mis koosnevad ainult sellistest propositsionaalsetest episoodidest, ongi see reaalsus, mis annab ainet mitmesugusteks teadvuseteooriate pähe pakutud ebaõigeteks kirjeldusteks, nii diletantlikeks kui akadeemilisteks. Minu seisukoht seisneb otsekoheselt öeldes selles, et ei ole olemas fenomenoloogilist muutkonda, millel oleks mingisugunegi seos meie ütlustega. On olemas meie avalikud ütlused, on olemas meie propositsioonilise teadlikkuse episoodid, meie otsustused — ja see ongi kõik; edasi tuleb, niivõrd kui jutt on introspektsioonist, pimedus. (1979, p. 95)

Selle kõige peale oskan ma öelda vaid niipalju, et Dennettil peab introspektsioon käima küll hoopis teistviisi kui minul. Mina leian introspektsioonile andudes endas eest aistinguid, valu- ja tundeelamusi ja palju mitmesuguseid muid asju, millega otsustused küll kaasnevad, kuid mis ei ole ainult otsustused — kui me just ei defineeri ümber otsustuse või “meie propositsionaalse teadlikkuse episoodide” mõistet nii, et see hõlmaks ka neid elamusi. Kui mõisted niiviisi ümber määratleda, muutub Dennetti positsioon mõistlikuks, ​**[lk.190]**​ kuid siis ei ole enam põhjust arvata, et meie otsustustele saab leida reduktiivsed seletused. Kui aga mõista otsustusi kui funktsionaalselt individuaalseid seisundeid, nagu seda on soodumus ütlusi anda — ja minu arvates Dennett seda just nii mõtlebki —, siis pole tema tees enam usutav, seisnedes lihtsalt kõigi nende andmete eitamises, mille seletamine on teadvuseteooria ülesanne.

Milles võiks küll olla asi, kui inimene väidab, et introspektsioon näitab talle ainult otsustusi? Võib-olla on Dennett sombi.[[245]](#footnote-246) Võib-olla on sõnal “otsustus” tema suus mingi ebaharilik tähendus. Kõige tõenäolisemalt aga peab ta introspektsiooniks midagi muud — nii-öelda *ekstrospektsiooni*​, protsessi, mille käigus inimene vaatleb omaenda kognitiivseid mehhanisme otsekui “väljastpoolt” ja jälgib, mis seejuures toimub. Jälgides vaid kognitiivseid mehhanisme, on kerge jõuda järeldusele, et kogu töö teevad ära otsustused. Asjaomastes kognitiivsetes protsessides ei toimu midagi peale rohke kategoriseerimise, eristamise ja reageerimise. On üsna usutav, et protsesse, mis kulgevad minus, kui ma tajun kollast objekti, saab täielikult seletada mõnede võrkkesta rakkude tundlikkusega, aistingute teisendumisega seesmisteks representatsioonideks ning nende representatsioonide kategoriseerimise ja nimetamisega. Kuid nii ei seleta me introspektsiooni sisu, vaid ainult selle käigus toimuvaid protsesse. Ekstrospektsioon ei ole introspektsioon, kuigi pole raske mõista, kuidas filosoof, kellel on kalduvus mõtiskleda oma seesmiste mehhanismide üle, võib pidada üht teiseks. Teadvuselamus jääb sellisest seletamisviisist puutumata. (Pole muide võimatu, et kui äsjased kirjeldused ei suudagi öelda midagi tavalise teadvuse kohta, siis seletaksid nad see-eest suurepäraselt *pimenägemise* (sellest nähtusest tuleb pikemalt juttu [6. peatükis](#PARA.125-6)) fenomenoloogiat!)

Sarnase sammu astub Dennett raamatu “Teadvuse seletus” (1991) võib-olla kõige kesksemas argumendis. Olles esitanud oma raporteerimisvõime​teooria, tuleb Dennettil seejärel näidata, et see seletab kõike, mida on vaja seletada, sealhulgas elamust, niivõrd kui see seletamist vajab. Pikkade ettevalmistavate manöövrite järel esitab ta otsustava argumendi (lk. 363-64): elamuse teooria peab seletama, miks asjad näivad niisugustena, nagu nad meile näivad. Ja ta väidab, et tema teooria suudabki seletada, miks asjad näivad niisugustena, nagu nad meile näivad. Seega — järeldab ta — seletab tema teooria kõike, mis seletamist vajab.

See on elegantne argument, millel, erinevalt paljudest reduktsionistlikest argumentidest teadvuse asjus, ei puudu näiline usutavus. Kuid argumendi elegantsus tuleneb viisist, kuidas Dennett kasutab ära mitte kergesti märgatavalt mitmemõttelist sõna “näima”, mis balansseerib fenomenaalse ja psühholoogilise maailma vahelisel noateral. Sõnal “näib” on fenomenaalne tähendus — kui miski näib nii- ja niisugusena, siis see tähendab, et meil on neist nii- ja niisugune elamus —, ja samas ka psühholoogiline tähendus — kui miski näib nii- ja niisugusena, siis tähendab see, et meil on soodumus otsustada, et ta on niisugune. Elamuseteooria ülesanne on seletada, mida tähendab näimine selle sõna esimeses mõttes. Kuid Dennetti teooria seletab näimist selle sõna teises mõttes.[[246]](#footnote-247).

​**[lk.191]**​ Niipea kui see peen riugas on läbi nähtud, kaotab argument suurema osa oma jõust. Kui Dennett ütleb, et tema teooria seletab seda, kuidas asjad meile näivad, siis tähendab see viimselt, et see seletab vaid, miks me ütleme, et asjad on nii, ning käitume ka igas muus suhtes sellele vastavalt. (Nagu Dennett ise märgib, on tema teadvuseteooria aluseks tema kvaasibiheivioristlik sisuteooria). Kuid see pole hoopiski see, mida teadvuseteooria seletama peab. Nii tulebki välja, et seesugust teooriat teadvuseteooriaks nimetades minnakse kõigist olulistest küsimustest juba ette mööda.

Ja üldse, igaüks, kes võtab oma teadvuseteooria *explanandum*​iks fenomenaalsed otsustused, jõuab vältimatult reduktiivsele seisukohale. Kuid viimseks *explanandum*​iks ei ole otsustused, vaid elamused ise. Mitte ükski teooria, mis seletab ainult käitumissoodumusi, ei suuda seletada, miks on olemas tunne olla üks või teine teadvusega agent.

### 4. Argumendid seletusliku irrelevantsuse vastu

[Nägime](#ORG.5.2), et fenomenaalse otsustuse paradoks viib intuitsioonivastaste järeldusteni. Kuid see ongi kõik, mida me oleme siiani näinud. Mõned arvavad, et need järeldused ei ole mitte lihtsalt intuitsioonivastased, vaid hukatuslikud. Selle kinnituseks vajavad nad argumenti, mis näitaks, miks teadvus lihtsalt ei saa olla seletuslikult irrelevantne.

Selliseid argumente leida on üllatavalt raske, kuid siiski võimalik. Üldine idee on näidata, et seletuslik irrelevantsus on vastuolus mõne faktiga, mida me endi kohta teame. Ma näen selleks kolme võimalust. Võidakse öelda, et seletuslik irrelevantsus on ühitamatu tõsiasjaga, et meil on oma teadvuselamuste kohta teadmine, või et me neid mäletame, või et me nendele viitame. Ma ei pea ühtki neist argumentidest veenvaks, kuid kõik nad toovad kaasa huvitavaid küsimusi ja kõik nad väärivad uurimist.

Mõnda neist argumentidest on kõige loomulikum väljendada pigem põhjusliku kui seletusliku irrelevantsuse mõiste abil. Selleks, et anda neile argumentidele võimalus ilmutada oma täit jõudu, nõustun ma ajutiselt elamuse põhjusliku irrelevantsusega, et teha kindlaks, kas need argumendid töötavad. Võib-olla saab sarnaseid argumente esitada ainuüksi seletusliku irrelevantsuse mõiste abil, kuid nii läheksid nad keerulisemaks. Igal juhul olen ma jätnud vähemalt lahtiseks võimaluse, et elamus osutub põhjuslikult irrelevantseks: huvitav on vaadata, kas see viiks fataalsete tagajärjedeni.

Et anda oponendi vastuväidetele võimalus ilmutada oma täit jõudu, kõnelen järgnevas mõnikord otsustuste asemel uskumustest. Nagu ​**[lk.192]**​ [ülal](#PARA.104-3) mainitud, ei toetu minu peamine kaitseliin vahetegemisele uskumuste ja otsustuste vahel, ja seepärast ei pane ma sellele siinkohal kuigi palju rõhku. Siiski võib see nüanss mängida abistavat rolli. Järgnevat arutlust jälgides tuleks meeles pidada vähemalt seda, et (1) me kõneleme sombide uskumustest ja otsustustest deflatiivselt​[[247]](#footnote-248), ja (2) on võimalik, et mingit osa minu enda fenomenaalsetest uskumustest (selle sõna täielikus mõttes) konstitueerib teadvuselamus.

### 5. Eneseteadvuse argument\*

Kõige keerulisem kõigist probleemidest, mida seletuslik irrelevantsus meie ette seab, on see, mille üle ma olen [juba](#PARA.107-5) arutlenud: meie teadmine omaenda teadvuselamustest. Pealtnäha tundub, et me mitte lihtsalt ei moodusta otsustust, et meil on teadvuselamused, vaid teame seda. Aga kui mittereduktiivne vaade on õige, siis on elamus selle otsustuse moodustamisel seletuslikult irrelevantne; sama otsustuse oleksime moodustanud ka siis, kui elamust ei oleks olnud. Seepärast tundub olevat raske mõista, kuidas saab seda otsustust pidada teadmiseks.

Seda mõtet võiks sõnastada lihtsalt probleemina: kui elamus on seletuslikult irrelevantne, siis kuidas on meil võimalik elamusest midagi teada? Probleem on juba iseenesest oluline, olles on üks keskseid küsimusi teadvuselamuse kohta. Kuid seesuguseid raskeid küsimusi on juba palju, ja meil ei pruugi olla võimalik leida neile vastust enne, kui meil on üksikasjalik teadvuseteooria välja töötatud. Sama mõtet võib sõnastada ka tugevamal kujul, argumendina: kui elamus on seletuslikult irrelevantne, siis ei ole meil võimalik isegi teada, et meil on elamused. Just sellistele argumentidele püüan ma siinkohal vastata. Ma ütlen välja ka mõned ideed probleemi lahendamiseks, kuid tulen [allpool](#ORG.5.5.2) selle projekti juurde veel tagasi.

Äsja sõnastatud argumendi arendamiseks näen ma kaht võimalust, mis on teineteisega küllalt sarnased. Esiteks võib oponent võtta lähtepunktiks otseselt minu sombist teisiku võimalikkuse. Minu sombist teisik moodustab samu fenomenaalseid otsustusi mis minagi. Juhtudel, kus mina otsustan, et mul on teadvus, otsustab ka tema, et tal on teadvus. Veel enamgi, tema otsustusi produtseerivad samad mehhanismid mis minul minu omi. Kui alus, millele toetub otsustuse õigsus, tuleneb ainuüksi mehhanismidest, mille alusel otsustus moodustub — nagu sageli arvatakse —, siis võib öelda, et sombi otsustused seisavad sama kindlal alusel nagu minu omad. Samas aga ei ole tema otsustustel üldse mingit alust — nad on täielikult ja süstemaatiliselt väärad. Pealtnäha võiks sellest järeldada, et ka minu otsustustel ei saa olla alust. Neid produtseerivad samad mehhanismid, millest sünnivad sombi ekslikud otsustused, ja seega ei saa neid kindlasti teadmisteks pidada.

Kui minu fenomenaalsed otsustused ei seisa kindlamal alusel kui sombi omad, siis kaob mittereduktiivsel positsioonil kogu pind jalge alt, kuna hävib ​**[lk.193]**​ lähtekoht — meie teadmine elamuse faktist. Siit järeldub, et see asjaolu annab meile korraga nii probleemi kui ka argumenti. Probleem kõlab nii: kuidas saavad minu otsustused seista kindlamal alusel kui sombi omad, kui tunnistada, et neid kujundavad samad mehhanismid? Argument kõlab nii: kui minu otsustusi kujundavad samad mehhanismid nagu sombil, siis ei saa nad seista kindlamal alusel kui sombi otsustused.

Teine argument apelleerib *põhjuslikule teadmuseteooriale*​. Sageli kinnitatakse, et üht või teist entiteeti puudutava uskumuse põhjendamisel on otsustavaks teguriks uskumuse õige põhjuslik seos selle entiteediga. Näiteks minu uskumusi laua kohta, mida ma hetkel vaatan, kinnitab vähemalt osaliselt tõsiasi, et laual on nende uskumuste tekkega põhjuslik seos. Põhjusliku teooria pooldajad ütlevad, et selleks, et otsustust mingi objekti või olukorra kohta võiks pidada teadmiseks, peab tal olema selle objekti või olukorraga põhjuslik suhe. (Võimaliku erandi moodustavad aprioorsete teadmiste valdkonnad, näiteks kontseptuaalsed või matemaatilised teadmised.) Tundub küll sedamoodi, et kui minu uskumusel, et John on ujulas, ei ole mingit põhjuslikku suhet Johniga või ujulaga, siis ei ole mul teadmist, et John on ujulas.

Kuid elamused on, nagu ma vähemalt siinse arutluse piires möönan, põhjuslikult irrelevantsed. Teadvuselamus ei mängi mingit põhjuslikku rolli otsustuse moodustumises selle elamuse kohta. Kui põhjuslik teadmuseteooria on õige, siis järeldub siit, et meil ei ole võimalik oma elamuste kohta midagi teada. Ja taas on meie ees probleem ja argument. Probleem: kuidas on mul võimalik elamusest teada — eeldusel, et elamus ei ole minu otsustuste põhjus? Argument: kui elamusel pole minu otsustuste moodustamises põhjuslikku rolli, siis ei saa neid otsustusi pidada teadmisteks.

Sellesarnaste argumentide abil püüabki Shoemaker (1975a) põhjendada materialistlikku vaadet teadvusele ning tegelikult ka reduktiivset funktsionalismi. Shoemaker kiidab selgesõnaliselt heaks põhjusliku teadmuseteooria, öeldes, et selleks, et meil oleks elamuse kohta teadmine, peab elamusel olema põhjuslik roll meie introspektiivsete uskumuste kujunemises elamuse kohta. Reduktiivse funktsionalismi põhjendamiseks võtab ta appi ka sombiargumendi ühe erikuju. Kui sombid või nende funktsionaalsed ekvivalendid on loogiliselt võimalikud, siis on elamus introspektsioonile kättesaamatu: kuna sombidel on samad introspektiivsed mehhanismid mis meilgi, siis ei võimalda need mehhanismid kindlaks teha, kas me oleme sombid või mitte. Shoemaker järeldab sellest, et sombid ja nende funktsionaalsed ekvivalendid on loogiliselt võimatud.

Vastus kõigile neile argumentidele on minu arvates kaunis ilmne. Omadusdualist peaks väitma, et põhjuslik teadmuseteooria ei ole rakendatav meie teadmisele teadvusest ning et meie otsustusi teadvuse kohta ei saa põhjendada mehhhanismidega, mille varal need otsustused moodustuvad. Teadmine teadvuselamusest on paljudes olulistes aspektides täiesti erinev teadmisest teistes valdkondades. Meie teadmine teadvuselamusest ei seisne põhjuslikus, vaid mingis hoopis teistsuguses suhtes elamusega.

​**[lk.194]**​ Selle järelduse toetuseks võib tuua sõltumatuid põhjendusi. Nende leidmiseks võime vaadelda veel üht võimalust, kuidas omadusdualist võiks püüda vastata: ta võib võtta appi *reliabilistliku*​[[248]](#footnote-249) teadmuseteooria. See võib esialgu tunduda paljutõotavana, kuid minu arvates on ilmne, et reliabilistlik teooria ei ole rakendatav meie teadmisele teadvusest, ning selgub, et põhjuslik teooria on sobimatu samal põhjusel.

Reliabilistliku teooria järgi on uskumusel alust, kui protsess, mille varal ta moodustus, oli usaldusväärne — see tähendab, niisugune, millest tavaliselt sünnivad tõesed uskumused. Näiteks pertseptuaalsetel uskumusel on alust, kui ta moodustub optilise stimulatsiooni tagajärjel, mida avaldavad ümbritsevad objektid, kuna see on protsess, mis produtseerib tavaliselt tõeseid uskumusi; kuid ta on alusetu juhul, kui ta on hallutsinatsiooni tagajärg, mis on väga ebausaldusväärne mehhanism. Mittereduktiivne elamuseteooria on täiel määral ühildatav meie fenomenaalsete otsustuste usaldusväärsusega aktuaalses maailmas: vähemasti tundub noomilist korrelatsiooni silmas pidades tõenäolisena, et kui inimene moodustab otsustuse, et tal on visuaalne elamus, siis tal ongi visuaalne elamus. Seevastu minu sombist teisiku fenomenaalsed otsustused on läbinisti ebausaldusväärsed ja üldiselt ka väärad.

Seepärast võib reliabilistlik teooria tundudagi pääseteena meie raskustest: sellest teooriast järeldub, et meie otsustustel elamuse kohta võiks olla alust isegi otsese põhjusliku seose puudumisel, ja ühtlasi on tal ressursid seletamaks tõsiasja, et minu otsustustel on alust, minu sombist teisiku omadel aga mitte. Paljud leiavad siiski, et apelleerimine reliabilistlikule teooriale ei ole rahuldav väljapääs; see näeb välja pigem kahtlase manöövrina, mis ei suuda kanda talle peale pandud koormat. Teadmine, mida reliabilistlik teooria meil möönab olevat, näib liiga haprana, et pidada teda võrdväärseks meie teadmisega oma teadvuselamuse kohta. Järele mõeldes pole ka raske mõista, miks see nii on.

Häda on selles, et kui meie uskumusi teadvuse kohta õigustaks ainult usaldusväärne seos, siis ei saaks meil olla täit kindlust selles, et meil on teadvus. Usaldusväärse seose pelgast olemasolust ei piisa tõsikindluseks, sest me ei saa välistada võimalust, et usaldusväärset seost siiski ei ole[[249]](#footnote-250) ja et seose teises otsas ei olegi teadvust. Ainus võimalus oma teadvuse olemasolus täiesti kindel olla oleks, kui meil oleks veel mingi teine võimalus pääseda ligi seose teisele otsale; kuid see tähendaks, et meie teadmisel teadvusest oleks veel mingi muu alus. Sellist olukorda peetakse sageli talutavaks, kui jutt on meie teadmistest välise maailma kohta: me ei tarvitse olla kindlad selles, et toolid on olemas, selleks et (igapäevases mõttes) teada, et toolid on olemas, ja seepärast ei tee meile muret see, et meil pole täit kindlust toolide usaldusväärses seoses meie otsustustega toolide kohta. Kuid selles, et meil on teadvus, oleme me täiesti kindlad; vähemasti on see kindlus üheks alustoeks positsioonile, mida ma olen siin kaitsnud. Teadmisele, et meil on teadvus, on ehk võimalik vastu seada mitmesuguseid “filosoofilisi” vastuväiteid, ent mitte kuigi otsesel viisil — analoogiliselt kahtlustega meie teadmises ​**[lk.195]**​ välismaailma olemasolust; see oleks aga võimalik, kui uskumustel poleks mingit muud põhjendust peale usaldusväärse seose.

Uskumused, mida õigustab ainult usaldusväärne seos, on alati ühildatavad *skeptiliste hüpoteeside* olemasoluga stsenaariumide kohta, kus asjad tunduvad subjektile täpselt samadena, kuid vastavad uskumused on väärad usaldusväärse seose puudumise tõttu. Näiteks juhuks, kui meid huvitab pertseptiivne teadmine, võib konstrueerida olukorra, kus usaldusväärne seos puudub — näiteks kus subjektiks on vaadiaju —, kusjuures subjektile näib kõik olevat ikka samamoodi. Subjekti episteemilises lähtesituatsioonis ei ole midagi, mis selle stsenaariumi välistaks. Kuid teadvuse puhul ei ole taoliste skeptiliste hüpoteeside konstrueerimine võimalik. Meie episteemilises lähtesituatsioonis on teadvuselamus juba sees. Ei ole olukorda, kus kõik tundub meile täpselt samamoodi, kuid meil ei ole teadvust, kuna seda, kuidas asjad meile tunduvad, konsituteerib (vähemalt osaliselt) just meie teadvuselamus.

Väärib mainimist, et muusuguste, näiteks välist maailma puudutavate teadmiste puhul hoolitsevad filosoofid skeptilisi stsenaariume konstrueerides alati püüdlikult selle eest, et stsenaarium oleks kõige suhtes, mis puudutab elamust, identne esialgse stsenaariumiga. Nagu tähendas juba Descartes, on olemas piir, millest skeptitsism ei saa kaugemale minna. Skeptilised stsenaariumid, milles on olulisel kohal hoopis teistsugused elamused — näiteks suur hulk kollaseid ja rohelisi sähvatusi ühes kõrvulukustava müraga — välistuvad automaatselt. Me teame (palju tugevamas mõttes kui enne), et see ei ole olukord, milles me viibime.

Sellest järeldub, et reliabilistlik kontseptsioon teadmisest ei saa anda meile teadmist, mis oleks piisavalt soliidne, et seda võiks pidada samaväärseks meie teadmisega teadvuselamusest, ning ei ole seepärast praegusel juhul rakendatav. Kuid kõik, mida ma olen öelnud reliabilistliku teadmiskontseptsiooni kohta, on kohaldatav ka põhjuslikule kontseptsioonile. Kus on põhjuslikkus, seal on ka kontingentsus: olemasolevat põhjuslikku seost ei pruugiks olla. Kui ainus allikas, millega me saame õigustada uskumust ​*X*​, on põhjuslik seos ​*X*​-ga, siis ei ole subjektil võimalik täie kindlusega teada, et see põhjuslik seos on olemas. Ainsa võimaluse seda kindlalt teada annaks talle sõltumatu ligipääs ​*X*​-le või vastavale põhjuslikule ahelale, kuid see eeldaks teadmist, mille toetuks millelegi enamale kui see põhjuslik ahel ise. Alati leidub skeptiline stsenaarium, kus subjektile näib kõik olevat täpselt sama, kuid põhjuslik seos puudub ja ​*X* ei eksisteeri, ja seega ei saa subjektil olla ​*X* kohta kindlat teadmist. Kuid me teame kindlalt, et meil on teadvus, ja seega ei pea selle teadmise põhjuslik seletus paika.

Loomulikult võib oponent lihtsalt mitte nõustuda, et meie teadmine teadvuse kohta on soliidne, ja teatada, et skeptilised stsenaariumid, mida me ei saa välistada, on olemas — näiteks stsenaarium sombidega. Kuid igaüks, kes sellele seisukohale asub, oli teadvuse küsimuses arvatavasti juba algselt eliminativist (või reduktiivne funktsionalist). Kui keegi tunnistab, et meie otsesed andmed ei välista võimalust, et me oleme sombid, siis peab ta omaks võtma ​**[lk.196]**​ järelduse, et me olemegi sombid: pealegi muutub ka maailmapilt seeläbi palju lihtsamaks. Kuid põhjus, miks teadvuse probleem üldse eksisteerib, on selles, et meie otsesed andmed[[250]](#footnote-251) just nimelt välistavad selle võimaluse. Võtta teadvust tõsiselt tähendab tunnistada, et meil on otsesed andmed, mis välistavad võimaluse, et teadvust ei ole olemas. Loomulikult saab kõige selle vastu vaielda tavapärasel viisil, kuid mõte on selles, et ei ole enam mingit erilist põhjust hakata selle vastu vaidlema nüüd, kus me oleme oma mõttekäiguga juba siiani jõudnud. Eliminativistide ja reduktiivsete funktsionalistidega on meie teed ammu lahku läinud. Kui me võtame teadvust tõsiselt, siis on meil vägagi palju alust uskuda, et meie fenomenaalse teadmuse põhjuslik või reliabilistlik seletus ei pea paika.

#### Millele toetuvad fenomenaalsed otsustused?

Ülalkirjeldatud kontseptsioonide peamine viga seisneb selles, et nad kujutavad meie ligipääsu teadvusele kaudsena, nagu seda on ka meie ligipääs keskkonna objektidele, kuna seda vahendab mingi põhjuslik ahel või usaldusväärne mehhanism. Selline vahetalitus on omal kohal siis, kui meie episteemilise lähtesituatsiooni ja kõnealuste fenomenide vahel on lünk, nagu see on välismaailma puhul: meid lahutab keskkonna objektidest mingi vahemaa. Kuid intuitiivselt on selge, et meie ligipääsu teadvusele ei vahenda miski. Teadvuselamus asetseb meie episteemilise universumi keskmes; meil on talle otsene ligipääs.

Siinkohal tuleb küsida: millele nimelt toetuvad meie uskumused oma elamuste kohta, kui nad ei saa toetuda põhjuslikule seosele nende elamustega ning mehhanismidele, mille varal need uskumused kujunevad? Ma arvan, et vastus sellele küsimusele on selge: neid uskumusi kinnitab vastavate elamuste olemasolu. Näiteks fakt, et mul on praegu punane elamus, kinnitabki minu uskumust, et mul on punane elamus. Vahetage punane elamus teistsuguse elamuse vastu või kaotage ta üldse, ja minu uskumuse peamine tugi on kadunud. Kui ma usun, et minu elamuseks on valju müra kuulmine, siis toetub see uskumus peamiselt elamusele, mille kutsub esile see vali müra. Tõepoolest — võiks küsida —, millele muule ta saakski toetuda?

Et seda asjaolu veelgi paremini esile tõsta, paneme, nagu ennegi, tähele, et elamus on üks osa meie episteemilisest lähtesituatsioo­nist. Asendage minu erepunased elamused tuhmroheliste elamustega, ja te olete kaotanud tõendusmaterjali mõnede minu uskumuste kasuks, sealhulgas uskumuse kasuks, et mul on erepunane elamus. See asjaolu kajastub faktis, et ei ole võimalik konstrueerida skeptilist stsenaariumi, kus ma oleksin kvalitatiivselt samaväärses episteemilises olukorras, ent minu elamused oleksid radikaalselt teistsugused. Minu elamused on üks osa minu episteemilisest situatsioonist, ja juba ainuüksi nende olemasolu on tõendus mõnede minu uskumuste kasuks.

See kõik tähendab, et elamuses on midagi olemuslikult episteemilist. Elamusega käib automaatselt kaasas mingit laadi ​**[lk.197]**​ intiimne episteemiline suhe selle elamusega — suhe, mida võiks nimetada *tutvuseks*​. Ei ole isegi kontseptuaalset võimalik, et subjektil oleks säärane punane elamus, ilma et tal oleks sellega mingisugustki episteemilist puudet: elamuse olemasolu tähendab ka nimetatud suhet selle elamusega.

Pange tähele: ma ei ütle, nagu tuleks elamusega automaatselt kaasa ka teadmine elamusest, samas mõttes, nagu teadmine nõuab uskumust. Ma arvan, et niisugune väide oleks ekslik: meil on palju elamusi, mille kohta meil ei ole uskumusi ja millest me seega ka midagi ei tea. Veelgi enam, meil võib olla elamusi, mida me mitte mingil viisil ei kontseptualiseeri. Midagi läbi elada ja järelikult selle läbielamisega tuttav olla tähendab olla temaga suhtes, mis on elementaarsem kui uskumine: suhtes, mis annab meie uskumuste kasuks tõenduse, kuid ei ole ise uskumus.

Tõepoolest, sellest, mida ma olen öelnud, ei järeldu kuidagi, nagu oleksid kõik uskumused elamuste kohta eksimatud, selles mõttes et elamus annaks igale sellisele uskumusele automaatselt täieliku kinnituse. Kuna uskumusi elamuste kohta lahutab elamustest endist mingi vahemaa, siis võib neid uskumusi kujuneda kõikvõimalikel põhjustel, ja mõnikord moodustub ka alusetuid uskumusi. Näiteks kui inimesel on mõtted mujal, siis võivad tema otsustused oma elamuste kohta olla täiesti ekslikud. Ma ei taha väita, nagu oleks elamuse olemasolu ainus tegur, millest oleneb elamust puudutava uskumuse põhjendatavus või põhjendamatus. Tahan öelda ainult niipalju, et elamus on üks niisugustest teguritest, võib-olla peamine tegur, ning et ta annab uskumusele elamuse kohta potentsiaalse põhjendusallika, mida elamuse puududes ei oleks.

Mõni leiab võib-olla, et see kõik on vaid *ad hoc*​ konstruktsioon ühe kahtlase väärtusega teooria päästmiseks. Mina leian, et minu jutt ei ole *ad hoc*​. Ka täiesti sõltumatult kõigist kaalutlustest seletusliku irrelevantsuse kohta on meil väga mõjuvaid põhjusi uskuda, et elamuse epistemoloogia on midagi erilist, et ta on väga erinev epistemoloogiast teistes valdkondades. Paljud on kõnelnud meie “otsesest teadmisest” või “tutvusest” elamusega, ja see pole olnud kaitsemanööver olukorra sunnil. Paljud on isegi väitnud, et teadmine elamusest on igasuguse teadmise aluspõhi just sellepärast, et meil on elamusega nii otsene episteemiline suhe. Oleks võib-olla liialdus öelda, et kõik teadmised tulenevadki ainult teadmisest elamuse kohta, kuid üldist mõtet, et meie teadmises oma elamuse kohta on midagi erilist, ei ole keegi kunagi ümber lükanud[[251]](#footnote-252).

Analoogiliselt on lihtne sõltumatult põhjendada ka väidet, et elamused ise ongi tõend meie uskumustele elamuste kohta. Näiteks Siewert (1994) — kes võtab teadvust tõsiselt, kuid ei ilmuta vähimatki sümpaatiat vaatele, mille järgi teadvus on seletuslikult irrelevantne — kaitseb põhjalikult seisukohta, mis ütleb, et meil on uskumustele oma elamuste kohta “esimese isiku garantii”, mis toetub vähemalt mingis osas asjaolule, et meil need elamused ​**[lk.198]**​ on. Seega ei ole tarvis mingeid erilisi kunstvigureid, et näha elamuses otsest tõendusallikat.

#### Vastused argumentidele

Kõike seda arvestades võib seletusliku irrelevantsuse vastastele argumentidele anda sirgjoonelise vastuse. Vastuseks põhjusliku teadmuseteooria argumendile mainime, et on sõltumatuid põhjusi arvata, et põhjuslik teooria ei suuda seletada meie teadmist elamusest: see teadmine toetub otsesemale suhtele. Ja vastuseks sombist teisiku argumendile märgime, et minu elamust puudutavate uskumuste põhjendamiseks on tarvis veel midagi peale mehhanismide, mille varal need uskumused moodustuvad: määrav koht on ka elamustel enestel. Kuna minu sombist teisikul elamused puuduvad, siis on ta episteemiliselt hoopis teistsuguses situatsioonis kui mina, ja tema otsustusi niiviisi põhjendada ei saa.

Ahvatlevana võib tunduda vastuväide, et kui minu uskumus kuulub füüsilisse valda, siis peab ka põhjendus olema füüsiline. Kuid selline järeldus oleks ekslik. Füüsilise aluse puudumisest võiks järeldada, et alus puudub minu füüsilise osa (näiteks minu aju) uskumusel. Kuid küsimus ei ole selles, kas minu ajul on alust seda uskuda, vaid kas seda on minul, ja kui omadusdualism on õige, siis olen ma midagi enamat kui minu aju. Mind konstitueerivad nii füüsikalised kui ka mittefüüsikalised omadused, ja ainult ühest poolest kõneldes ei saa öelda minu kohta kõike. Ja see tähendab, et minu uskumusele ei anna alust mitte ainult minu füüsikalised, vaid ka mõned minu mittefüüsikalised omadused — nimelt elamused ise.

Sellele saaks veelgi vastu vaielda: “Aga uskumus oleks ikkagi kujunenud, isegi kui elamust poleks olnud!”. Sellele on vastus: “Ja mis siis?” Ühel juhul on mul minu uskumuse kinnituseks tõend, nimelt minu vahetu tutvus elamusega. Teisel juhul see tõend puudub. Kui uskumus oleks mingil teisel juhul võinud tekkida ka tõendusmaterjali puududes, siis ei tähenda see, et tõendusmaterjal ei suuda põhjendada minu uskumust praegusel juhul.[[252]](#footnote-253) Ma tean, et mul on teadvus, ja selle teadmise ainsaks aluseks on minu vahetu elamus. Öelda, et elamus ei mõjuta minu psüühika talitlust, ei ole sama mis öelda, et elamus ei puutu minusse endasse.

Ja lõpuks võib neis olukordades ikka ja jälle refräänina kuulda: “Aga sinu sombist teisik ütleks sedasama!”. Kui ma ütlen, et ma tean, et mul on teadvus, siis juhitakse minu tähelepanu sellele, et minu sombist teisik ütleb sedasama. Kui ma ütlen, et minu uskumust õigustab minu vahetu elamus, märgitakse, et minu sombist teisik ütleb sedasama. Vastus sellele on jällegi: “Ja mis siis?” Kõige rohkem näitab see, et kolmanda isiku seisukohalt on minu sombist teisik minuga ​**[lk.199]**​ identne, mistõttu teie ei saa olla kindlad, et mul on teadvus; aga seda me teadsime ju nagunii kogu aeg. Kuid sellest ei järeldu ühelgi moel, et mul pole võimalik esimese isiku seisukohalt teada, et mul on teadvus. Esimese isiku seisukohalt vaadates oleme mina ja minu sombist teisik väga erinevad: minul on elamused, temal mitte. Seepärast on minul oma uskumuse kohta tõendid, temal aga mitte. Hoolimata sellest, et ta ütleb sedasama mis mina, tean ma (olgugi et sina ei pruugi selles kindel olla), et ma ei ole tema, sest minule on minu elamused tuttavad otseselt, esimeses isikus. See võib esialgu kõlada mõnevõrra paradoksaalselt, kuid tegelikult on jutt ilmselgest asjast: tänu oma teadvuselamusele me teame, et meil on teadvus.

Isegi kui keegi ütleb, et minu sombist teisik usuks sedasama, mida usun mina, ei tee see kuidagimoodi esimese isiku seisukohalt usutavaks skeptilist hüpoteesi, et ma olen sombi. Sellise vastuväite aluseks on võib-olla vaikiv eeldus, et minu episteemilist situatsiooni määratlevadki esmajärjekorras uskumused, nii et kui on olemas olukord, kus ma usun täpselt sedasama, mida ma usun praegu, siis on see olukord tõenduste mõttes samaväärne minu praeguse olukorraga. Kuid see on mõistagi väär. Minu elamusi puudutavate uskumuste tõendusmaterjal on midagi palju algsemat kui need uskumused ise. Primaarne on elamus, uskumused on aga suuremalt jaolt sekundaarne fenomen.

Samuti tuleb meeles pidada, et uskumuse mõiste, millest me siin lähtume, on deflatiivne (s.t. funktsionaalne), nii et kui öelda, et minu sombist teisik usub kõike sedasama mis mina, siis on see ikkagi vaid väide selle kohta, mis mul on temaga ühine kolmanda isiku seisukohalt: ta on sooduv esitama samalaadseid ütlusi, tegema samalaadseid järeldusi jne. Sellega pole öeldud midagi selle kohta, kuidas asjad seestpoolt paistavad. Tunne, nagu annaks viide sellele, et sombil oleksid samad uskumused, paikapidava vastuväite, võib tulla sellest, et oponent lähtub uskumuse inflatiivsest käsitusest, mille järgi uskumus on vähemalt mingis osas elamuslik fenomen. Ainult selles mõttes võiks uskumuste identsus teha olukorrad esimese isiku seisukohalt eristamatuks; kuid muidugi ei ole selles mõttes mingit põhjust arvata juba sedagi, et sombil oleksid samad uskumused mis minul.

Kõige selle kokkuvõtteks võib öelda, et argumendid, mis kõnelevad iseenda tundmisest, ei anna põhjust lükata tagasi seisukohta, mida ma siin kaitsen. Kui keegi võtab teadvust tõsiselt, siis ongi tal juba mõjuv põhjus võtta omaks niisugune teadvuse epistemoloogia, mis neil argumentidel hambad murrab. Kuigi ahvatlevaid argumente, mida siin esitada võiks, on palju, ei paista ükski neist kriitilisele uurimisele vastu pidavat.

Muidugi tuleb teadvuse esimese isiku epistemoloogia täpsustamiseks veel väga palju teha. Parimal juhul olen ma siin visandanud vaid nende asjade üle mõtlemiseks sobiva mudeli pelgad kontuurid. Paljud küsimused on veel lahendamata. Eelkõige oleks soovitav analüüsida järgmist: mil viisil nimelt elamus uskumusele kinnitust pakub; millised muud tegurid ​**[lk.200]**​ elamusi puudutavate uskumuste põhjendamisel olulist rolli mängivad; millistel tingimustel on inimesel elamust puudutavaks uskumuseks täielik alus; jne. Kõik need on olulised küsimused, mis väärivad teadvuse epistemoloogia uurimisel põhjalikku käsitlemist. Kuid kõik nad kuuluvad meie ülesande juurde, mida ei paista olevat põhjust lahendamatuks pidada. Oluline on see, et argumendid, mida iseenda tundmine elamuse mittereduktiivse käsitluse vastu pealtnäha pakub, ei pea paika.

### 6. Mäluargument\*

Teiseks argumendiks elamuse põhjusliku või seletusliku irrelevantsuse vastu on vastuolu tõsiasjaga, et me mäletame oma elamusi. Mulle kahtlemata tundub, et mulle meenuvad sageli minu kunagised elamused, näiteks riidekapi vürtsikas naftaliinilõhn lapsepõlvest või eriliselt ere oranž elamus eile õhtul päikeseloojangut vaadates. Kuid sageli kinnitatakse, et selleks, et midagi mäletada, tuleb olla sellega sobivas põhjuslikus suhtes; seda seisukohta nimetatakse mõnikord *põhjuslikuks mäluteooriaks*​. Kui aga elamused on minu psühholoogilisele talitlusele põhjuslikult irrelevantsed, siis tuleb välja, et minu kunagistel elamustel ei ole ühegi minu praeguse seisundiga põhjuslikku seost. See tähendaks, et meil ei olegi võimalik oma elamusi mäletada.

Põhjuslik mäluteooria ei ole siiski kivisse raiutud. Tema analüüsid on väga paljudel juhtudel usutavad. Põhjuslik teooria töötab hästi mitmel alal — näiteks seletab ta hästi teadmist —, kuid see ei tähenda, et teda on kohane rakendada kõikjal.​[[253]](#footnote-254) Ja just elamuse puhul pole tema sobivus kuigi ilmne. Elamuste mäletamise​[[254]](#footnote-255) seletamiseks ei tarvitse põhjuslikud teooriad olla nii sobimatud nagu teadmuse seletamiseks, kuna pole kahtlust, et meie suhe meenuva elamusega toimub millegi vahendusel, ja võib arvata, et selle vahenduse oluliseks osaks on mingi põhjuslik ahel. Kuid see ei tähenda, nagu seletakski see põhjuslik ahel kõik ära.

Elamuste puhul, mida me mäletame, on kindlasti olemas põhjuslik seos psühholoogia tasandil: algse elamuse aegne kognitiivne alusseisund on põhjuslikus seoses kognitiivse seisundiga meenutamise ajal. Ja tundub, et seesugusest põhjuslikust seosest võiks elamuse mäletamiseks piisata. Põhjuslik seos võib olla näiteks varasema fenomenaalse uskumuse ja hilisemate uskumuste vahel; ja kui see, mida ma eelmises jaos ütlesin, on õige, siis võib seda algset uskumust lugedagi teadmiseks, kuna tema aluseks on tutvus elamuse endaga. Sedaliiki põhjuslik seos tutvusega põhjenduvalt uskumuselt hilisemale ​**[lk.201]**​ uskumusele näib olevat täiesti piisav, et pidada seda hilisemat uskumust mäletamiseks. Seega on nähtavasti piisavalt alust arvata, et ühe või teise elamuse mäletamine ei eelda põhjuslikku seost selle elamusega.

Mida lugeda mälestuseks ja mida lihtsalt tõeseks ja õiguspäraseks uskumuseks mineviku kohta — see on muidugi suures osas semantilise otsuse küsimus. Oluline on, et mittereduktiivne teooria võib päästa oma näo, osutades mehhanismile, mille varal kujunevad tõesed uskumused varasemate elamuste kohta. Senikaua kui mittereduktiivne teooria seda suudab, jääb iga mälule toetuv argument tema vastu hambutuks. Kui keegi ikka veel visalt kinnitab, et mälu eeldab põhjuslikku seost objektiga, siis võib talle vastuseks öelda lihtsalt seda, et meil ei ole elamustest mälestused, vaid “pseudomälestused”; ka nii ei kaota me midagi olulist. Kuid nii või teisiti tundub mulle, et põhjuslikust seosest vastava algse psühholoogilise seisundiga piisab täiesti, et pidada neid uskumusi mälestusteks.

### 7. Kolmas argument: referendid\*

Kolmas argument teadvuse nii põhjusliku kui ka​[[255]](#footnote-256) seletusliku irrelevantsuse vastu on, et need on vastuolus meie võimega oma teadvuselamustele viidata. Kahtlemata on meil võime oma teadvuselamustest mõelda ja neist kõnelda — seda olen minagi teinud kogu käesolevas raamatus. Kuid mõnikord kinnitatakse, et viide entiteedile eeldab põhjuslikku seost selle entiteediga; seda väidet nimetatakse *põhjuslikuks osutusteooriaks*​. Kui see oleks tõsi, siis ei oleks elamustele viitamine nende põhjusliku irrelevantsuse tõttu võimalik.[[256]](#footnote-257)

Ei tundu olevat mingit põhimõttelist põhjust, miks peaks osutus mingile entiteedile eeldama põhjuslikku seost selle entiteediga. Põhjuslik seos kaasneb osutusega sageli, kuid ei ole ilmne, et see peab alati nii olema. Mingile entiteedile viitamiseks on vaja vaid seda, et meie mõistetel oleksid intensioonid (eelkõige primaarsed intensioonid), mida see entiteet rahuldada võiks. Näiteks minu mõistel “suurim kinnistäht universumis” on primaarne intensioon, mis osutab igas keskpunktiga maailmas mingile referendile. Aktuaalses maailmas osutab see intensioon ühele kindlale tähele ​*S* — olgu mul sellega põhjuslik seos või mitte —, ja seega on ​*S* selle mõiste referent. Niipea kui on olemas primaarne intensioon, mida võiks rahuldada mingi entiteet aktuaalses maailmas, on meil kõik osutuseks vajalikud komponendid olemas.

Tegelikult saabki paljude meie mõistete primaarseid intensioone kirjeldada põhjuslikult: kindlas keskpunktiga maailmas osutavad nad sobivale entiteedile, millel on keskpunktiga põhjuslik seos. Sellest ideest ongi välja kasvanud nn. põhjuslik osutusteooria. Kuid pole mingit põhjust, miks peaks primaarne intensioon alati töötama just nii. Leidub palju funktsioone, mis osutavad hüpoteetilises keskpunktiga maailmas entiteedile, millel pole keskpunktiga ​**[lk.202]**​ mingit põhjuslikku seost. Niisugused funktsioonid võivad anda väga häid primaarseid intensioone, millel on täiesti korralik referent.

Ja veel, primaarse intensiooni olemasolu ei sõltu — isegi juhtudel, kus tema kirjeldus toetub põhjuslikkusele — ühelgi moel põhjuslikust seosest referendiga. Primaarne intensioon ei sõltu sellest, mis toimub aktuaalses maailmas. Põhjuslik seos võib tõepoolest sageli mängida rolli primaarse intensiooni kindlakstegemisel ühes või teises maailmas, kuid see on midagi hoopis muud kui mängida rolli primaarse intensiooni enda määratlemisel. Tõepoolest, leidub mõisteid, millel ei olegi referenti (näiteks “jõuluvana”), kuid primaarne intensioon on olemas — nimelt intensioon, mis osutaks referendile, kui selguks, et maailma asjade seis on niisugune, nagu selleks tarvis.

Sageli mängib põhjuslik seos referendiga rolli mõiste loomisel ja seega primaarse intensiooni moodustamisel. Võiks väita, et isegi “universumi suurima kinnistähe” puhul aitavad põhjuslikud seosed maailmaga mingil määral kujundada alusmõisteid, millest see liitmõiste moodustub. Kuid jällegi ei näi olevat mingit sügavat põhjust, miks peaks primaarse intensiooni olemasolu lausa nõudma põhjuslikku seost vastava objektiga. Isegi vaadiajul võib, hoolimata oma põhjuslikust isolatsioonist, olla mõisteid, millel on primaarne intensioon olemas (kuigi enamikule neist intensioonidest ei vastaks meie maailmas miski[[257]](#footnote-258)). Ka siin ei sõltu primaarse intensiooni koosseis säärastest põhjuslikest seostest.

Põhjuslikkus on nii paljudes meie mõistetes kesksel kohal ühel väga loomulikul põhjusel: üldiselt me osutame sellele, mida me teame​, ja meie teadmised puudutavad tavaliselt asju, millel on meiega põhjuslik seos. Kuid, nagu juba nägime, on kaalukaid põhjusi mitte kohaldada põhjuslikku teadmusmudelit vähemasti teadvusele: siin on meie teadmine otsesemat liiki. Seepärast ei pea me teadvusele osutama intensiooni kaudu, mis osutaks millelegi, millel on põhjuslik seos maailma keskpunktiga, vaid võime viidata talle intensiooni kaudu, mis osutab millelegi, millega see keskpunkt on otseselt tuttav.

Nii või teisiti on oluline see, et (1) minu mõistel “teadvus” võib primaarne intensioon olemas olla, olgu mul mõiste referendiga põhjuslik seos või mitte (sest primaarne intensioon ei sõltu kunagi sellisest põhjuslikust seosest); ja (2) primaarne intensioon võib osutada referendile olenemata sellest, kas referendiga põhjuslikku seost on või ei ole (sest pole põhjust, miks peaks primaarne intensioon osutama oma referendile just põhjusliku seose alusel). Intensioon annab väga hea funktsiooni keskpunktiga maailmade hulgast nende maailmade omaduste hulka; meie maailmas leidub midagi, mis sellele intensioonile vastab, ja seega on minu mõistel referent olemas. Teadvus on, nagu oleme näinud, enam-vähem algmõiste (nagu võib-olla aeg ja ruum), ja seega pole lootust kirjeldada talle vastavat intensiooni ​**[lk.203]**​ üksikasjalikult nii, nagu seda saaks teha mõne teise mõiste puhul; kuid pole põhjust arvata, et teadvuse mõiste​[[258]](#footnote-259) ei ole väga hästi suuteline osutama referendile mõnes maailmas.

### 8. Fenomenaalsete uskumuste sisu\*

Isegi kui nõustuda, et omadusdualism on ühitatav osutustega teadvusele, seab meie fenomenaalsete mõistete ja uskumuste sisu meie ette ikka veel palju huvitavaid lahendamata mõistatusi. Kõigepealt kerkib küsimusi meie mõistete  — niihästi üldiste (nagu “teadvus”) kui ka spetsiifilisemate (nagu “punane elamus”) — intensiooni olemuse kohta: millele need mõisted konkreetses maailmas ikkagi osutavad? Mõtlema paneb ka meie mõistete sisu koosseis: kas sellesse kuulub ainult meie psühholoogiline loomus või meie psühholoogiline loomus koos fenomenaalse loomusega, ja missugune on kummagi roll? Mul ei ole neis asjus kindlaid arvamusi, kuid ma püüan neid küsimusi siinkohal vähemasti riivata.

Üks huvitav võimalus nende mõistatuste kallale asumiseks on küsida: kas minu fenomenaalsete uskumuste sisus on erinevusi võrreldes sombi omadega, ja kui on, siis milles need seisnevad? Ma ütlen nüüd jälle “uskumused”, mitte “otsustused”, sest küsimus on just selles, kas fenomenaalse uskumuse sisus leidub veel mõni element lisaks fenomenaalse otsustuse sisule, mis on, nagu me kokku leppisime, puhtalt psühholoogiline entiteet. Ma möönan vähemalt siinse arutelu raames, et sombil on uskumused olemas (kuigi on selge, et tema uskumustes pole midagi, mida poleks tema otsustustes). Küsimus on, kas tema ja minu uskumuste sisu vahel on erinevusi. Ja milles on erinevus — kui seda on — minu ja tema vastavate uskumuste tõesustingimuste vahel ning meie vastavate mõistete intensioonide vahel?

Üks võimalik vastus on, et meie uskumuste ja arusaamade sisud kattuvad täpselt. Selle vaate järgi on sombi mõistetel “teadvus” ja “punane elamus” samasugused primaarsed intensioonid nagu minu vastavatel mõistetel, ja tema uskumustel on samad (primaarsed) tõesustingimused. Näiteks osutab mõiste “teadvusega olend” ka sombil teadvusega olenditele vastavas keskpunktiga maailmas. Vahe on ainult selles, et sombi maailmas niisuguseid olendeid ei ole, või vähemasti ei ole tema selline olend. Seega on uskumuse “mul on teadvus” (primaarsed) tõesustingimused sombil samasugused nagu minul; ainus erinevus on see, et tema uskumus on väär, minu oma aga tõene.

Selle mõttekäigu üle otsustamiseks tuleb uurida mõningaid spetsiifilisemaid küsimusi. Esiteks, kas sombi kõneleb ebatõtt, kui ta ütleb: “Mul on teadvus”? Mõni võib-olla eitaks seda: sombi ütlust tuleb tõlgendada heatahtlikult, otsekui viitaks tema teadvusemõiste mingile temas realiseeruvale funktsionaalsele omadusele — ja sellisel juhul on tema ütlus tõene.[[259]](#footnote-260) Kuid sombi (või vähemalt minu sombist teisik) ​**[lk.204]**​ kinnitab kindlasti, et tema teadvusemõiste ei ole funktsionaalne: ta peab silmas omadust, mis ulatub tema eristamis-, kategoriseerimis-, raporteerimis- jne. võime piiridest kaugemale. Tundub, et pole põhjust tema sõnu mitte uskuda. Me võime loomulikult aktsepteerida ka tema sõnade “deflatiivset” tõlgendust, nii et ütlused nagu “ma tulin jälle teadvusele” võivad igapäevases kontekstis tõesteks osutuda; kuid vähemalt filosoofilistes kontekstides on arvatavasti mõistlik hinnata tema mõisteid kõrgema standardi järgi kui see, mida silmas peab tema, ja siis tuleb välja, et tema uskumused on väärad. Asi pole nii, nagu oleks ta kontseptuaalses segaduses, millesse saaks tuua selgust täpsema kontseptuaalse analüüsi teel. (Kui ta seda suudaks, suudaksin seda ka mina; kuid meie arutelu lähtub eeldusest, et minu mõtetes sellist kontseptuaalset segadust ei ole). Seega tundub, et mingis mõttes võib tema väiteid teadvuse kohta ikkagi ekslikeks pidada.[[260]](#footnote-261).

Sellest aga ei järeldu, et teadvuse mõiste intensioonid minul ja sombil kattuvad, vaid ehk ainult niipalju, et temal tähendab see mõiste umbkaudu “omadust, mis väljub kõigi füüsiliste ja funktsionaalsete omaduste piirest”. Kuid täisväärtuslikku teadvusemõistet tal ikkagi ei ole. Tegelikuks proovikiviks on küsimus, kas on olemas keskpunktiga maailmu, kus kõik relevantsed üksikasjad on identsed, ent sombi uskumuse tõeväärtus erineb minu uskumuse tõeväärtusest. Näiteks kuidas on asi siis, kui me mõlemad ajame juttu olendiga, kellel on teadvus? Mina ütlen: “Sul on teadvus,” ja minu jutt on tõene. Kui tema ütleb: “Sul on teadvus,” kas tema jutt on tõene? Mõni ütleks, et ei ole, sest tal puudub vahetu tutvus teadvusega, mis annaks talle täisväärtusliku mõiste. Teistsugune intuitsioon ütleb aga, et mõtet oleks ka jaataval vastusel.

Kõige asjakohasem oleks ehk vaadelda näiteid hüpoteetilise olendiga, kellel on mingi olemuslik omadus, mis pole ei struktuurne, funktsionaalne ega ka mitte fenomenaalne — kui niisuguseid omadusi saaks olla —, kuid mille suhted kõige muuga oleksid samasugused nagu fenomenaalsetel omadustel tavaliselt. Kui ma ütlen säärasele olendile: “Sul on teadvus”, siis ei ole minu jutt tõsi; aga kui seda ütleb minu sombist teisik ja kui tema jutt oli eelmisel juhul tõene, kas tema ütlus on ka seekord tõene? Kuna sombil puudub tutvus teadvusega, siis on raske näha, kuidas saaks sombi teadvusemõiste olla piisavalt spetsiifiline, et nende kahe juhu vahel vahet teha. Ja sellisel juhul ei küüni sombi arusaam teadvusest täisväärtusliku teadvusemõisteni.

Kui meie tutvus teadvusega mängib mingit rolli meie mõiste primaarse intensiooni konstitueerimises, siis tähendab see, et teadvuse mõistel on huvitavaid erinevusi teistest mõistetest, näiteks vee mõistest. Neil teistel juhtudel ei sõltu primaarne intensioon aktuaalsest referendist; kognitiivsel süsteemil, millel asjaomase mõiste referent on teine või puudub üldse, võib primaarne intensioon olla sama. Kuid teadvus ei ole tavaline referent: meie suhe temaga on vahetu ja ta on meie vaimse elu keskpunktis, mistõttu tal võib olla tavatult oluline roll primaarse intensiooni konstitueerimisel. Igal juhul jätan ma selle küsimuse lahtiseks.

​**[lk.205]**​ Aga kuidas on lugu spetsiifiliste fenomenaalsete mõistetega, näiteks punase elamuse mõistega? Siin on olukord mõnevõrra keerulisem, kuna selliste terminitega võib seonduda mitu mõistet. Üks võimalus neis mõistetes selgusele jõuda on vaadelda küsimust, kuidas tuleks kirjeldada invertspektriga indiviide — näiteks niisugust, kellel punase eseme vaatamine kutsub esile samasuguse elamuse nagu minul rohelise eseme vaatamine. Kui ma ütlen, et säärasel inimesel on (punast roosi vaadates) punane elamus, siis võib seda ütlust pidada mingis mitteranges mõttes tõeseks; selles mõttes tähendaks “punane elamus” enam-vähem “elamust, mida tavaliselt kutsuvad (indiviidis, kelle elamus see on) esile punased asjad”. Võib-olla leidub umbes sellisel viisil töötav punase elamuse mõiste avalikus keeles​[[261]](#footnote-262), ent nii või teisiti jätan ma selle küsimuse sinnapaika.

Loomulikum on ehk mõtteviis, mille järgi niisugune ütlus on väär: kõnealusel inimesel ei ole punane, vaid roheline elamus. Ühe võimalus sellise arusaama selgitamiseks on öelda, et minu mõiste “punane elamus” primaarne intensioon võiks olla umbes nagu “elamus, mida (minus) tavaliselt põhjustavad punased asjad”. Sellel mõistmisviisil on mõned huvitavad omadused: näiteks kui sinu ja minu spektrid on teineteise suhtes invertsed, siis osutab sinu mõiste “roheline elamus” sellele, mida mina nimetan punaseks elamuseks. Sellest järeldub, et kui sina ütled “Rohi põhjustab minus rohelisi elamusi”, siis võib sinu jutt olla tõene, samas kui mina ütlen “Rohi põhjustab sinus rohelisi elamusi”, siis on minu jutt väär. Niisugust mõtteviisi peetakse mõnikord maotuks[[262]](#footnote-263), kuid see on kõnealuse mõiste indeksikaalsuse loomulik tagajärg. (Sarnane fenomen tuleb esile ütlustega, kus figureerib sõna “mina”, ning ütlustega vee kohta, kui neid lausun mina ja kui seda teeb minu Teisikmaa-kaksik). Kui keegi tahab niisugust olukorda vältida, siis võib ta alati jääda ülalkirjeldatud “avalikuma” mõiste juurde[[263]](#footnote-264).

Selle mõiste — s.t. punase elamuse mõiste — primaarne intensioon peaks olema üsna lihtne, kui elamuse üldises mõistes on juba kokku lepitud. Sinu mõistel “punane elamus” võib primaarne intensioon olla sama mis minu omal, isegi kui meie spektrid on teineteise suhtes invertsed: mõlemad osutavad ühes kindlas keskpunktiga maailmas samadele entiteetidele (elamustele, mida maailma keskpunktis põhjustavad tavaliselt punased asjad), ehkki kuna keskpunktiga maailmad, milles me elame, on erinevad (minu maailmas olen keskpunktiks mina, sinu maailmas sina), siis on meie mõistete referendid muidugi erinevad. Seepärast võivad meie mõistetel olla ka erinevad sekundaarsed intensioonid: minu intensioon osutab kontrafaktilises maailmas punastele elamustele, sinu oma aga rohelistele. Isegi sombil võib primaarne intensioon olla samasugune nagu meil kahel, vähemalt kuivõrd tal üldse on mingi ettekujutus elamusest — kuigi temal ei osuta intensioon tema keskpunktiga maailmas muidugi millelegi ja tema punase elamuse mõistel puudub referent.

Sellega asi aga veel ei piirdu. Seesugune punase elamuse mõiste relatsiooniline tõlgendus on ikkagi suhteliselt perifeerne. Selle mõistega on suguluses veel üks mõiste​, mida relatsiooniline kirjeldus ei ammenda, ​**[lk.206]**​ kuid mis on võib-olla üldse kõige tähtsam mõiste kõigis neis diskussioonides — nimelt punase elamuse kvaliteedi mõiste. Ülaltoodud relatsioonilises iseloomustuses ei ole midagi, mis seda mõistet hõlmaks — seda tõendab asjaolu, et ülalkirjeldatud primaarne intensioon on ühildatav paljude erinevate kvaliteetidega. Invertspektriga inimesel oleks see intensioon samasugune, kuid mõiste, mis on sellest kvaliteedist minul — ütleme, ​*R* —, erineb minu invertspektriga paariku vastavast mõistest ​*G*​.

Esimesel pilgul võiks arvata, et seda kvaliteeti hõlmab minu mõiste “punane elamus” sekundaarne intensioon, nagu me seda äsja kirjeldasime. Nagu nägime, on minu invertspektriga paarikul punase elamuse mõiste sekundaarne intensioon minu omast erinev, vastavalt erinevatele omadustele, millele see mõiste meil kummalgi osutab. Kuid selle asjaolu fikseerimisest sekundaarses intensioonis ei piisa. Et seda mõista, pangem tähele, et teadmine punaste elamuste (s.t. elamuste, mida põhjustavad punased objektid) omaendi spetsiifilisest kvaliteedist on informatiivne, või teisiti öeldes, pole kaugeltki ​*a priori* ilmne, et punased elamused peaksid olema ​*R*​. See äratundmine kitsendab meie aktuaalse maailma mudelit: elamuse liik, mida põhjustavad punased asjad, oleks võinud olla teist- või kolmatsugune, aga ta on just niisugune. Ja seesugune informatiivsus eeldab primaarsete intensioonide erinevust: kui kahe mõiste primaarsed intensioonid ühtivad, siis on ​*a priori* tõsikindel, et nende mahud kattuvad.[[264]](#footnote-265).

Seda võib mõista ka, kui panna tähele, et kui Mary saab esimest korda tundma punast elamust, siis erineb see, mida ta teada saab, sellest, mida saab teada tema invertspektriga teisik, kellel on Mary punaste elamuste asemel rohelised elamused. Mary saab teada, et punased asjad põhjustavad niisuguseid elamusi, tema teisik aga, et naasuguseid. Nende maailmamudelid kitsenevad erineval viisil: Mary tunnistab nüüd üht keskpunktiga maailmade kogumit, teisik teist. Sellest järeldub, et kõnealuste omaduste primaarsed intensioonid peavad Maryl ja tema teisikul olema erinevad. Maryl osutab primaarne intensioon igas keskpunktiga maailmas ühtliiki (just niisugustele) elamustele, tema teisikul aga teistliiki (naasugustele) elamustele.

Roll, mida minu kvalitatiivne mõiste​ *R* suhtluspraktikas otseselt mängib, on väike. Selles mõttes sarnaneb see mõiste Wittgensteini “mardikaga toosis.”[[265]](#footnote-266) [[266]](#footnote-267) [[267]](#footnote-268) Minu invertspektriga teisikul on teine sarnane mõiste ​*“G*​”, kuid teised mõistavad teda nii, nagu ütleks ta sama, mida ütlen nende arusaamise järgi mina — eeldades, et nad ise on mõlemas stsenaariumis ühesuguses olukorras. Siin tulebki esile see “tabamatus”, millest ma kõnelesin [1. peatükis](#PARA.25-3): punaste elamuste külluslikust seesmisest olemusest hoolimata ei ole mul selle erinevuse edasiandmiseks võimalik öelda peaaegu midagi — ma võin talle vaid mitmesuguste relatsiooniliste omaduste kaudu osutada, eeldades, et teistel seostuvad nendega samasugused elamused. Seega tundub, et kommunikatiivset koormust kannab ​**[lk.207]**​ siin just relatsiooniline punase elamuse mõiste. Seda tabamatust võib võtta kaudse kinnitusena elamuse seletuslikule irrelevantsusele: tõsiasi, et meil ei ole kuigi palju öelda elamuse olemusliku kvaliteedi edasiandmiseks, haakub hästi asjaoluga, et see kvaliteet ei mängi kognitiivsete protsesside juhtimisel mingit otsest rolli.

(Kõnelda võib neist omadustest muidugi ikka, nagu minagi olen seda aeg-ajalt teinud kogu käesolevas raamatus. Ma võin väljendada oma fenomenaalseid uskumusi keeleliselt; asi on aga selles, et minu keel edastab teistele minu uskumuste kogu sisu ainult eeldusel, kui nad ka ise valdavad vastavaid kvaliteete ja kui nende kvaliteetide vahelised asjaomased suhted on samasugused.)

See on ilmne näide selle kohta, kuidas meie mõistete ja uskumuste sisu konstitueerib miski, mis jääb väljapoole meie füüsilise ja funktsionaalse struktuuri piire, mistõttu uskumuste sisu reduktiivne käsitlemine on võimatu.[[268]](#footnote-269) Minu invertspektriga teisik ja mina võime olla füüsiliselt identsed, kuid meie vastavad kvalitatiivsed mõisted on erinevad, kusjuures mitte ainult referentidelt, vaid ka primaarsetelt intensioonidelt. Meie ees on olukord, kus fenomenaalse uskumuse sisu konstitueerib fenomenoloogia ise; siin on see veelgi selgemini näha kui teadvuse mõiste puhul. Siin on vaieldamatult toimumas midagi väga huvitavat: mingi elamus, mida võiks pidada kvalitatiivse mõiste referendiks, pääseb mingil moel mõiste sisse ja konstitueerib tema mõtte (mille all tuleb siinkohal mõista primaarset intensiooni). See on midagi hoopis muud kui tavalised juhud, kus mõiste objekt võib mängida rolli sekundaarse, kuid mitte primaarse intensiooni konstitueerimises.[[269]](#footnote-270) Selle on võimalik ainuüksi sellepärast, et elamus on vaimu kese.

Näeme seega, et fenomenaalsetes uskumustes kätkeb tõepoolest midagi enamat kui fenomenaalstes otsustustes, vähemalt meie juhtudel. Võimalik, et see asjaolu aitab teadvuse epistemoloogia mõistmisele kaasa. Kõige spetsiifilisem sellise konstitutiivse suhte juht leiab aset siis, kui fenomenaalse mõiste ​*S* konstitueerimises osaleb ainult üks ​*S*​-elamus (just täpselt “seda” liiki elamus[[270]](#footnote-271)). Otsene konstitutiivne seos — viis, kuidas elamus nii-öelda mõiste sisse pääseb — võiks aidata meil taibata, kuidas uskumust, et mingi elamus on ​*S*​, saaks põhjendada selle elamuse endaga. Kindlasti seostub elamus niiviisi tihedalt mingi sedaliiki uskumusega, mida võiks pidada vastavas olukorras sobivaks. Ja kui seesugune spetsiifiline legitiimne fenomenaalne uskumus on olemas, siis võib mõista, kuidas sellest saaks tuleneda üldisemaid legitiimseid fenomenaalseid uskumusi (näiteks inimese uskumus, et tal on teadvus). Ma jätan selle küsimuse arutamise sinnapaika, sest siin on palju segast; kuid elamuste säärane suhe fenomenaalsete mõistetega pakub rohkesti mõtlemisainet.

(Võiks ehk välja pakkuda teesi, et uskumus, mille järgi mingi elamus on ​*S*​, kus​ *S*​ mõistet konstitueerib ülalkirjeldatud viisil elamus ise, seisab alati kindlal pinnal. See ei kaotaks alusetuid ​**[lk.208]**​ fenomenaalseid uskumusi, kuid need moodustuksid nüüd teistel juhtudel, kus mõiste ja elamuse vahel on teistsugune suhe — näiteks kui mingit mõistet, mida konstitueerib üks elamus (või üks elamuste kogum, või mingi väline kirjeldus) kohaldatakse teisele elamusele. Võib-olla suudaks see suhteliselt nõrk tees tabada tavapärastes “mittekorrigeeritavus”-teesides leiduvat tõepärast elementi, jättes samaaegu ruumi kõigile tavapärastele vastunäidetele; temast võiks saada isegi keskne lüli elamuse epistemoloogia üksikasjalikus käsitluses. Kuid selles ma ei ole kindel.)

Asjaolust, et minu uskumuses on element, mis minu sombist teisiku vastavas uskumuses puudub, võib samuti olla kasu kõigi ülalmainitud epistemoloogilisi argumente toestavate intuitiivsete kujutluste vääramisel. Ainult meie otsustused (funktsionaalses mõttes) on samad: ei ole tõsi, et ka juhul, kui meil elamust poleks olnud, oleksid meil ikkagi samad uskumused.[[271]](#footnote-272) Kuid uskumused ongi siin ju kõige tähtsamad. Me nägime, et selline uskumuste eristuvus ei ole epistemoloogiliste argumentide kummutamiseks vajalik (ja ülaltoodud aruteluls ma seda ka ei eeldanud), kuid aitab võib-olla siiski hajutada selle küsimusega seonduvaid kahtlusi.

Loomulikult hakkab meid nüüd huvitama, kui kaugele selline sisu elamuspõhine konstitueerimine ulatuda võib. Asjaolu, et see konkreetsete fenomenaalsete mõistete puhul toimub, annab tuge mõttele, et nii võiks olla ka teadvuse mõistega, mis on üldisem, kuigi endiselt jääb ebaselgeks, missugune on suhe selle mõiste kahe kuju — minu mõiste ja sombi mõiste — vahel. Edasi võiks küsida, kas elamus võib mängida rolli mittefenomenaalsete mõistete, näiteks väliste liigimõistete sisu konstitueerimisel, nagu on oletanud mõned filosoofid. Ei ole selge, kuidas selline laiendamine võiks välja näha, kuid võib-olla annab asjaolu, et elamus mängib seda rolli ühel juhul, mõningat kinnitust mõttele, et ta võib teha seda ka teistel juhtudel.

Igatahes ei ole siin midagi, mis eeldaks tingimata just põhjuslikku osutusteooriat. Seesuguse otsese seose võimaldajana, mille me leiame olevat elamustel fenomenaalsete mõistete primaarsete intensioonidega, oleks põhjuslik seos elamusega arvatavasti isegi kohatu: tavalistel juhtudel, kus põhjuslik seos on olemas (näiteks vee mõiste puhul), sellist seost ei ole. Vastupidi, tundub, et seesuguse konstitueeringu teeb võimalikuks meie vahetu tutvus elamusega. Seega ei ole põhjuslik seos elamusega vastavate primaarsete intensioonide konstitueerimiseks tarvilik, ning ammugi ei ole põhjuslikud seosed tarvilikud selleks, et primaarsed intensioonid osutaksid neil juhtudel referendile. Seega ei ole elamusele osutamiseks põhimõttelisi takistusi, isegi kui asuda omadusdualismi positsioonile.

Seega näib, et kuigi elamuse seletuslik irrelevantsus füüsilisele käitumisele võib esialgu tunduda intuitsioonivastasena, ei ole sellele ka tugevaid argumente vastu seada. Epistemoloogilised ja osutustest lähtuvad argumendid, mis võiksid ​**[lk.209]**​ tunduda tugevatena, osutuvad vaid probleemideks. Nende punktide uurimise käigus kerkib suur hulk huvitavaid küsimusi, ent, nagu me oleme kokkuvõttes näinud, on kaalukaid põhjusi arvata, et elamuse epistemoloogia ja semantika ei saa olla olemuselt põhjuslikud, vaid neid tuleb mõista hoopis teisiti. Ma olen siin veidi kõnelnud sellest, kuidas saaks neid asju mõista omadusdualistlikult positsioonilt. Nende küsimuste ammendav käsitlus nõuaks pikemat omaette uurimistööd, kuid ma loodan, et ka sellest, mida ma siin öelnud olen, piisab selgitamaks, et mittereduktiivne kontseptsioon annab nende küsimuste selgitamiseks loomuliku mudeli.

# ​**[lk.211]**​ III OSA: Teel teadvuseteooria poole

## ​**[lk.213]**​ 6. peatükk. Teadvuse ja kognitsiooni vaheline koherentsus

### 1. Tee mittereduktiivse teooria poole

Isegi kui teadvusele ei saagi leida reduktiivset seletust, võib teadvuseteooria loomine siiski võimalik olla. Me peame lihtsalt asuma otsima mittereduktiivset teooriat. Me võime loobuda eesmärgist püüda seletada teadvuse olemasolu ainuüksi millegi fundamentaalsema kaudu, tunnistada fundamentaalseks hoopis teadvuse enda ning püüda seletada, missugused on tema suhted kõige muuga maailmas.

Niisugune teooria sarnaneks põhimõtteliselt teooriatega, mida füüsika annab meile aine, liikumise või aegruumi kohta. Füüsikalised teooriad ei tuleta nende omaduste olemasolu millestki fundamentaalsemast, küll aga annavad olulisi ja üksikasjalikke selgitusi parameetrite ja nende omavaheliste seoste kohta, tänu millele meil on olemas rahuldavad seletused paljudele konkreetsetele fenomenidele, milles figureerivad mass, aeg ja ruum. Seda suudavad nad tänu lihtsale ja võimsale seaduste süsteemile, millest tuleneb mitmesuguste parameetrite abiga kõikvõimalikke spetsiifilisi fenomene.

Analoogiliselt on ka teadvuseteooria nurgakiviks mingi hulk teadvuse ja füüsikaliste süsteemide vahelist suhet reguleerivaid psühhofüüsikaseadusi. Me oleme juba jõudnud arusaamisele, et teadvus on füüsika suhtes looduslikult (kuigi mitte loogiliselt) superventne. Selle superventsuse peavad tagama psühhofüüsika seadused; nende seaduste kirjeldus ütlebki meile, mil viisil teadvus füüsikalistest protsessidest sõltub. Kui süsteemi füüsikalised faktid on teada, siis saab nende seaduste abiga järeldada, missugused teadvuselamused, kui üldse mingisugused, selle süsteemi juurde kuuluvad. Need seadused on universumi aluskorra osana võrdsel jalal füüsikaseadustega.

​**[lk.214]**​ Sellest järeldub, et kuigi see teooria ei seletaks teadvuse olemasolu — selles mõttes, et ta ei vastaks küsimusele, miks teadvus eksisteerib —, suudaks ta seletada teadvuse spetsiifilisi juhte, lähtudes füüsilisest alusstruktuurist ja psühhofüüsika seadustest. See on jällegi analoogne seletustega füüsikas: seletamaks, miks konkreetsed mateeria või liikumise juhtumid on niisugused, nagu nad on, viitavad nad üldistele alusprintsiipidele ühenduses mingite lokaalsete omadustega. Füüsika alusseaduste mõistetes saab seletada igatliiki makroskoopilisi füüsikalisi fenomene; samamoodi võiks oodata, et teadvuseteooria suudab seletada kõikvõimalikke “makroskoopilisi” elamusfenomene psühhofüüsika seaduste mõistetes.

Nendes seadustes tarvitse olla midagi eriliselt üleloomulikku. Nad moodustavad, nagu füüsikaseadusedki, ühe osa looduse aluskorrast. Kindlasti on neis midagi elementaarset​. Mingil tasandil tuleb seadusi lihtsalt tõdedena aktsepteerida ja neid mitte enam edasi seletada. Kuid nii on see ka füüsikas: mingist punktist edasi tunduvad looduse kõige põhilisemad seadused alati meelevaldsetena. Just sellepärast nad ongi loodusseadused, aga mitte loogikaseadused.

Teaduses ei saa kunagi midagi tasuta: alati tuleb kusagil võtta midagi aksioomiks. On tähelepanuväärne, et enamikus teaduse valdkondades tuleb aksioomideks pidada viimselt ainult füüsikaseadusi ja võib-olla mõningaid rajatingimusi​. Kuid pole mingit põhjust, miks peaks füüsika seadustel olema selles suhtes absoluutne eelisseisund. Kui teadvuse uurimisel selgub, et füüsikaprotsesside ja teadvuse vahelise seose üks või teine aspekt tuleb võtta aksioomiks, siis nii ongi. See on teooria hind.

Muidugi tahame me hoida aksioomide arvu siiski võimalikult väikesena. Lõplik teooria ei jäta seost tasemele, kus tohutu hulga keeruliste füüsiliste seisundite ja nendega kaasnevate elamuste puhul öeldakse vaid, et “ajuseisund ​*X* põhjustab teadvusseisundi ​*Y*​”. Ta süstematiseerib seda seost seletusliku alusskeemi abil​[[272]](#footnote-273), näidates ära lihtsad alusseadused, millede varal seos töötab. Ei taha ju füüsika olla lihtsalt hulk tähelepanekuid kõikvõimalike objektide asendite, kiiruste ja laengute kohta eri aegadel, vaid ta süstematiseerib need tähelepanekud ja näitab, kuidas need tulenevad võimalikult lihtsatest ja võimsatest alusseadustest. Nii peaks see olema ka teadvuseteooriaga. Me peame püüdma seletada teadvuse superventsust füüsika suhtes võimalikult lihtsa seadustiku abil.

Meie otsingute lõppeesmärgiks on fundamentaalsete seaduste süsteem. Füüsikud otsivad alusseadusi, mis oleksid nii lihtsad, et mahuksid T-särgi rinnaesisele; teadvuseteoorialt tuleks oodata sama. Mõlemal juhul huvitab meid universumi alusstruktuur ja meil on põhjust arvata, et see on väga lihtne. Ent alusseaduste avastamine on võib-olla pigem kauge eesmärk. Füüsikas leiti kõigepealt makroskoopilisi reeglipärasid iseloomustavad seadused, alles ​**[lk.215]**​ hiljem jõuti fundamentaalsete seadusteni. Ka teadvuseteoorias leitakse arvatavasti kõigepealt mittefundamentaalsed seadused, mis iseloomustavad füüsikaliste protsesside ja teadvuselamuse vahelisi suhteid üsna kõrgel tasemel. Ent isegi mõni säärane kõrgtaseme printsiip võib esialgu olla seletuste mõttes küllalt suur saavutus, nii nagu termodünaamika printsiipideski oli kasu ammu enne statistilise mehhaanika alusprintsiipideni jõudmist. Ja need kõrgtaseme seadused lõpuks avastatud, hakkavad nad seadma rangeid kitsendusi kõigile fundamentaalsele alusseadusele, juhtides nõnda meie edasist liikumist lõpliku teooria poole.

Kui meil on füüsika ja teadvuse fundamentaalsed teooriad lõpuks käes, oleme võib-olla jõudnud millenigi, mida võib õigusega pidada “kõige teooriaks”. Füüsika alusseadused seletavad ära füüsikaliste protsesside iseloomu, psühhofüüsika seadused seletavad ära nendega kaasnevad teadvuselamused, ja kõik muu tuleneb neist kahest.

Muidugi võib sellise teooria otsimise käigus juhtuda midagi, mis muudab meie ettekujutust lõplikust teooriast. Näiteks võib juhtuda, et me leiame kõikehõlmavad seadused, mis ühendavad füüsika ja teadvuse fenomenid mingi veel üldisema teooria alla, samamoodi nagu me leidsime teooria, mis ühendas elektri ja magnetismi, või nagu füüsikud otsivad praegu teooriat, mis ühendaks kõik füüsikalised alusjõud. Pole võimatu, et juhtub midagi veelgi üllatavamat. Kuid meie praegune mudel annab vähemalt lähtepunkti teooria otsimiseks, mis annaks meile esimese lähenduse ning trampliini, mis lubaks jõuda järgnevate, radikaalsemate teooriateni.

#### Kuidas ehitada üles teadvuseteooriat?

Selle peale võidakse öelda, et kõik see metafüüsiline hiilgus on ju tore, aga kuidas asi praktikas välja näeks? Eelkõige: kuidas leida psühhofüüsikaseadused, mis konstitueerivad teadvuseteooria? Seisab ju teadvuseteooria ees tohutu raskus, mis füüsikalisi teooriaid ei vaeva, nimelt andmete puudumine. Kuna teadvus ei ole katsesituatsioonides otseselt vaadeldav, siis ei ole meil võimalik korraldada lihtsalt katseid mitmesuguste füüsikaliste protsessidega kaasnevate elamuste mõõtmiseks, et nii- või naasuguseid psühhofüüsikalisi hüpoteese kinnitada või kummutada. Kas ei või olla nii, et isegi kui need seadused on olemas, jäävad nad igavesti teadmata? Tõepoolest võibki tunduda, et mittetestitavuse tõttu on iga teadvuseteooria, millega me võiksime välja tulla, pseudoteaduslik.

Siin on, mille pärast muretseda: just sel põhjusel ongi teooria loomisega teadvuseteaduses raskem algust teha kui füüsikas. Ent see ei tähenda veel, et teadvuseteooria otsimisest tuleks üldse loobuda. Alustuseks on meist igaühe valduses ligipääs rikkalikule andmeallikale iseenda juhul. ​**[lk.216]**​ Meil on üksikasjalikud ja spetsiifilised teadmised omaenda teadvuselamustest ning on teadmisi ka neid kandvatest füüsikalistest protsessidest, ja seega on meil suur hulk reeglipärasid omast käest võtta. Neist reeglipäradest lähtudes saab juba mingil moel abdutseerida võimalikult lihtsad alusseadused, millest nad võiksid tuleneda. Praeguseks ei ole meil ainsatki seaduste süsteemi, mis säärase ülesandega toime tuleks, ja see tähendab, et see ei ole sugugi triviaalne kitsendus. Võib väga hästi selguda, et leidubki ainult üks küllalt lihtne seaduste kogum, mis annab õiged tulemused; säärasel juhul oleks meil põhjust uskuda, et need seadused moodustavadki ühe osa õigest teooriast.

Võidakse öelda ka, et teooriaid, mida annab ühitada esimese isiku andmetega, on väga palju ja väga erinevaid: solipsistlikest (mille järgi teadvus on ainult minul) panpsühhistlikeni (mille järgi teadvus on kõigel), biokeemilistest (mille järgi teadvuse sünniks on tarvilik mingi kindel biokeemiline keskkond) informaatilisteni (mille järgi teadvuse sünniks piisab õigest informaatilisest keskkonnast), samuti kõikvõimalikud ekstravagantsed teooriad nagu see, mille järgi inimestel on teadvus ainult paarituarvulistel aastatel. (Praegu on 1995. aasta.) Kuidas oleks võimalik mõndki neist teooriatest välistada, kui meil pole võimalik teiste vaimudesse piiluda, et nendel teavuselamusi mõõta?

Kõik need teooriad on andmetega loogiliselt ühildatavad, kuid sellest ei piisa, et neid veenvaks teha. Näiteks solipsistlikud teooriad teeb äärmiselt ebapraktiliseks nende täielik meelevaldsus (miks peaks teadvus olema just ainult sellel isikul?) ja suur aegruumiline heterogeensus (minu teadvuselamusel on süstemaatilised seosed minu füüsilise struktuuriga, kuid minu füüsilisel teisikul, kes paikneb kusagil mujal, ei ole üldse teadvust). Väga paljudes valdkondades mängivad meie teooriate kujundamises lisaks empiirilistele tõenditele rolli ka kõikvõimalikud tõepärasuskaalutlused. Mõelgem näiteks sellele, et me tunnistame evolutsiooniteooriat, mitte aga teooriat, mille järgi maailm loodi koos kõigi meie mälestustega ja kõigi fossiilsete ladestustega viiskümmend aastat tagasi. Ka kvantmehhaanikas eelistame me mõningaid lihtsaid teooriaid teistele, mis on empiiriliselt samaväärsed, kuid väga keerukalt ja meelevaldselt kokku klopsitud. Teooria loomiseks ei piisa empiirilistest tõenditest; on ka muid kaalutlusi, teiste seas tõepärasus, lihtsus ja esteetilisus.

Eelkõige ei saa üle hinnata lihtsuse rolli. Ilma selle kitsenduseta jääks igasugune teaduslik teoretiseerimine masendavalt laialivalguvaks. Igale teaduslikule teooriale on ju väga hõlpus leiutada alternatiivne hüpotees, mis on temaga empiiriliselt samaväärne, kuid täiesti *ad hoc*​. Kuid mitte keegi ei võtaks säärast hüpoteesi omaks just tema tarbetu keerukuse pärast. See tähendab, et niipea kui me leiame lihtsad alusseadused, mis on ühitatavad olemasolevate andmetega, on meil tõsist põhjust loobuda keerulisematest alternatiividest.

On ka teisi tõepärasuskitsendusi, mis võivad teadvuseteooria loomisel tublisti edasi aidata. Kõige ilmsem neist on põhimõte, millele me toetume alati, kui ​**[lk.217]**​ võtame teadvuselamuse kriteeriumiks kellegi sõnalise ütluse: põhimõte, mille järgi inimeste ütlused oma elamuste kohta kajastavad üldjoontes õigesti nende elamuste sisu. Me ei saa selle printsiibi paikapidavust tõestada, kuid ta on juba ette palju usutavam kui alternatiiv. Mingil määral toetub selle printsiibi tõepärasus järeldustele, mida me teeme iseenda juhu põhjal, kuid seda võib pidada metodoloogiliseks kitsenduseks ka teadvuseteooria väljatöötamisel. Kui see põhimõte osutuks täielikuks eksituseks, oleks olukord täiesti lootusetu: maailm minetaks igasuguse ratsionaalsuse ja teadvuseteooria loomine käiks meile üle jõu. Teooriate väljatöötamisel eeldame me alati, et maailm on mõistuspärane — et planeedid ei ilmu koos kõigi fossiilikihtidega eikusagilt ning et maailmas ei valitse veidrad ja keerulised seadused, mis reprodutseerivad täpselt neidsamu ennustusi, mida saab teha lihtsate seaduste põhjal. Vastasel korral on kõik lubatud.

Seesugused tõepärasuseeldused seavad teadvuseteooriale ühe väga kasuliku kitsenduse ning annavad lisaks ka rikkaliku andmeallika isegi kolmanda isiku seisukohalt: selleks, et teada saada, kas kellelgi on mingile stiimulile vastav teavuselamus, tuleb seda lihtsalt tema käest küsida! See põhimõte võimaldab meil teha palju tugevamaid järeldusi teadvuselamuste seoste kohta oma füüsiliste kandjatega. See eeldus ongi niivõrd tõepärane, et teadlased toetuvad sellele alatasa ja pole palju neid, kellele tuleks mõttesse seda kahtluse alla seada. On ka teisi sarnaseid eeldusi, mis võivad kasulikku rolli mängida — näiteks põhimõte, et inimeste mälestused oma elamustest ei ole üldjoontes radikaalselt ekslikud. Loomulikult võib neist printsiipidest olla aeg-ajalt erandeid; kuid me vähemasti eeldame, et raportid ja mälestused kajastavad elamusi õigesti, kui pole kaalukaid põhjusi uskuda vastupidist.

Tõepärasuseelduste hulka võiksid ehk kuuluda veel järgmised: fundamentaalseadused on ajas ja ruumis homogeensed; teadvuselamus sõltub ainult organismi seesmisest füüsilisest seisundist; suure tõenäosusega ei kajastu teadvuselamuses juhuslikud tegurid — näiteks molekulide jaotus neuronites —, kui nad ei mõjuta neuroni toimimist; jne. Muidugi on loogiliselt võimalik, et mõni neist eeldustest on ekslik, ent kuna puuduvad põhjused uskuda vastupidist, siis on mõistlik eeldada, et nad on kõik tõesed. Üheskoos seavad need tõepärasuseeldused teadvuseteooriale tugevaid kitsendusi ja võivad meid selle loomisel olulisel määral aidata.

Aga mida teha murega teadvuseteooria mittekontrollitavuse pärast? See mure hakkab endast tõsiselt tunda andma ainult juhul, kui peaks selguma, et leidub kaks võrdselt lihtsat teooriat, mis mõlemad sobivad suurepäraselt andmetega ja rahuldavad kõiki asjakohaseid tõepärasuskitsendusi. Seda ei pruugi juhtudagi: võib esile kerkida üksainus teooria, mis on kõigist alternatiividest selgelt üle. Kui aga ilmub kaks võrdselt head teooriat, siis seisab meie ees raske valik, kuid isegi säärasel juhul lubavad meie kitsendused ​**[lk.218]**​ jõuda üpris hea arusaamiseni teadvusest. Igal juhul on mure mittetestitavuse pärast ilmselgelt ennatlik, senikaua kui meil pole ainsatki teooriat, mis suudaks teadvuse fenomenidega vähegi rahuldavalt toime tulla.

Loomulikult tähendab säärane esimese isiku andmetele ja tõepärasuskitsendustele toetumine seda, et teadvuseteoorial hakkab olema spekulatiivne iseloom, erinevalt teooriatest enamikus teistes teadusvaldkondades. Kuna range intersubjektiivne kontroll pole võimalik, siis ei või me kunagi olla päris kindlad, et meie teooriad on õigel teel. Sel põhjusel ei hakka teadvuseteadusel tõenäoliselt kunagi olema seda empiirilist garantiid, mis on olemas teistel teadustel, ja enamik kõige kangekaelsemaid teadlasi jääb temast alatiseks kõrvale. Kuid teadvus on niivõrd keskne fenomen, et parem on mõista teda mingilgi määral kui üldse mitte: kui suudetakse välja töötada mõistlik teadvuseteooria, mis osutub paremaks kui kõik konkureerivad teooriad, siis on see oluline saavutus isegi juhul, kui talle ei õnnestugi leida absoluutselt veenvat kinnitust. Niisuguses paadis me juba kord oleme, kui püüame mõista maailma: me võtame materjalid, mis võtta on, ja nendega töötamegi.

Käesolevas ja kahes järgnevas peatükis astun ma mõned esimesed sammud teadvuseteooria loomise teel. Käesolev ja [järgnev](#ORG.7) peatükk arutlevad võimalike mittefundamentaalsete psühhofüüsikaseaduste üle ning püüavad põhjendada mõningaid printsiipe, mis väljendavad kõrgtasemelisi reeglipärasid teadvuse sõltuvuses füüsikalistest protsessidest. [Ülejärgmises peatükis](#ORG.8) asun spekuleerima fundamentaalsete seaduste üle, millele need reeglipärad toetuvad. See kõik on väga esialgse iseloomuga, kuid millestki tuleb alustada.

### 2. Koherentsusprintsiibid

Kõige paljutõotavam on alustada teadvuseteooria väljatöötamist tähelepanu koondamisega teadvuselamuse märkimisväärsele koherentsusele meie kognitsiooni struktuuriga. Fenomenoloogia ja psüühika ei hõlju teineteisest lahus, nende vahel on süstemaatilised vastavused. Teadvuse ja kognitsiooni vahel on palju seaduspäraseid seoseid, millega teadvuseteooriale saab esialgse hoo anda. Seda suhet uurima asudes tuleks koondada tähelepanu kõigepealt fenomenaalsetele otsustustele. Need otsustused kuuluvad psühholoogiasse, kuid seonduvad tihedalt fenomenoloogiaga ning aitavad sellistena nende kahe valdkonna vahele silda rajada. Mõeldes fenomenaalsetele otsustustele ja nende toimimisele meis endis, võime fenomenaalse maailma ühendamiseks psühholoogia sfääriga välja pakkuda terve hulga printsiipe.

Kõige ilmsem neist printsiipidest on see, mida ma [mainisin 1. jaos](#PARA.121-6): meie teist järku otsustused teadvuse kohta on üldjoontes õiged. Me võime nimetada seda *usaldusväärsusprintsiibiks*​. Kui ma moodustan otsustuse, et mul on ​**[lk.219]**​ mingi kuulmisaisting, siis tavaliselt mul ongi see kuulmisaisting. Kui ma arvan, et ma elasin äsja läbi valu, siis tavaliselt ma elasingi äsja läbi valu. On olemas ka vastupidine põhimõte, mida võiksime nimetada *tuvastatavusprintsiibiks*​: kui meil on mingi elamus, siis suudame me tavaliselt moodustada tema kohta mingi teist järku otsustuse. Muidugi libisevad paljud elamused meie tähelepanust mööda, kuid tavaliselt oleme me võimelised neid märkama: elamuses, mis oleks meile põhimõtteliselt märkamatu, oleks midagi õige kummalist.[[273]](#footnote-274)

Minu visandatud printsiibid ei ole absoluutsed. Meie teist järku otsustused on mõnikord ekslikud, demonstreerides seega erandjuhte, kus usaldusväärsusprintsiip ei kehti. Seda võib juhtuda tähelepanematusest (kui mul on mõtted hajevil, võin ma uskuda, et mul oli äsja valuelamus, tegelikult aga oli elamuses ainult vali müra), eksimisest asjakohastes kategooriates (näiteks kui ma nimetan tulipunast elamust ekslikult kastanpruuniks), vaimuhaigusest või neurofüsioloogilisest häirest (nagu näiteks juhtudel, kus patsiendid salgavad, et nad on pimedad​[[274]](#footnote-275), ning annavad oma elamuste kohta valeütlusi) ning mitmesugustel muudel põhjustel. On ka vastupidiseid olukordi: elamused võivad jääda märkamata, kui nad leiavad aset siis, kui inimene magab, või vilksatavad mööda nii kiiresti, et ta ei jõua neid tähele panna. Kuid siiski hõlmavad need printsiibid vähemalt olulisi reeglipärasid. Tüüpilistel juhtudel on teist järku otsustused enamasti õiged ja elamust on tavaliselt võimalik märgata. Need reeglipärad ei ole eranditeta seadused, kuid nad peavad liiga sageli paika selleks, et olla vaid pelk kokkusattumus. Siin on midagi süstemaatilist.

Ma ei hakka neid väiteid üksikasjalikult põhjendama, sest minu huvi keskmes on teistsugused koherentsusprintsiibid. Kuid esimese isiku juhtumi vaatlus näitab, et vähemalt neil korril on need printsiibid usutavad ning et homogeensus- ja lihtsusprintsiibi varal saab neid loomulikul teel laiendada teistele juhtudele. Mõningane kaal on ka sellel, et terve mõistus kiidab need printsiibid heaks; muidugi võime me terve mõistuse otsusest mööda minna, kui selleks on mõjuvaid põhjusi, ent muudel võrdsetel tingimustel tuleks siiski püüda olla terve mõistuse poolt, mitte vastu. Ja lõpuks on neil printsiipidel, nagu ma [ülal](#PARA.121-6) juba ütlesin, teadvuseteooria väljatöötamisel omamoodi metodoloogilise kitsenduse staatus. Kui teist järku otsustused oleksid alati ebausaldusväärsed või kui enamikku meie elamusi oleks täiesti võimatu märgata, siis oleks meie otsustustel elamuste kohta tegelikkusega nii vähe kokkupuudet, et teadvuseteooria loomist poleks kusagilt alustadagi.

#### Teadvuse ja teadlikkuse omavaheline koherentsus

Kõige fundamentaalsem teadvust ja kognitsiooni ühendav koherentsusprintsiip ei puuduta teist järku fenomenaalseid otsustusi, vaid teadvuse suhet esimest järku otsustustega. Printsiibid, millega meil siin tegemist tuleb, puudutavad ​**[lk.220]**​ teadvuse ja teadlikkuse vahelist koherentsust. Tuletame meelde, et *teadlikkus* on teadvuse psühholoogiline korrelaat, mida võib üldjoontes iseloomustada kui seisundit, kus mingi informatsioon on otseselt kättesaadav käitumise tahteliseks juhtimiseks ja verbaalseks raporteerimiseks. Teadlikkuse sisu vastab esimest järku fenomenaalsete otsustuste sisule (ühe mööndusega, millest [kohe](#PARA.123-3) juttu tuleb), s.t. sisukatele seisunditele — näiteks “See objekt on punane” — mille objektiks ei ole teadvus, vaid mis on teadvusega paralleelsed.

Kus on teadvus, seal on ka teadlikkus. Visuaalse elamusega, mille kutsub esile punane raamat minu laual, kaasneb raamatu funktsionaalne tajumine. Minu tajumissüsteemid töötlevad ja teisendavad optilisi stiimuleid, registreerivad, et laual on nii- ja niisuguse kuju ja värvusega objekt, ja teevad selle informatsiooni kättesaadavaks käitumise juhtimisele. Sama võib öelda elamuse konkreetsete üksikasjade kohta. Igal üksikasjal on teadvuses oma kognitiivne representatsioon. Mõistmaks, et igal üksikasjal peab selline representatsioon olema, piisab, kui panna tähele, et ma võin nende üksikasjade kohta ütlusi anda ja nendest olenevalt oma käitumist suunata: näiteks võin ma osutada raamatu vastavatele osadele. Säärane informatsiooni süstemaatiline kättesaadavus eeldab seda sisu kandva seesmise seisundi olemasolu.

See seesmine seisund on esimest järku fenomenaalne otsustus — vähemalt esimeses lähenduses. Lisasin selle täpsustuse sellepärast, et võidakse kahelda, kas seda seisundit saab rangelt võttes üldse otsustuseks nimetadagi. Selle seisundi sisu ei pea olema midagi niisugust, mida subjekt reflektiivselt kinnitaks, ja ta ei tarvitse seda isegi kontseptualiseerida. Sellist seisundit võib pidada otsustuseks ainult mingis nõrgas mõttes; võib-olla on parem nimetada teda mingitliiki informatsiooniliseks registreeringuks või parimal juhul kaudseks või subpersonaalseks otsustuseks. Ma arutan selle küsimuse üle põhjalikumalt [käesolevas peatükis allpool](#ORG.6.3.2), seni aga tuleb minu juttu neist seisunditest kui otsustustest mõista avaras mõttes osutusena representatsioonilistele seisunditele, mida subjekt ei tarvitse reflektiivselt kinnitada ja millel ei tarvitse kontseptuaalset sisu olla.

Kõik, mis kehtib visuaalse elamuse kohta, kehtib samavõrra kõigi teistegi meeleliste elamuste kohta. Kõigest, millest me saame kuulmiselamuse, on meie kuulmissüsteemis representatsioonid, mis on hilisematele protsessidele käitumise juhtimiseks, sealhulgas sõnaliseks raporteerimiseks kättesaadavad. Inimene, kes ei teaks teadvusest mitte midagi, saaks põhimõtteliselt ikkagi uurida meie kognitiivseid protsesse ja teha kindlaks meie teadlikkuse sisu — selleks tuleb panna tähele rolli, mida informatsioon mängib hilisemate protsesside juhtimisel. Enam-vähem samamoodi võib käsitleda hallutsinatsioone ja teisi olukordi, kus aistingud on, kuid reaalne aistitav objekt puudub. Kuigi reaalset objekti, millega taju sisu seonduks, ei ole, on representatsioon meie tajusüsteemis ikkagi olemas. Macbethil[[275]](#footnote-276) oli esimest järku kognitiivne seisund sisuga “siin on pistoda”, kuid tajutavat pistoda selle elamuse taga ei olnud.

​**[lk.221]**​ Sama rubriigi alla liigituvad isegi mittepertseptuaalsed elamused. Kuigi valuelamusel ei pruugi objekti olla, kuulub tema sisu juurde kognitiivne representatsioon, umbes nagu “miski teeb haiget” — või õigemini ehk “midagi on halvasti”. Seda näitab juba asjaolu, et me võime oma valu kohta ütlusi anda ja vastavalt oma käitumist suunata. Nagu visuaalses tajus, nii on teadlikkus mängus ka siin, kuigi selle teadlikkuse objekt ei ole nii selgepiiriline. Sama võib öelda meie emotsioonide ja muude “seesmiste” elamuste kohta. Kõigil neil juhtudel vastavad elamustele mingid kognitiivsed seisundid; kui neid ei oleks, ei saaks elamuse sisu käitumises kajastudagi.

Paneme tähele, et kõnealune printsiip ei seisne selles, et iga kord, kui meil on teadvuselamus, oleme me sellest elamusest teadlikud. Olulised on siin esimest, mitte teist järku otsustused. Printsiibi sisu on, et kui meil on elamus, siis me oleme selle sisust teadlikud. Kui elamus puudutab raamatut, oleme teadlikud sellest raamatust; valu läbi elades oleme teadlikud millestki valusast; mõtet läbi elades oleme teadlikud sellest, millest me mõtleme. Jutt ei ole elamusest, millele järgneb eraldi veel otsustus, nagu see võiks olla teist järku otsustuste puhul; need esimest järku otsustused on elamuste kaasnähtused ja eksisteerivad nendega kõrvuti.

Seos elamuste ja teist järku otsustuste vahel on palju kaudsem: me oleme küll võimelised oma elamusi märkama, ent märkame neist enamasti ainult sisu, mitte elamust ennast. Ainult aeg-ajalt astume sammu tagasi ja märkame elamust, mida punane raamat meis esile kutsub; enamasti me lihtsalt mõtleme raamatust. Erinevalt teist järku otsustustest, mis on harvad, on esimest järku otsustused täiesti tavalised. Kõige otsesemas seoses on teadvus seega esimest järku otsustustega.

Olen siiani püüdnud näidata, et kus on teadvus, seal on ka teadlikkus. Ent see käib mõlemat pidi: kus on teadlikkus, seal on tavaliselt ka teadvus. Kui me oleme teadlikud millestki oma ümbruses ning meie käitumist suunab mingi sisu, mille kohta me oskame ütlusi anda, siis on meil tavaliselt ka vastav teadvuselamus. Kui minu kognitiivne süsteem representeerib koera haukumist, siis on mul ka koera haukumise elamus; kui ma olen teadlik palavusest enda ümber, tunnen ma palavust; ja nii edasi.

Siin tekib väikesi keerukusi. Võib tunduda, et on olemas teadlikkuse liike, millega vastavaid elamusi ei kaasne. Üheks näiteks võib võtta teadlikkuse mälus sisalduvast informatsioonist. Ma olen teadlik sellest, et Clinton on president, selles mõttes et see informatsioon on mulle kättesaadav, ma võin selle kohta ütlusi anda ja seda oma käitumise tahtelisel suunamisel kasutada. Kui ma aga parajasti sellele asjaolule ei mõtle, siis ei tundu olevat ka vastavat teadvuselamust, või kui ongi, siis äärmiselt nõrk. Samamoodi võin ma olla (hetkel sellele mõtlemata) teadlik sellest, et allkorrusel on jalgratas, ilma et sellega kaasneks elamust. Selline ​**[lk.222]**​ elamusteta teadlikkus avaldub kõige selgemal kujul propositsionaalses teadlikkuses — näiteks kui ma olen teadlik, et minu jalgratas on allkorrusel —, kuigi seda võiks pidada ka mingitlaadi objektiteadlikkuseks, kuna tundub mõistlik öelda, et ma olen teadlik oma jalgrattast.

Võiksime jätta selle küsimuse sinnapaika, kuid parem on seada teadlikkuse mõistele kitsendusi, nii et teadlikkus oleks teadvusega veelgi paralleelsem. Tundub olevat mõistlik öelda, et protsessid, mis neil kahel puhul kulgevad, on mingis suhtes funktsionaalsed erinevad: sellest annab tunnistust juba ainuüksi see, et ma oskan nendevahelise erinevuse kohta ütlusi anda. Just see funktsionaalne erinevus tuleb esile tuua.

Võib-olla kõige silmatorkavam erinevus on see, et juhtudel, kus teadlikkusega kaasneb teadvus, on toimimas mingi otsene ligipääs, mis teadvuseta teadlikkuse juhtudel puudub. Näiteks informatsioon, et Clinton on president, tuleb mälust “välja tellida”, et ta mõjutaks käitumise teadlikku juhtimist, vähemalt siis, kui me vastavale asjaolule hetkel ei mõtle. See informatsioon ei ole käitumise juhtimise mõjutamiseks nii otseselt valmis kui elamuste ja tegelike mõtetega seonduvad kognitiivsed seisundid. Teisiti öeldes, kognitiivne juurdepääs informatsioonile on siin mõnevõrra kaudsem. See asjaolu annabki aluse aktuaalsete ja mitteaktuaalsete​[[276]](#footnote-277) mõtete funktsionaalseks eristamiseks.

Seepärast võime lisada teadlikkuse täpsustatud mõiste määratlusse otsese ligipääsu nõude. Selles täpsustatud tähenduses mitteaktuaalsed mõtted teadlikkuse sisusse ei kuulu, aktuaalsed mõtted aga küll. Sellest tulenevalt võiks oodata, et aktuaalsed mõtted kaasnevad elamustega, mitteaktuaalsed aga mitte. Ja just seda me täheldamegi. Minu mitteaktuaalne mõte, et Clinton on president, ei mõjuta minu fenomenoloogiat, kuid samasisulise aktuaalse mõttega käib kaasas ka vastav elamus. Et seda mõista, paneme tähele, et on olemas tunne mõelda, et Clinton on president; kui ma ei oleks just praegu seda mõtet mõelnud, siis oleks minu tunne olnud mingis raskesti tabatavas mõttes teistsugune.[[277]](#footnote-278).

Seega tundub olevat tõepärane, et kui teadlikkust asjakohaselt defineerida, siis kaasneb teadvusega alati teadlikkus, ja ümberpöördult. Meid huvitavat liiki teadlikkuse kirjeldust saab veelgi täpsustada; ma teen seda [allpool](#ORG.6.3), kus ma vaatlen ka mitmesuguseid huvitavaid näiteid. Kuid juba siinne jäme käsitlus näitab, et see suhe annab kasuliku vaatepunkti, mis lubab teadvuse ja kognitsiooni omavahelise koherentsuse hästi fookusesse saada.

#### Struktuurse koherentsuse printsiip

Siiani oleme vaadelnud hüpoteesi, mis ütleb, et kus on teadvus, seal on ka teadlikkus, ning et kus on (õiget liiki) teadlikkus, seal on ka teadvus. ​**[lk.223]**​ Nende kahe vaheline korrelatsioon ulatub kaugemalegi. Näiteks vastavad teadvuse mitmesugused struktuursed omadused otseselt struktuursetele omadustele, millel on teadlikkuses representatsioon.

Tavaliselt ei ole üksikisiku teadvuselamus üksainus homogeenne kämp, vaid tal on keerukas sisestruktuur. Näiteks minu nägemisväljal on kindel geomeetria. Siin on suur punane laik, selle läheduses väike kollane laik, nende kahe vahel veidi valget; kusagil on triipe, ruute, kolmnurki jne. Elamuses on ka ruumilise sügavuse ilmin­guid: seal on kuupe ja teisi kolmemõõtmelisi kujundeid, ühed asjad on teiste taga jne. Minu nägemisväljas on tohutu mass üksikasju, mis moodustavad ühtekokku kõikehõlmava struktuuri.

Oluline on see, et kõigil neil üksikasjadel on kognitiivsed representatsioonid milleski, millest me võime mõelda kui teadlikkuse struktuurist. Näiteks on mitmesuguste värvilaikude suurusel ja kujul representatsioon minu nägemissüsteemis — võib-olla kaunis täpse topograafilise kaardi kujul, ent kui ka mitte seda, siis mingi representatsioon neil ikkagi on. Seda, et see neil olema peab, kinnitab asjaolu, et vastav informatsioon on kättesaadav käitumise juhtimiseks. Sama kehtib ka triipude, kuubikujuliste figuuride jne. tajumise kohta. Kõik need struktuursed üksikasjad on kognitiivsele süsteemile kättesaadavad ja neid saab kasutada käitumise juhtimiseks, seega on igal neist teadlikkuse sisus oma representatsioon.

Kui kellelgi oleksid täielikud teadmised minu kognitiivsetest protsessidest, siis oleks tal põhimõtteliselt võimalik välja selgitada kõik need struktuursed üksikasjad. Nägemisvälja geomeetria saaks kindlaks teha, analüüsides informatsiooni, mille visuaalne süsteem hilisematele juhtprotsessidele kättesaadavaks teeb; juba ainuüksi asjaolust, et iga selline detail võib kajastuda subjekti käitumuslikes võimetes — näiteks võib subjekt mitmesugustele struktuuridetailidele käega osutada või nende kohta ütlusi anda —, järeldub, et see informatsioon peab kusagil leiduma. Loomulikult oleksid säärase analüüsi üksikasjad väga keerulised ja praegusaegsetele meetoditele täiesti kättesaamatud, kuid me teame, et see informatsioon on olemas. Seega näeme, et teadvuse struktuur peegeldab teadlikkuse struktuuri.

Sama kehtib ka fenomenaalse välja implitsiitse struktuuri, näiteks värvuste omavaheliste suhete kohta. Isegi kui ma näen mingil hetkel ainult üht värvust, oleksin ma võinud näha tervet hulka värvusi, millel on selle värvusega kindel struktuurne suhe. Iga värvus on väga sarnane mõne teise värvusega ja väga erinev kolmandast. Kaks värvust võivad teineteist mingis mõttes täiendada; üks värvuste grupp võib tunduda “soojana” ja teine “külmana”. Lähemal analüüsil selgub, et meie fenomenaalsed värvused moodustavad kolmemõõtmelise struktuuri, reastudes mööda punase–rohelise, kollase–sinise ja valge–musta telge (telgede valik on mõnevõrra meelevaldne, kuid neid on alati kolm), ning et see kolmemõõtmeline fenomenaalne struktuur peegeldub meie tajumissüsteemides, kus töödeldaval värvusinfol ​**[lk.224]**​ on samuti kolmemõõtmeline struktuur. Seda oligi oodata — me teame, ju et vastav informatsioon on käitumise juhtimiseks kättesaadav —, kuid huvitav on jälgida, kuidas selle struktuuri üksikasju on püütud välja selgitada uuemaaegsetes uurimistöödes nägemissüsteemist (vt. arutelu Hardin 1988). Võiksime öelda, et värvusinfol on meie teadvuselamuses omamoodi *diferentsstruktuur* (võimalike elamuste vaheliste erinevuste ruum), millele vastav diferentsstruktuur on olemas ka teadlikkuses: värvuselamuste ja nendevaheliste suhete muutkonnale vastab värvusrepresentatsioonide ja nendevaheliste suhete muutkond.

Sarnaseid implitsiitseid struktuure ja nende vasteid infotöötluse tasemel võib täheldada ka teistes fenomenaalsetes valdkondades. Näiteks peab muusikalise akordi fenomenoloogiline struktuur kajastuma tema representatsiooni struktuuris, et ta oleks raporteeritav ja saaks mõjutada teisi juhtprotsesse. Sama kehtib ka maitseaistingute implitsiitse struktuuri kohta. Empiirilistes uurimistöödes sellekohaste protsesside kohta leitakse sääraseid vastavusi tähelepanuväärselt sageli; ent isegi kui neid uuringuid ei oleks, kinnitab asjaolu, et need struktuursed üksikasjad on juhtprotsessidele kättesaadavad, ikkagi, et mingid vastavused peavad siin olema. Analoogilised arutluskäigud viivad üldisele järeldusele, et igal detailne struktuur, mida fenomenaalses vallas leida võib, peegeldub teadlikkuses sisalduvate representatsioonide struktuurides.

Elamusel on ka mitmesuguseid spetsiifilisemaid omadusi, mis peegelduvad samuti teadlikkuses. Kõige ilmsem neist on elamuse intensiivsus. On selge, et intensiivsus mõjutab hilisemaid protsesse, ja seega peab ta kuidagi representeeruma teadlikkuse struktuuris. Võibki pidada tõenäoliseks, et elamuse intensiivsus vastab otseselt jõule, millega tema aluseks olev esitus püüab juhtrolli mängida, hõivates hilisematele protsessidele kättesaadavaid ressursse. (Mõelge näiteks erinevusele intensiivse valu ja nõrga valu või kogu isiksust valdava emotsiooni ja taustemotsiooni vahel). Üks selline omadus on ka elamuse detailsus, mille näiteks võib tuua erinevuse nägemisvälja kõrge lahutusega keskosa ja madala lahutusega äärealade vahel. Võiks oodata, et see erinevus kajastub ka alusesituste lahutuste erinevuses, ja selgub, et nii see tõepoolest ongi.

Kui elamused ongi mingis mõttes “tabamatud”, siis elamustevahelised suhted seda üldiselt ei ole: meile ei valmista raskusi nende suhete üle arutlemine, olgu jutt sarnasus- ja erinevussuhetest, geomeetrilistest suhetest, intensiivsuste vahekordadest vm. Nagu tähendas Schlick (1938), on teistele hõlpus edasi anda elamuse vormi, selle sisuga (seesmise kvaliteediga) ei tarvitse see aga nii olla: ma võin iseloomustada suhet punase ja rohelise elamuse vahel, aga mitte punasust ja rohelisust ennast.[[278]](#footnote-279) Seega tuleks oodata, et neil suhetel on mingi kognitiivne representatsioon, ja seda me ka tõepoolest täheldame. Elamuste sarnasused ja erinevused ​**[lk.225]**​ vastavad nende elamuste representatsioonide sarnasustele ja erinevustele teadlikkuses: elamuse geomeetria vastab teadlikkuse geomeetriale, jne. Kui teadvuse mõistet ülalöeldud mõttes täpsustada, nii et teadlikkuse seisunditega kaasnevad alati elamusseisundid, siis hakkab tunduma, et paika peab ka vastassuunaline struktuurne vastavus: kui teadlikkus sisaldab mingi struktuuri representatsiooni, siis kajastub see struktuur ka elamuse struktuuris.

Niisiis oleme lisaks üldisele printsiibile, mis ütleb, et kus on teadvus, seal on ka teadlikkus ja ümberpöördult, leidnud ka spetsiifilisema printsiibi: teadvuse struktuur peegeldab teadlikkuse struktuuri ja teadlikkuse struktuur peegeldab teadvuse struktuuri. Ma hakkan seda printsiipi siitpeale nimetama *struktuurse koherentsuse printsiibiks*​.[[279]](#footnote-280) See on keskse tähtsusega ja süstemaatiline seos fenomenoloogia ja psühholoogia vahel ning viimselt taandatav fenomenoloogia ja fenomenoloogiat kandvate füüsikaliste protsesside vahelisele seosele. Allpool saame näha, et sellest printsiibist on kasu mitmeti.

### 3. Veel teadlikkuse mõistest

Üks huvitavamaid filosoofilisi projekte teadvuseteaduses on teadlikkuse mõiste täpsustamine nii, et sellest saaks teadvusele veelgi täiuslikum psühholoogiline korrelaat. Oma esialgses määratluses on teadlikkuse vastavus teadvusega ebatäielik, kuid tema mõistet on võimalik täpsustada, et saada paremini hakkama raskemate juhtudega. Me tahaksime välja jõuda niisuguse teadlikkuse mõisteni, mis oleks teadvuselamusega kõikjal tõepärases korrelatsioonis, vähemalt meile tuttavatel juhtudel.

Esialgselt määratlesin ma teadlikkust kui seisundit, kus mingi informatsioon on kättesaadav sõnaliseks raporteerimiseks ja käitumise tahteliseks juhtimiseks. Juhtumid, kus on olemas propositsionaalne teadlikkus, kuid puudub elamus, viisid mõttele muuta teadlikkuse mõistet nii, et see eeldaks tingimata just otsest ligipääsu. Võimalikud on veel teisedki muudatused. Kõige ilmsem neist tuleneb sellest, et, nagu näitavad kaalutlused elamuse kohta imetajatel, ei ole teadvuselamuse puhul kättesaadavus sõnaliseks raporteerimiseks rangelt nõutav, kuigi ta on hea heuristiline kriteerium juhtudel, kus keel on olemas. Loomulik oleks muuta teadlikkuse määratlust umbes nii: teadlikkus on otsene kättesaadavus globaalseks juhtimiseks. Teisiti öeldes, subjekt on mingist informatsioonist teadlik siis, kui see informatsioon on otseselt kättesaadav väga mitmesuguste käitumuslike protsesside suunamiseks. See lubab kõnelda elamusest ka teistel loomadel peale inimese, ja see klapib hästi kokku ka raporteeritavuskriteeriumiga. Juhtudel, kus informatsioon on raporteeritav, on ta üldjuhul ka kättesaadav globaalseks juhtimiseks, näiteks väga mitmesuguste käitumisaktide sihiteadlikuks suunamiseks. Vastupidine järeldus ei kehti alati, näiteks kui jutt on loomade käitumisest, kuid vähemalt subjektide puhul, kes on suutelised ütlusi andma, tähendab informatsiooni kättesaadavus globaalseks juhtimiseks üldiselt ka kättesaadavust raporteerimiseks.

​**[lk.226]**​ Muidugi on taolisel täpsustamisprojektil omad piirid, kuna meil puudub “elamusmeeter”, millega neid hüpoteese empiiriliselt kinnitada ja täpsustada. Siiski on meil esimese isiku juhuks hea ettekujutus seisunditest, millega meil kaasnevad elamused, ja seisunditest, mille puhul neid ei ole; analüüsides, mis neil juhtudel toimub, võime tavaliselt jõuda nende seisundite funktsionaalse iseloomustuseni. Seega annab elamuse ja funktsiooni vahelise seose uurimine tuttavatel juhtudel meile märkimisväärset tuge. Samuti võime proovida neid hüpoteese empiiriliselt täpsustada esimese isiku eksperimentide abil. Näiteks võime asetada end mingisse funktsionaalsesse seisundisse ja panna tähele, missugune on seejuures meie elamus. Homogeensus- ja usaldusväärsusprintsiipi pisut appi võttes võime teha järeldusi ka uurimustest analoogiliste juhtumite kohta teistel.

Mingit rolli võib mängida ka vähem tavapäraste juhtude empiiriline vaatlemine. Võime näiteks küsida, misliiki elamusi võiks tõenäoliselt olla subjektidel, kes kannatavad nii- või teistsuguste haiguslike seisundite all, või loomadel. Loomulikult ei saa me kunagi täiesti kindlalt öelda, missugused on nendel juhtudel elamused, kuid mõned järeldused on teistest palju tõepärasemad. Põhimõtteliselt aitavad need juhtumid meil mõtlemist suunata ja fantaasial välja destilleerida tõepärased printsiibid, mis puudutavad elamuse ja funktsiooni omavahelist seost. Need printsiibid võivad viimselt toetuda mitte-empiirilisele analüüsile, kuid tähelepanu koondamine empiirilistele juhtudele aitab vähemalt siduda meie mõttekäike tegeliku maailmaga.

Näiteks mõtisklused elamuste omistamise üle imetajatele annavad tulemusi, mis sobivad kokku minu [ülal](#PARA.125-2) pakutud täpsustatud kriteeriumiga. Me oleme üldiselt valmis omistama imetajatele mingi tajulise stiimuli läbielamist juhul, kui käitumise juhtimist on võimalik panna sellest stiimulist sõltuma, ja eriti veel juhul, kui see on võimalik väga mitmesugustes käitumissituatsioonides. Kui me oleme suutnud panna informatsiooni stiimuli kohta avalduma ainult ühes, suhteliselt väheolulises käitumuslikus reaktsioonis, võime veel oletada, et see informatsioon läks teadvusest täiesti mööda. Ent sedamööda, kuidas informatsiooni kättesaadavus suureneb, hakkab oletus, et sellega kaasneb mingi elamus, tunduma järjest tõepärasemana. Seega on teadvuse ja sääraselt defineeritud teadlikkuse omavaheline koherentsus kooskõlas nii esimese isiku andmetega kui ka loomulike arutlustega loomade kohta.

On veel hulk teisigi huvitavaid probleemseid juhtumeid, mis analüüsimiseks huvi pakuvad. Üks seesugune on *pimenägemine* (mida on kirjeldanud Weiskrantz (1986)). See puue tuleb nägemiskorteksi kahjustusest, mille puhul tavaline visuaalse informatsiooni töötlus­trajektoor on viga saanud, kuid visuaalse informatsiooni töötlus näib siiski mingil vähesel määral toimuvat. Pimenägija ei näe mõnes oma nägemisvälja piirkonnas vähemalt omaenda ütluse järgi midagi. Kui süüdata tema “pimedas piirkonnas” punane või roheline tuli, ütleb ta, et ei näe midagi. Ent kui panna teda ikkagi mõistatama, mis selles piirkonnas on, näiteks kas seal on punane või roheline tuli, siis selgub, et ta annab ikkagi palju sagedamini ​**[lk.227]**​ õige vastuse kui eksliku. Kuigi ta ei näe, mis selles piirkonnas on, mingil moel ta ikkagi midagi “näeb”.

Pimenägemist tuuakse mõnikord näiteks teadvuse irdumise kohta oma funktsionaalsest rollist. On ju pimenägemise puhul toimumas eristamine, kategoriseerimine ja mingis mõttes isegi sõnaline raporteerimine, kuid teadvuselamust nähtavasti ikkagi ei ole. Kui funktsionaalne roll oleks siin elamusest täiesti lahus, siis oleks meil tõsist põhjust koherentsusprintsiibis kahelda. Õnneks on järeldus, et selles olukorras on teadlikkus olemas, teadvust aga mitte, alusetu. Kõigepealt ei ole sugugi ilmne, et elamust säärastel juhtudel üldse ei ole: mingi nõrk elamus, millel on verbaalse raporteerimisega mingi iseäralik seos, võib siin siiski olla. Kuid, mis veelgi olulisem, see ei ole kaugeltki tüüpiline teadlikkuse juhtum. Siin on funktsionaalsed rollid väga erinevad nendest, mis etenduvad tavalistel juhtudel. Just selle funktsionaalsete rollide erinevuse tõttu me ju märkamegi, et siin on midagi valesti.[[280]](#footnote-281).

Muu hulgas puudub pimenägijatel nähtavasti tavaline ligipääs olemasolevale informatsioonile. Nendel on ligipääs eriskummaliselt kaudne, nagu näitab asjaolu, et vastav informatsioon ei ole sõnaliseks raporteerimiseks ja käitumise tahteliseks juhtimiseks otseselt kättesaadav. Ta on kättesaadav palju väiksemale arvule juhtprotsessidele kui tavaline tajuinformatsioon; infot saab teha teistele protsessidele kättesaadavaks, kuid ainult ebatavaliste meetoditega, näiteks ergutades või sundides isikut valikut tegema. Seega ei saa pidada seda informatsiooni globaalsele juhtimisele otseselt kättesaadavaks, ja katseisikud ei ole sellest informatsioonist päriselt teadlikud selle sõna meid huvitavas mõttes. Elamuse puudumine vastab otseselt teadlikkuse puudumisele. On ka võimalik, et pimenägemisega isikutel on mingi nõrk elamus siiski olemas, ja sellisel juhul võiks öelda — kui valida otsesuse ja globaalsuse mõõdupuud sobival viisil —, et neil on ka mingi nõrgavormiline teadlikkus. Ligipääsu puudumise tõttu asjakohastele faktidele jääb olukorra kirjeldus mõnevõrra ebamääraseks, ent nii või teisiti on see juht ühitatav teadvuse ja teadlikkuse vahelise koherentsusega.

Üldiselt ei saa seesugused juhtumid olla tunnistuseks funktsionaalse organisatsiooni ja teadvuselamuse vahelise seose vastu, sest järeldused teadvuse olemasolu või puudumise kohta teeme me just funktsionaalsetel alustel. Näiteks toetuvad tõendid tavatute teadvusseisundite kohta säärastel patoloogilistel juhtudel enamasti ainuüksi tõenditele tavatute teadlikkusseisundite kohta. Seega ei saa need juhtumid koherentsusprintsiipi ohustada, vaid ainult toetada ja täpsustada.

Üheks keerukaks probleemiks on elamused unes. Tõenäoliselt on meil unenägude ajal elamused olemas (kuigi vt. Dennett 1978b), kuid puudub raporteerimisvõime, ning tegevuse täieliku puudumise tõttu pole elamustel mingit rolli tegevuse juhtimisel. Siiski on need juhtumid tõepäraselt analüüsitavad, kui koondada tähelepanu kättesaadavusele globaalseks juhtimiseks; tõsi, enamik asjaomaseid juhtprotsesse ​**[lk.228]**​ on une ajal seiskunud. Võib-olla jõuab informatsioon sinnasamasse, kust ta on tavaliselt juhtimiseks kättesaadav; seda oletust toetab äsjase unenäo sisu kättesaadavus une ja ärkveloleku vahepealses seisundis. Kui see on nii, siis võib ikkagi vaadelda järgmist kontrafaktuaali: kui raporteerimisvõime oleks olnud olemas ja juhtimine võimalik (näiteks kui motoorne ajukoor oleks talitlenud normaalselt), siis oleks informatsioon võinud midagi mõjutada. Kuid küsimus väärib hoolikamat analüüsi, samuti empiirilisi uuringuid, mis selgitaksid välja, mis inimesega une ajal tegelikult toimub.

Mõningaid huvitavad juhtumeid mainib Block (1995)[[281]](#footnote-282) oma pikemas arutelus fenomenaalse teadvuse ja “ligipääsuteadvuse” vahelisest erinevusest. Blocki käsitluses on üks või teine seisund *ligipääsuteadvuslik*​, kui tema sisust on võimalik lähtuda kui mõtlemise eeldusest ja see sisu on sobiv tegevuse ja kõne ratsionaalseks juhtimiseks. Seega vastab ligipääsuteadvus enam-vähem minu esialgsele teadlikkuse definitsioonile, kuigi minu oma paneb ratsionaalsusele vähem rõhku. Block vaatleb mõningaid juhtumeid, kus need kaks teadvuse liiki võivad teineteisest irduda. On õpetlik uurida, kuidas aitaks nendega toime tulla koherentsusprintsiip.

Kõneldes võimalikust olukorrast, kus on olemas ligipääsuteadvus, kuid puudub fenomenaalne teadvus, apelleerib Block ainult kontseptuaalselt võimalikele juhtudele, näiteks sombidele; on ilmne, et need mitteaktuaalsed juhtumid ei saa koherentsusprintsiipi mitte kuidagi kõigutada. Ta mainib ka pimenägemist, kuid tähendab, et see fenomen lubab kõnelda ligipääsuteadvusest ainult mingis nõrgas mõttes. Ta vaatleb ka *ülipimenägemist* — see on pimenägemine, mille korral subjekt on treenimise teel saavutanud palju parema ligipääsu informatsioonile nägemisvälja pimeda osa kohta. Mõned juhtumid, millest Block kõneleb, on ilmselgelt vaid mõeldavad olukorrad, kus teadvus ei figureeri; kuid Block märgib ka ise, et pole alust uskuda, et niisuguseid juhte võiks tegelikkuses ette tulla. Huvitaval kombel märgib ta empiirilistele näidetele kõige lähemate juhtude kohta (ahv, keda kirjeldas Humphrey (1992) ja inimesest patsient, kellest kõneleb Weiskrantz (1992); vt. ka Cowey ja Stoerig 1992), et nende puhul on põhjust uskuda, et mängus on tõepoolest fenomenaalne teadvus.

Mis puutub fenomenaalsesse teadvusse ilma ligipääsuteadvuseta, siis selle kohta mainib Block mõningaid tegelikkuses aset leidnud juhtumeid. Üks neist on olukord, kus subjekt saab äkki teadlikuks sellest, et tema kõrvus on juba mõnda aega kõlanud vali puurimine kusagil eemal. Block oletab, et enne, kui subjekt seda märkas, oli see heli tema fenomenaalses teadvuses, kuid mitte tema ligipääsuteadvuses. Ent kui määratleda teadlikkust nii, nagu mina siin pakkusin, siis tundub mõistlik öelda, et subjekt oli puurimisest teadlik kogu aeg. Võib vist öelda, et asjaomane informatsioon puurimise kohta oli kogu aeg kättesaadav, ta lihtsalt ei kasutanud seda. Nii et kui defineerida ligipääsuteadvust või teadlikkust soodumuslikult, siis ei valmista see juhtum koherentsusprintsiibile mingeid raskusi. Block mainib ka juhtumit, kus katseisiku ees vilgutatakse lühikest aega ruudustikku, milles on 3×3 kirjatähte (Sperling 1960). Kui isikutel palutakse nimetada, mis tähed olid ülemises reas, tulevad nad sellega toime, kuid ei oska seejärel nimetada tähti teistes ridades, ja samamoodi on see ka teiste ridadega. Block ​**[lk.229]**​ kinnitab, et katsealuse fenomenaalse teadvuse sisus on kõik üheksa tähte, ligipääsuteadvuses aga ainult kolm tähte korraga. Kuid jällegi tundub siin tõenäolisemana oletus, et algselt oli kättesaadav informatsioon kõigi üheksa tähe kohta, kuid katseisik kasutas ainult informatsiooni kolme tähe kohta, ja protsess, mille varal ta seda tegi, hävitas ligipääsu ülejäänud informatsioonile. Seega on ka see juhtum koherentsusprintsiibiga ühitatav, kui käsitada teadlikkust soodumuslikult.

Vaadelda võiks veel paljusid teisigi juhtumeid. Siin kõnelesin ma vaid mõnedest ning tõin paar väikest näidisanalüüsi, et anda aimu ühe huvitava filosoofilise projekti kontuuridest. Hoolikama analüüsi korral võiks püüda seada rangemaid kitsendusi teadvuselamusega kaasneva ligipääsetavuse laadile ning selle rollile globaalses juhtimises. Teadlikkuse käsitlemine otsese kättesaadavuse mõttes globaalsele juhtimisele on alles lähtekoht. Edasist analüüsi on ootamas viljakas tööpõld.

#### Suhted funktsionalistlike teadvuseteooriatega\*

Kirjeldatud projektis võib näha funktsionalistliku teadvusekäsitluse otsingut. Soovitud käsitlus ei ole reduktiivne — ta ei ütle, et teadvus pole midagi enamat kui funktsionaalne roll, või et teadvuse funktsionaalse rolli seletamisest piisab teadvuse seletamiseks —, vaid ta on, vastupidi, mittereduktiivne käsitlus, mis püüab leida teadvuse tekke funktsionaalseid kriteeriume. Kuid mõneti mängib ta reduktiivfunktsionalistlike käsitlustega samal väljakul; kõrvuti pretensioonikamate metafüüsiliste taotlustega püüavad ka need leida teadvuse funktsionaalseid kriteeriume. Huvitav on jätta metafüüsilised erinevused kõrvale ning võrrelda mitmesuguseid käsitlusi ainuüksi pakutud funktsionaalsete kriteeriumide alusel.

Näiteks idee, et, teadvus käib käsikäes otsese kättesaadavusega globaalseks juhtimiseks, meenutab Dennetti (1993b) ütlust, et teadvus on “tserebraalne staar” [​*cerebral celebrity*​]: “Teadvus on tserebraalne staar — ei rohkem ega vähem. Teadvusse jõuavad need protsessid, mis kestavad ja monopoliseerivad ressursse küllalt pikka aega, et avaldada tüüpilisi ja “sümptomaatilisi” toimeid mälule, käitumise juhtimisele jne.” (lk. 929).

Jättes kõrvale asjaolu, et Dennett peab toodud väidet kontseptuaalseks tõeks, on tema vaated üsna lähedased minu omadele. Peamine erinevus seisneb selles, et minu arusaama järgi käib teadvus käsikäes juba potentsiaalsegi tserebraalse staariga. Selleks, et mingi sisu jõuaks teadvusse, ei ole tingimata tarvis, et tal oleks mingi tegelik roll globaalses juhtimises — piisab, kui ta on selleks kättesaadav. Niisugune nõue näib paremini klappivat elamuse omadustega. Näiteks on meie elamuses ka nägemisvälja äärealad, kuigi neil enamasti globaalses juhtimises erilist rolli ei ole — kuid vajadusel on nad selleks kättesaadavad. Paljud ​**[lk.230]**​ taustmürad meie elamuses ei tarvitse avaldada olulist mõju mälule, käitumisele ja muule sellisele, kuid samas oleks vastav informatsioon võinud seda teha. Muidugi pole võimatu, et Dennett kasutab sõna “teadvus” tugevamas mõttes, nii et tema arusaamise järgi need äärealad ja taustmürad teadvusse ei jõuagi (peab ju Dennett elamuse ideed üldse kahtlaseks), kuid võrdlus pakub siiski huvi.

Veel ühe funktsionalistliku käsitluse on välja pakkunud Rosenthal (1996), kes väidab, et teadvusse jõuavad niisugused seisundid, mis on mingi kõrgemat järku mõtte objektiks. Minu terminoloogias tähendab see, et esimest järku seisund jõuab teadvusse parajasti siis, kui ma olen moodustanud tema kohta teist järku otsustuse. See on tunduvalt tugevam nõue kui minu oma, samas mõttes nagu on tugevam ka Dennetti nõue. Pealtnäha pole just palju põhjust uskuda, et me moodustame kõigi oma elamuste kohta — sealhulgas iga viimsegi üksikasja kohta nägemisväljas, iga taustheli kohta jne. — teist järku otsustusi. Rosenthal arvab, et teist järku otsustused tavaliselt ise teadvusse ei jõua, mistõttu me nende olemasolu ei märka. Kuid tundub, et nende vastu tunnistavad isegi kolmanda isiku positsioonist lähtuvad kaalutlused. Tundub, et kõik need teist järku otsustused on kognitiivse süsteemi ehitusplaanis täiesti tarbetud. Võiks oodata, et süsteem on suuteline moodustama neid otsustusi siis, kui neid tarvis läheb, nagu see toimub meie märgatavamate elamuste puhul, kuid süsteem, mis moodustaks nägemisvälja iga detaili kohta teist järku otsustuse, oleks äärmiselt liiane (joonis 6.1).

**Joonis 6.1.** Zippy the Pinhead kõrgemat järku teadvuseteooriate kohta. (​Avaldatud King Features Syndicate’i eriloal.​)

 - — Ma jälgin end jälgimas ennast joonistamas jälgivana end ennast jälgimas!

 — Aga mida sina teed, Zippy?

 - — Mina jälgin ennast jälgimas sind jälgivana end ennast jälgimas.

 — Räägi oma eneseteadvusest!!

 - — Ma jälgin enesejälgimist, joonistades iseennast ja kaht teist iseennast jälgimisprotsessi jälgimas!!

 — Kes tellis[[282]](#footnote-283) ananassidega singi??

Rosenthali kontseptsiooni järgi on teadvusseisundid need seisundid, millest me oleme teadlikud [​*conscious of*​]. See kõlab võib-olla usutavaltki, kuid usutavalt minu arvates ainult selles nõrgas mõttes (millest oli juttu [eelmises peatükis](#PARA.113-4)), mille järgi kõik meie elamused on meile tuttavad.[[283]](#footnote-284) Ei ole sugugi ilmne, et enamik meie elamustest on meie mõtete objektid. Oletus, et meil on elamuse ​**[lk.231]**​ iga üksikasja kohta kaks kognitiivset seisundit — esimest järku ja teist järku otsustus —, annaks psüühikast väga kaootilise pildi. On raske mõista, miks oleks evolutsioon pidanud jändama kõikjal nende teist järku otsustustega, kui tema eesmärke oleks sama hästi teeninud lihtne kättesaadavus globaalseks juhtimiseks. Parem on näha Rosenthali käsitluses introspektiivse teadvuse seletust, kuigi ta ise pakub seda välja just kui teadvuse käsitlust “tundena olla”.

Kasulik on jaotada teadvuse funktsionaalsed käsitlused *esimest järku* ja *teist järku* käsitlusteks. Teist järku käsitlustes (millede hulka kuulub Rosenthali käsitlus teadvusest kui kõrgemat järku mõttest, samuti Lycani (1995) jt. kõrgemat järku tajude kontseptsioon) on teadvuseks esmatähtis mingi teist järku kognitiivse seisundi olemasolu. Esimest järku teooriate järgi on teadvuseks nõutav ainult esimest järku kognitiivne seisund, kusjuures selle rollile rakenduvad mõningad kitsendused. Teist järku teooriad võivad anda hea käsitluse introspektsioonist või reflektiivsest teadvusest, kuid esimest järku teooriad tunduvad pääsevat teadvuselamusele lähemale[[284]](#footnote-285).

Loomulikult ei kaasne mitte iga esimest järku kognitiivsele seisundiga teadvuselamus; võib leiduda esimest järku otsustusi maailma kohta, mis ei vasta ühelegi elamusele. Seepärast vajab säärane teooria veel mingit komponenti, mis lubaks tõsta esimest järku seisundite hulgast esile relevantsete seisundite klassi. Ilmne viis selleks on kitsendada nende seisundite rolli​. Seda ma olengi teinud, oletades, et relevantsed on parajasti need esimest järku otsustused, mis on otseselt kättesaadavad globaalseks juhtimiseks. Nagu allpool näeme, pakuvad ka teised esimest järku teooriad välja samalaadseid kitsendusi. Võidakse väita, et põhjusliku mõju nõue on vaid veel üks selline kitsendus kõrgemat järku mõtetele; kuid häda on selles, et see kitsendus tundub olevat liiga tugev.

Üks huvitav kompromissoletus on, et teadvusseisund vastab esimest järku otsustusele, mille kohta on potentsiaalselt võimalik moodustada teist järku otsustust.[[285]](#footnote-286) [[286]](#footnote-287) Nii väldiksime eelmise ettepaneku kaootilisusest ja pilt saaks tõepärasem. Tõepoolest, see oletus ei olegi nii väga erinev ettekujutusest, et teadvusseisund vastab esimest järku otsustusele, mis on kättesaadav globaalseks juhtimiseks: eeldatavasti käib kättesaadavus globaalseks juhtimiseks enamasti käsikäes kättesaadavusega teist järku otsustuse moodustamiseks. Selle ideega võib aga tulla raskusi selliste süsteemide puhul nagu imikud ja loomad, kellel on elamused nähtavasti olemas, kuid puudub võime moodustada teist järku otsustusi; tundub, et see nõuab suuremat kontseptuaalset keerukust kui elamusteks tarvis. Kui nii, siis on parem võtta kirjelduse aluseks siiski kättesaadavus globaalseks juhtimiseks.

Võib arvata, et kõik funktsionalistlikud esimest järku teooriad toovad sisse kitsendusi, mis seonduvad mingisuguse kättesaadavusega. Seda kinnitab olemasolevate reduktiivsete käsitluste vaatlus. Näiteks Kirk (1992) oletab, et selleks, et mingi taju jõuaks ​**[lk.232]**​ teadvusse, peab saabuv informatsioon olema “kohal” peamiste protsesside jaoks, mis kujundavad süsteemi otsuseid, Kirk (1994) aga ütleb, et see informatsioon peab olema “otseselt aktiivne”. Midagi sarnast pakuvad välja ka Dretske (1995), kes ütleb, et elamus on informatsioon, mis representeerub süsteemile, ja Tye (1995), kelle järgi informatsioon peab olema mingil sobival viisil “valmis” kognitiivseks töötluseks. Kõik need oletused on arvatavasti suuremate raskusteta ühitatavad. Kõik nad näivad püüdvat väljendada sarnast ideed.

Igal juhul on huvitav, et teadvuse mittereduktsionistliku käsitluse pooldaja ei pea vaatama mitmesuguste esimest ja teist järku funktsionalistlike teooriate vahelistele lahkarvamustele kui perspektiivitute teooriate vahelisele sisesõjale. Kuigi need teooriad ei suuda teadvust seletada, on nad siiski täiesti väärilised kandidaadid teadvuse kognitiivse baasi teooria rollile, ja mõned neist tulevad oma ülesandega paremini toime kui teised. Isegi omadusdualist võib tunnistada, et neis on oma tõetera, ja omistada nendevahelistele erinevustele mõningat tähtsust.

#### Esimest järku otsustused ja esimest järku registreeringud\*

Olen püüdnud näidata, et elamusseisundid on otseses vastavuses neid kandvate kognitiivsete seisunditega, mida ma olen nimetanud esimest järku otsustusteks. Kuid nagu ma [ülal](#PARA.123-3) märkisin ja nagu on toonitanud ka Dretske (1995), on võib-olla eksitav nimetada neid seisundeid otsustusteks. Tuletagem meelde, et me defineerisime otsustusi esialgselt nii, et nad ei erinenud kuigi palju uskumustest (kui mitte arvestada fenomenaalset elementi). Ent kuigi on mõistlik oletada, et vaateväljaelamuses vastab igale üksikasjale omaette representatsiooniline seisund, pole ilmne, et subjektil oleksid kõigi nende detailide kohta ka uskumused. Näiteks ei ole mul arvatavasti uskumusi oma nägemisvälja äärealade sisu kohta, vähemalt senikaua kui ma neile tähelepanu ei pööra. Sellegipoolest on isegi uskumuste puududes olemas mingi asjakohast informatsiooni sisaldav kognitiivne seisund, sest see informatsioon on vähemasti kättesaadav.

Võiksime lihtsalt leppida kokku kasutada terminit “otsustus” avaramas mõttes, mis katab lisaks selgesõnalistele uskumustele ka seesugused kognitiivsed seisundid. On ju ikkagi usutav, et nägemisvälja äärealade representatsioone võib vaadelda kui “mikrootsustusi” või implitsiitseid otsustusi, mida moodustavad kognitiivsetes süsteemides kulgevad protsessid, isegi kui need ei ole subjekti kui terviku otsustused. Kuid vististi on parem vältida segadust otsustuse mõistega ning võtta seeasemel tarvitusele avarama tähendusega termin representatsiooniliste seisundite jaoks, mis ei ole tingimata otsustused. Selleks võtan ma edaspidiseks tarvitusele termini “registreering”. Näiteks kannavad tajutavate seisundite kognitiivset sisu esimest järku registreeringud, mitte esimest järku otsustused. Esimest järku registreering on seisund, mida subjekt ei tarvitse tunnistada, kuid mis on siiski ​**[lk.233]**​ sisukas ja subjektile kättesaadav seisund ning millel on tema kognitiivsüsteemis mingi roll.

Esimest järku registreeringud võivad esimest järku otsustustega mõnikord isegi vastuollu minna. Selgeks näiteks selle kohta on optilised illusioonid: subjekt võib teada, et kaks objekti on ühesuurused, kuid tajus representeeruvad nad ikkagi erisuurustena. Veel ühe näite toob Dretske: te tõstate korraks seitse sõrme ja ma näen kõiki seitset; ent kuna ma ei jõua neid kokku lugeda, siis otsustan ma ekslikult, et sõrmi oli kaheksa. See tähendab, et oma otsustuse järgi nägin ma kaheksat sõrme, kuigi fenomenaalse elamuse sisus oli seitse sõrme. Seega ei ole otsustused fenomenoloogiaga alati täpselt paralleelsed. Ent kusagil tajusüsteemis leidub representatsioon, mis vastab visuaalsele informatsioonile seitsmest sõrmest, ning see on teistele protsessidele edaspidiseks kasutuseks kättesaadav. Seda varasemat representatsiooni ma nimetangi esimest järku registreeringuks. Me võime kujutleda esimest järku registreeringuid taju- ja introspektsiooniprotsesside otsese produktina, mis moodustub enne nende ratsionaalset integreerumist ühtseks tervikuks.[[287]](#footnote-288).

Seega ei moodustu teadlikkuse sisu rangelt võttes mitte esimest järku otsustustest, vaid esimest järku registreeringutest. Esimeses lähenduses moodustavad teadlikkuse sisu need esimest järku registreeringud, mis on otseselt kättesaadavad globaalseks juhtimiseks. Kui asju niiviisi defineerida, siis on teadlikkuse sisu otseses vastavuses teadvuse sisuga. Muidugi jäävad mõned esimest järku registreeringud teadlikkuse sisust välja, näiteks subliminaalsed tajud. Nagu otsustuste puhul, nii võime ka siin kõnelda esimest järku fenomenaalsetest ja mittefenomenaalsetest registreeringutest, eristamaks registreeringuid, millele vastab mingi elamus, registreeringutest, millele vastavaid elamusi ei ole. Ent kuna mind huvitavad ainult need registreeringud, mis kuuluvad esimesse klassi, siis ütlen ma enamasti lihtsalt “esimest järku registreeringud” ega täpsusta alati otsesõnu, et jutt on fenomenaalsetest registreeringutest.

(Esimest järku registreeringute representatsioonilist sisu, nagu ka sellega paralleelset elamuse sisu, on arvatavasti kõige parem mõista kui mittekontseptuaalset sisu. Võtan selle ja veel mõned teised küsimused teadvuse ja elamuse sisu kohta arutuse alla lõppmärkuses.)[[288]](#footnote-289) [[289]](#footnote-290).

### 4. Koherentsusprintsiipide seletuslik roll

Koherentsusprintsiibid, mida ma olen siin laias laastus kirjeldanud, ei ole pelgad metafüüsilised mängunupud. Empiirilises töös teadvuselamusega võivad nad täita keskset rolli. Selleks, et teadvuse empiiriline uurimine võiks edasi liikuda, on tarvis mõningaid eeleksperimentaalseid arutlusi, mis võimaldaksid teha füüsikaliste andmete põhjal järeldusi teadvuselamuse kohta. Vajaliku lähtepunkti annavadki koherentsusprintsiibid. Kui need on paigas, on meil käes metodoloogiline vundament ​**[lk.234]**​ teadvuselamuse empiiriliseks uurimiseks mitmes valdkonnas. Suur osa sellest tööst on juba teoksil; koherentsusprintsiibid toovad vaid päevavalgele eeldused, millest see töö vaikivalt lähtub.

Peamisi projekte, kus need printsiibid võivad seletamises rolli mängida, on vähemalt kolm. Esiteks võib struktuurse koherentsuse printsiibist olla abi projektides, kus mingit kindlat liiki elamuste struktuuri seletamiseks võetakse appi faktid füüsikaliste protsesside kohta. Teiseks toimib teadvuse ja teadlikkuse vaheline koherentsus omamoodi episteemilise hoovana, mis lubab uurijatel teha järeldusi elamuse kohta kolmanda isiku andmete põhjal. Ja kolmandaks võib teadvuse ja teadlikkuse vaheline koherentsus teenida meid taustprintsiibina teadvuse füüsikaliste korrelatsioonide otsimisel. Vaatleme neid järgemööda.

Esimene neist projektidest pakub meieaegses praktikas kõige selgemat näidet. On täiesti tavaline asi, et elamuste struktuursete omaduste väljaselgitamiseks võetakse appi uurimistööd neurobiolooga ja kognitiivsete protsesside alal. Näiteks on fenomenaalse värvusruumi struktuuri seletamisel, nagu [juba mainitud](#PARA.124-5), väga palju kasu värvusnägemise aluseks olevate protsesside uurimisest. Analoogiliselt aitab nägemiskorteksi topograafiliste kaartide uurimine heita valgust fenomenaalse nägemisvälja struktuurile, ja kuulmiskorteksis toimuvate protsesside uurimine aitab mõista paljusid kuulmiselamuste struktuurseid aspekte (näiteks helikõrguste suhteid ja helide lokaliseerimist). Midagi sarnast toimub ka paljudes teistes fenomenaalsetes valdkondades.

See jutt võib teha lugeja nõutuks: kuidas võib lugu füüsikalistest protsessidest, olgu see missugune tahes, heita valgust elamuse omadustele — kui arvestada kõike, mida ma olen öelnud reduktiivse seletamise võimatuse kohta? Struktuurse koherentsuse printsiip lubab meil mõista, mis siin toimub. Sisuliselt on see printsiip siin tausteelduseks, mis loob silla füüsikaliste protsesside omadustelt elamuse omadusteni. Kui eeldada, et teadvuse struktuur ja teadlikkuse struktuur on omavahel koherentsed, siis tuleb teadvuse ühe või teise konkreetse aspekti seletamiseks seletada ära ainult vastav teadlikkuse aspekt. Kogu ülejäänud töö teeb ära sildprintsiip.

Näiteks värvuste puhul annab lugu füüsikalistest protsessidest meile teadlikkuse struktuuri reduktiivse kirjelduse, seletades asjassepuutuvaid sarnasusi ja erinevusi visuaalsete stiimulite vahel, mida värvussüsteem töötleb ja järgmistele süsteemidele kättesaadavaks teeb. Kui ülevaade värvustaju relevantsest struktuurist on käes, siis ütleb koherentsusprintsiip, et see struktuur kajastub ka värvuselamuse struktuuris. Seega annab visuaalsete protsesside funktsionaalne käsitlus meile, kui tunnistada koherentsusprintsiipi, kaudselt ka fenomenaalse värvusruumi struktuuri kirjelduse. Samal meetodil saab seletada ka paljusid teisi elamuse omadusi.

​**[lk.235]**​ Tugeva mulje mõjul, mida mõnele autorile on avaldanud teadvuse struktuuri ja kognitsiooni struktuuri vaheline koherentsus, on nad väljendanud arvamust, et teadvuse füüsikaliseks seletamiseks polegi midagi muud tarvis. Näiteks Van Gulick (1993) nendib, et meie värvusruumi struktuur vastab otseselt struktuurile, mida representeerib visuaalne infotöötlus, ja oletab, et see asjaolu täidabki seletuslünga, andes värvusaistingutele funktsionaalse seletuse. Clark (1993) pühendab sellele strateegiale terve raamatu, kus ta püüab näidata, et sensoorseid kvaliteete on võimalik täielikult seletada, vaadeldes sarnasus- ja erinevussuhteid kvaliteediruumides.

Kui see, mida ma enne ütlesin, on õige, siis on need väited pisut liiga tugevad. Esiteks ei seleta see meetod värvuselamuse seesmist olemust, nagu tõestab struktuuri säilitava spektriinversiooni võimalikkus. Parimal juhul seletab ta struktuurseid sarnasusi sedaliiki elamuste või liitelamuse osade vahel; seega on teadvuse täielikuks seletamiseks tarvis midagi enamat. Teiseks, mis veelgi tähtsam, ükski teadlikkuse struktuuri kirjeldus ei seleta, miks üldse kaasneb teadlikkusega elamus — just sellepärast, et ta ei suuda seletada, miks struktuurse koherentsuse printsiip üldse kehtib. Seda printsiipi tausteelduseks võttes oleme juba lahkunud reduktiivse seletamise pinnalt: see printsiip eeldab teadvuse olemasolu juba ette ega aita teda mitte kuidagi seletada. Seda võib pidada üheks mittereduktiivse seletamise erikujuks, mis lähtub teadvuse olemasolust kui eeldusest ning püüab seletada tema mõningaid omadusi.

Neis piirides annab struktuurse koherentsuse printsiip meile tohutult kasuliku seletusliku seose füüsikalise ja fenomenaalse valla vahel. Selleks, et seletada mõnd ilmset struktuuri fenomenaalses maailmas — näiteks seoseid, mida me leiame valitsevat muusikalistest akordidest saadavate elamuste vahel  —, võime uurida vastava psühholoogilise valdkonna funktsionaalset organisatsiooni ning püüda kognitiiv- ja neuroteaduse saavutuste abiga seletada teadvuse struktuuri selles vallas reduktiivselt. Selliselt tegutsedes oleme seletanud ära kogu fenomenaalse maailma struktuuri, kui mitte arvestada struktuurse koherentsuse printsiibi panust. Kuna me lähtusime sellest printsiibist kui eeldusest, siis ei ole me andnud seletust teadvusele endale, kuid oleme siiski seletanud ära suure osa ühe konkreetse fenomenaalse valdkonna eripäradest.

Sellise tõlgenduse korral võib struktuurse koherentsuse printsiibist saada selgroog projektile, mida Crick ja Koch[[290]](#footnote-291) nimetavad “kvaalide looduslooks”. Kui neuroteadus ei suudagi seletada elamuse olemasolu, suudab ta siiski seletada paljusid elamust puudutavaid fakte. Ta suudab kaudselt seletada elamuste sarnasus- ja erinevussuhteid; elamusruumide, näiteks maitse- või värvusruumi geomeetriat, elamusväljade, näiteks nägemisvälja struktuuri peeni üksikasju, üksikelamuste tajutavat paiknemist neil väljadel, elamuste intensiivsust, elamuste ​**[lk.236]**​ kestust, elamuste vahelisi assotsiatsioone ja palju muud. Nagu ütlevad Crick ja Koch: neuroteadus võib seletada kõiki neid elamuste omadusi, mida on võimalik objektiivselt edastada. Just nende tunnuste edastatavus näitab, et nad kajastuvad süsteemi füüsikalistes omadustes ning ühtlasi teadlikkuse omadustes. Teadvuse ja teadlikkuse omavaheline struktuurne koherentsus on alus, millele selline seletus vaikivalt või sõnaselgelt toetub. (Erilise tähtsuse omandab selline alus psühhofüüsikas; kõnelen sellest põhjalikumalt lõppmärkuses.) [[291]](#footnote-292).

Nende meetodite abil võime saada isegi aimu, mis tunne on olla nahkhiir! Funktsionaalne organisatsioon võib öelda meile palju nahkhiirele kättesaadava informatsiooni kohta — mida ta suudab eristada, kuidas ta asju liigitab, missugused on tema tajuvälja kõige olulisemad omadused jne. — ja kuidas ta seda informatsiooni kasutab. Selle kõige põhjal peaks olema võimalik luua üksikasjalik pilt teadlikkuse struktuurist nahkhiire kognitiivses süsteemis. Struktuurse koherentsuse printsiibi abil saame siis hea ettekujutuse ka nahkhiire elamuste struktuurist. Me ei saa teada kõike selle kohta, mis tunne on olla nahkhiir — meil ei ole näiteks selget ettekujutust tema elamuste seesmisest olemusest —, kuid siiski üsna palju. Akinsi (1993) huvitavat uurimust nahkhiirte vaimse elu kohta võib pidada üheks panuseks sellesse projekti.

Ka Cheney ja Seyfarthi raamat “Kuidas ahvid maailma näevad” [​*How Monkeys See the World*​] (1990) pakub vastust analoogilisele küsimusele, viies meid teise bioloogilise liigi siseilma. Õigupoolest rakendab see teos algusest lõpuni tausteeldusena struktuurse koherentsuse printsiipi: kirjeldades ühtesid või teisi funktsionaalseid protsesse ja nendega kaasnevat teadlikkuse struktuuri, kutsub ta meid tuletama sellest vastavat elamusstruktuuri. Meile juba tuttavatel põhjustel ei leia me niiviisi vastust küsimusele, mis Nagelit tegelikult vaevab, kuid see on siiski väljapaistev saavutus. Me ei pea seisma silmitsi teadvuse viimse müsteeriumiga iga kord, kui soovime tuua selgust mõnesse konkreetsesse fenomenaalsesse valdkonda.

#### Koherentsusprintsiibid kui episteemilised hoovad

Neuroteaduses, psühholoogias, etoloogias ja neile lähedastel aladel esitatakse empiirilistes uurimistöödes aeg-ajalt väiteid teadvuselamuse olemasolu kohta süsteemis. Kuigi teadvus jääb neis valdkondades enamasti tagaplaanile, leidub arvestatav hulk uurimistöid, kus empiiriliste tulemuste põhjal tehakse järeldusi teadvuselamuse kohta. Kuidas on see võimalik, hoolimata kõigist raskustest elamuse otsese vaatlemisega? Kui vaadelda on võimalik ainult füüsikalisi protsesse, mis annab siis kellelegi mingisugusegi õiguse sääraseid järeldusi teha?

Vastus saab olla ainult selles, et iga kord, kui keegi teeb empiiriliste tulemuste põhjal järeldusi elamuse kohta, rakendab ta ​**[lk.237]**​ mingit sildprintsiipi, mis seostab füüsikalised protsessid elamusega. Sildprintsiip annab füüsikalisel tasandil rakendatava kriteeriumi teadvuse olemasoluks süsteemis. Selline printsiip toimib episteemilise hoovana, mis võimaldab liikuda teadmiselt füüsikaliste protsesside kohta teadmistele elamuse kohta. Seda episteemilist hooba ennast ei ole võimalik katseliselt kontrollida, vähemalt mitte kolmanda isiku positsioonilt; ta toimib vaid omamoodi aprioorse tausteeldusena. Neid eeldusi ei öelda iga kord otsesõnu välja, kuid ainult nende abil on kuigagiviisigi võimalik seostada seesuguseid uurimistöid teadvuselamusega.

Sildprintsiipidel on siin nii otsustav tähtsus, et tasub nad selge sõnaga välja öelda. Igaüks, kes toetub mingile sildprintsiibile — see tähendab, igaüks, kes teeb väliste vaatluste põhjal järeldusi elamuse kohta — tegeleb “filosoofiaga”, sest sildprintsiibid ei ole ise katsetulemuste järeldused. Sellised printsiibid saavad toetuda ainult esimese isiku positsioonist tulenevatele kaalutlustele ja üldistele usutavusprintsiipidele. Tegelikkuses käivad need printsiibid igasuguste katsetulemuste eel, sest alles printsiibid ütlevadki meile, kuidas neid tulemusi tõlgendada. Loomulikult lähtutakse katselises uurimistöös alati mingitest aprioorsetest eeldustest, kuid siin on nende roll eriliselt kaalukas. Seepärast on oluline neid eeldusi hoolika analüüsi teel võimalikult hästi põhjendada. Kõike, mida ma olen käesolevas peatükis öelnud, võibki tõlgendada ka selles vaimus.

Mina olen soovitanud võtta sildprintsiibiks teadvuse ja teadlikkuse vahelise koherentsuse printsiibi: kui süsteem on mingist informatsioonist teadlik — selles mõttes et see on otseselt kättesaadav globaalseks juhtimiseks —, siis jõuab see informatsioon ka teadvusse. Ma kahtlustan, et kui keegi võtaks sildprintsiibid, mida empiiriliste uurimistööde tegijad ja tulemuste tõlgendajad kasutanud on, hoolika uurimise alla, siis selguks, et peaaegu kõik nad oleksid selle printsiibiga kooskõlas ja isegi sellest tuletatavad. Kõige tavalisemaks sildprintsiibiks on muidugi raporteeritavus kui elamuse kriteerium: vähemalt keelevõimeliste süsteemide puhul valitseb üldiselt seisukoht, et kui mingi informatsioon on raporteeritav, siis on ta ka teadvuses. Raporteeritavus on üks teadlikkuse erikujusid — informatsioon, mis on raporteeritav, on alati kättesaadav ka juhtimiseks —, ning seega on kõnealune kriteerium, kuigi tema rakendusala on ahtam, koherentsusprintsiibiga selgelt kooskõlas.

Aeg-ajalt lähtutakse ka teistsugustest kriteeriumidest; mõnikord soovivad teadlased öelda midagi elamuste kohta keeleta loomadel või inimestel, kelle raporteerimismehhanismid ei talitle normaalselt. Sellistel juhtudel loetakse elamuse parimaks tunnuseks tavaliselt mingi informatsiooni märgatavat mõju käitumise juhtimisele. Näiteks Logothetis ja Schall (1989) ütlevad, et nende töö teemaks on “subjektiivse nägemistaju neuronkorrelaatide” isoleerimine ahvidel. Siin on eeldatud, et läheduses liikuv objekt on kutsunud ahvil esile tajuelamuse, kui ta suudab ennustatavalt teha mingi silmaliigutuse või reageerida objekti liikumisele kangivajutusega. See ​**[lk.238]**​ ühildub jällegi suurepäraselt kriteeriumiga, mille annab teadlikkus või otsene kättesaadavus globaalseks juhtimiseks.

Mõni leiab võib-olla, et apelleerimine eeleksperimentaalsetele sildprintsiipidele rikub eksperimentaalteaduse puhtust[[292]](#footnote-293); ja tõepoolest, vajadus sääraste episteemiliste hoobade järgi ongi võib-olla põhjus, miks need teadused teadvusest nii sageli eemale hoiduvad. Kuid just selles paadis me juba kord oleme, ja parem on teha järeldusi nende printsiipide alusel kui neid üldse mitte teha. Kuid asjaomased printsiibid tasub selge sõnaga välja öelda ja hoolika analüüsi teel põhjendada, mitte neid kalevi all hoida. Nii jõuame arutlustes, mis viivad empiirilistele järeldustele teadvuselamuse kohta, suuremale selgusele.

#### Teadvuse füüsikalised korrelaadid

Mis on teadvuse neuraalsed ja infotöötluslikud korrelaadid? See on üks keskseid küsimusi teadvuse kohta empiirilistes uuringutes. Empiirilisi hüpoteese on esitatud mitmesuguseid. Näiteks Crick ja Koch (1990) on pakkunud välja hüpoteesi, et elamuse neuronkorrelaatideks on teatavad 40-hertsised võnkumised ajukoores. Üks Baarsi (1988) võimalikke tõlgendusi on, et globaalne töölaud on elamuse infotöötluslik alus ning et elamuse sisu vastab otseselt töölaua sisule. Farah (1994) väidab, et teadvus seondub “kõrgekvaliteediliste” representatsioonidega ajus. Libet (1993) on esitanud neuraalse “​kestusteooria​”, mille järgi teadvus seondub neuraalsete protsessidega, mis kestavad küllalt kaua, vähemalt umbes 500 millisekundit. Sarnases vaimus on esitatud palju teisigi hüpoteese.

Teadvuse ja teadlikkuse vaheline koherentsus lubab mõtestada suurt osa sellest tööst loomulikul viisil. Tähelepanuväärsel kombel on ka kõiki neis hüpoteesides välja pakutud kandidaate endid võimalik pidada üheks faktoriks teadlikkuses, s.t. otseses kättesaadavuses globaalseks juhtimiseks.​[[293]](#footnote-294) Crick ja Koch pakkusid oma võnked välja rolli tõttu, mida need nende arvates mängivad informatsiooni ühendamisel ja töömällu viimisel; ja muidugi ei olegi töömälu midagi muud kui süsteem, mis teeb selle sisu juhtimisele kättesaadavaks. Libeti piisava ajalise kestusega neuronaktiivsus võib asjasse puutuda just sellepärast, et see avaldab järjekindlalt ja ulatuslikult mõju teadlikkuseks tarvilikule kognitiivsüsteemile. Sama võib öelda ka Farahi “kõrgekvaliteediliste” representatsioonide kohta; on mõeldav, et “madalakvaliteedilised” representatsioonid ei pruugi olla suutelised avaldama tarvilikku kõikehõlmavat mõju kognitiivsüstemi talitlusele. Kõige selgem on see seos Baarsi globaalse töölaua puhul: töölaua mudeli mõte ongi mängida rolli globaalse ligipääsu ja juhtimise vahendamisel.

Seda asjaolu deflatiivselt tõlgendades võiks öelda, et öeldes “teadvus”, peavad need uurijad silmas vaid teadlikkust, mistõttu ​**[lk.239]**​ need kokkusattumused ei olegi nii huvitavad. Kuid minu arvates on konteksti põhjal selge, et enamik neist — vähemalt Crick ja Koch ning Farah, nähtavasti ka Libet — kõnelevad just teadvusest selle sõna täies fenomenaalses mõttes ja püüavad välja selgitada selle füüsikalisi korrelaate. Kõik nad ütlevad aeg-ajalt midagi niisugust, millest on näha, et kontseptuaalne vahetegemine teadvuse ja teadlikkuse vahel, nagu mina neid siin defineerin, oleks neile vastuvõetav.

Huvitavam tõlgendus olekski, et need uurijad kõnelevad teadvusest fenomenaalses mõttes, ning panna tähele, et kõik nende hüpoteesid on ühitatavad teadvuse ja teadlikkuse vahelise koherentsusega kui kõikehõlmava sildprintsiibiga. Tõepoolest, kõiki neid hüpoteese on võimalik tuletada koherentsusprintsiibist ja vastavatest empiirilistest andmetest. Tunnistades tausteeldusena koherentsusprintsiipi, tunnistame ühtlasi, et elamus on vahetus seoses otsese kättesaadavusega globaalseks juhtimiseks. Kui empiirilised andmed ütlevad, et 40-hertsised võnkumised teenivad ühel või teisel liigil (näiteks *Homo sapiens*​il) globaalset kättesaadavust, siis on meil põhjust arvata, et need võnkumised on sellel liigil üks elamuse korrelaate. Kui katseandmed näitavad, et küllalt suure ajalise kestusega aktiivsus aitab kaasa globaalsele kättesaadavusele, siis on meil põhjust uskuda, et seesugune aktiivsus on üks elamuse korrelaate. Ja nii edasi.

Loomulikult võib õigeks osutuda rohkem kui üks neist hüpoteesidest. Võib-olla aitavad nii võnkumised kui ka piisava kestusega aktiivsus kaasa globaalsele kättesaadavusele, ent erinevatel juhtudel, või tulevad nad mängu ligipääsu- ja juhtimisprotsesside eri staadiumides. Võib-olla soodustavad võnkumised kõrgekvaliteedilisi representatsioone globaalsel töölaual. Need on empiirilised küsimused. Kuid need hüpoteesid võivad osutuda ka ekslikuks. Võib selguda, et võnkumised ei mängi mingit erilist rolli mitte globaalses juhtimises, vaid ainult perifeersetes protsessides. Võib-olla on nende mõju hilisematele protsessidele ja käitumisele väga väike.

Väärib märkimist, et kui meil oleks põhjust uskuda, et võnked ei kuulu selles mõttes teadlikkuse juurde, siis oleks meil ka põhjust uskuda, et nad ei kuulu elamuse juurde. Kui selguks, et võnkumistel puudub eriline seos näiteks raporteerimisvõime ja teadvusega, siis kaoks korrelatsioonihüpoteesil pind jalge alt. Pole meil selle hüpoteesi kasuks ju sõltumatuid tõendeid: kõik meie tõendid toetuvad eeldatavale seosele raporteerimisvõime ja teadlikkusega. Kuna meil “elamusmeetrit” ei ole, siis tuleb jääda toetuma säärastele kaudsetele kriteeriumidele, milledest parimad ongi nähtavasti raporteerimisvõimest ja teadlikkusest lähtuvad kriteeriumid. Sellest järeldub, et meil saab olla empiirilisi tõendeid protsessi ​*N* ja teadvuse vahelise seose kohta ainult juhul, kui meil on juba tõendeid seose kohta ​*N* ja teadlikkuse vahel.

See asjaolu annab meile selge metoodika elamuse füüsiliste korrelaatide otsimiseks. Eeleksperimentaalsed kaalutlused viivad mõttele, et teadvuse peamiseks informatsiooniliseks korrelaadiks on teadlikkus või globaalne kättesaadavus. Empiirilised ​**[lk.240]**​ tulemused näitavad, et füüsikalised seisundid, mis ühe või teise liigi teadlikkuses mingit rolli mängivad, on teatavat tüüpi ​*N*​. Aprioorne sildprintsiip ja empiirilised tulemused viivad üheskoos oletuseni, et kõnealuse liigi puhul on ​*N* teadvuse füüsikaline korrelaat.

(Huvitaval kombel on Dennett [1993b] välja pakkunud peaaegu samasuguse metoodika. Ta ütleb, et konkreetsed oletused teadvuse aluse kohta (näiteks needsamad võnkumised) on tõepärased just sedavõrd, kuivõrd asjaomased protsessid mängivad rolli “tserebraalse staari” ilmumises. Ma võin siin temaga peaaegu kõiges nõustuda; vahe on selles, et mina eeldan ainult potentsiaalse tserebraalse staari olemasolu ning usun, et teadvuse ja staari vaheline seos on noomiline printsiip, mitte kontseptuaalne tõde. Tõepoolest, niipea kui keegi hakkab teadvust empiiriliselt uurima, läheb tal tarvis mitteempiirilist sildprintsiipi, mille varal tõlgendada füüsikalisi tulemusi teadvuselamuse seisukohalt. Reduktsionisti jaoks on see kontseptuaalne tõde, omadusdualisti jaoks aga noomiline printsiip, mis põhineb esimese isiku kaalutlustel ja usaldusväärsel analüüsil. Kuid paljud punktid, mida ma siin esitan, jäävad kehtima mõlemal juhul.)

On vististi loomulik oletada, et keskseks korrelatsiooniks füüsiliste protsesside ja elamuse vahel on teadvuse ja teadlikkuse vaheline koherentsus. Otseselt ei loo elamust mitte võnkumised, piisava kestusega aktiivsus või kõrgekvaliteedilised representatsioonid, vaid protsess, mis tagab otsese kättesaadavuse globaalseks juhtimiseks. Iga spetsiifilisem füüsikaline seisund kvalifitseerub korrelatsiooniks ainult kuivõrd ta mängib rolli globaalses kättesaadavuses; seega on spetsiifilisemad korrelatsioonid mainitud kõikehõlmava korrelatsiooni järelmid.

Selliseid korrelatsioone võib olla palju. Näiteis võivad eri tüüpi füüsikalised protsessid teenindada erinevate modaalsuste kättesaadavust. Samuti võivad töötlemisprotsessi eri etappidel mängida rolli erinevad korrelaadid; isegi kui vaadelda ainuüksi nägemiselamust, võib nägemiskorteksis toimida ühtliiki korrelatsioon ja kaugemates piirkondades teistsugune. Loomulikult ei ole mingit garantiid, et isegi ühe kindla modaalsuse piires toimib mingi lihtne neuraalne korrelatsioon. Võib-olla ei saagi elamust teenindavaid protsesse, mis kulgevad nägemiskorteksis, otseselt iseloomustada; võib selguda, et neid saabki rühmitada ainult nende funktsionaalse rolli järgi (s.t. toetudes asjaolule,, et nad teenindavad teadlikkust). Kuid me võime vähemalt loota, et suudame leida neile selgema iseloomustuse: kui mitte tajukorteksis, siis ehk mõnes töötlustrajektoori hilisemas punktis.

Võib ka juhtuda, et me leiame korrelatsioone neurontasemest kõrgematel tasemetel. Sobiva kognitiivse või informaatilise mudeli abil võime leida mingi viisi kirjeldada teadlikkust tagavaid ja seega elamust kandvaid üksusi informaatika mõistetes. Kõige lihtsam selline kirjeldus on muidugi tautoloogiline — teadvust teenindavad need protsessid, mis tegelevad globaalse ligipääsu, juhtimise ja sõnaliste ütluste andmisega. Kuid sisukas kognitiivne mudel võib anda meile ka kirjeldusi, mis pole nii ​**[lk.241]**​ tautoloogilised. Võib-olla on selle abil võimalik iseloomustada suhteliselt lokaalsetes mõistetes seda informatsioonitüüpi, mis süsteemi üldist ülesehitust arvestades osutub mängivat olulist globaalset rolli. Näiteks võib tuua Shallice’i (1972) idee, mille järgi teadvuse sisu vastab mingite kindlate tegevussüsteemide “selektorsisendite” sisule. Selle mudeli järgi määravad selektorsisendid kindlaks, millised tegevussüsteemid muutuvad “dominantseteks”, s.t. omandavad mingi rolli globaalses juhtimises. Kui see on nii, siis aitavad selektorsisendid kaasa teadlikkuse tekkele ja ning on tõenäoliselt teadvuselamuse korrelaadid.

Teised informatsioonilised korrelatsioonid jäävad kuhugi tautoloogiliste ja mittetautoloogiliste vahepeale. Üheks näiteks võib tuua Baarsi globaalse töölaua. Me võime määratleda globaalset töölauda kui omamoodi “virtuaalset” ruumi, mis on täpses vastavuses üle kogu süsteemi levivate sisudega; sellisel juhul ühtib töölaua sisu peaaegu definitsiooni järgi teadvuse sisuga. Baarsi mudelil on empiirilist haaret aga rohkem: tema järgi on töölaud ühtne (ja vähemasti informaatiliselt lokaliseeritav) süsteem, mis integreerib ja levitab informatsiooni. See võib osutuda ekslikuks ning seega on ettepanekul empiiriline sisu, mida katselised uurimused võivad kinnitada või ümber lükata. Kuid oma kirjelduselt on töölaud ikkagi küllalt sarnane teadlikkusega — sellest tulebki hüpoteesi mõningane aprioorne hõng; mõnikord tundub, et selle mudeliga annaks ühildada peaaegu ükskõik milliseid empiirilisi tulemusi. (Loomulikult esitab Baars töölaua talitluse kohta ka palju spetsiifilisemaid väiteid, mis kannavad olulist empiirilist sisu.) Mõningat tautoloogilisust on tunda ka Farahi “kõrgekvaliteedilise representatsiooni” hüpoteesil, kuigi see oleneb sellest, kuidas me seda kõrgekvaliteedilist representatsiooni defineerime: kui see on lihtsalt võime täita olulist globaalset rolli, siis on aprioorsuse hõng tugev, ent kui definitsiooni aluseks võtta representatsiooni moodustumise viis, jääb see palju nõrgemaks.

Isegi kognitiivsel tasandil pole erilist põhjust uskuda, nagu oleks elamuse ainsaks aluseks üks spetsiaalne mehhanism. Schacter (1989) oletab küll, et teadvuse taga võib olla omaenda eriline mehhanism, näiteks mingi moodul, kuid see on ainult üks võimalus. Võib selguda, et igale rollile globaalses juhtimises vastab alati mingi keskne mehhanism, mis sellele kaasa aitab (näiteks Baarsi globaalne töölaud), kuid täpsete andmete puududes võib sama tõenäoliseks pidada ka võimalust, et vajaliku kättesaadavuse tagavad isegi üheainsal liigil või katseisikul eri aegadel väga erinevad protsessid.

Mõnikord esitatakse füüsikaliste korrelaatide asjas ka tugevamaid järeldusi kui need, millest olen kõnelnud mina. Näiteks kui me avastaksime, et 40-hertsised võnkumised on teadaolevatel juhtudel elamuse aluseks, kas ei võiks siis oletada, et 40-hertsised võnkumised ongi elamuse viimne alus? Võib-olla sünnib neist võnkumistest elamus isegi siis, kui nendega ei kaasne ​**[lk.242]**​ teadlikkus, ja võib-olla puudub veatult talitlevatel süsteemidel, kus need võnkumised puuduvad, ka elamus? Kuid selline järeldus oleks põhjendamatu. 40-hertsiseid võnkumisi hakati oluliseks pidama sellepärast, et neil on mingi seos teadlikkusega; meil pole põhjust arvata, et neis on midagi erilist ka siis, kui nad seda rolli ei mängi. Kindlasti pole mingit põhjust uskuda, et 40-hertsised võnkumised kutsuksid ka katseklaasis esile samataolisi elamusi nagu minul! Ja isegi vahepealsetel juhtudel, näiteks loomade või anestetiseeritud süsteemide puhul, oleks riskantne järeldada nende võnkumiste olemasolust midagi elamuse kohta, kui just võnkumiste toimumine ise ei anna meile põhjust uskuda mingitlaadi teadlikkuse olemasolu.

Üldiselt ei maksa oodata, et neilt empiirilistelt meetoditelt universaalseid psühhofüüsikaprintsiipe, mis osutuksid ühitamatuteks nendega, millest alustasime meie. Nad võivad aidata kaasa spetsiifilisemate printsiipide leidmisele, mis kehtivad mingi ühe bioloogilise liigi piires, kuid sellisel juhul piisab nende tuletamiseks juba teada olevate​[[294]](#footnote-295) sildprintsiipide otsesest rakendamisest. Kuna need printsiibid kannavad füüsikalistest andmetest elamuse kohta järelduste tegemisel kogu koormust, siis ei saa andmed toetada järeldusi, mis neile printsiipidele vastu räägiksid.

Seega ei maksa pidada teadvuse neuronkorrelaadi otsinguid teeks universaalse teooria püha graali juurde. Me võime saada siit hinnalist abi teadvuse mõistmiseks konkreetsetel juhtudel, näiteks inimese puhul: teadvust kandvate protsesside tundmaõppimine aitab meil kindlasti mõista näiteks teadvuse struktuuri ja dünaamikat. Kuid füüsikaliste protsesside seostamisel teadvuselamusega jäävad alati keskset rolli mängima eeleksperimentaalsed koherentsusprintsiibid.

### 5. Koherentsus kui psühhofüüsikaseadus

Olen käsitlenud koherentsust seni peamiselt enam-vähem tuttavatel juhtudel, sealhulgas inimesel ja teistel bioloogilistel süsteemidel. Kuid on loomulik oletada, et koherentsusprintsiipidel on üldkehtiva seaduse staatus. Kui teadvusega kaasneb alati teadlikkus ja ümberpöördult nii minu kui ka kõigi teiste inimeste puhul, siis võib siin aimata midagi süstemaatilist. Mingi seadusega sarnanev korrelatsioon on neil tuttavatel juhtudel kindlasti olemas. Seega võime püstitada hüpoteesi, et see koherentsus on looduse seadus: igas süsteemis kaasneb teadvusega teadlikkus ja ümberpöördult. Sama kehtib struktuurse koherentsuse printsiibi kohta tema täielikus mõttes. Tähelepanuväärne korrelatsioon teadvuse struktuuri ja teadlikkuse struktuuri vahel tundub olevat liiga spetsiifiline, et olla juhuslik. On loomulik oletada, et selle aluseks on seadus: iga süsteemi puhul aegruumi ​**[lk.243]**​ igas punktis peegeldab teadvuse struktuur teadlikkuse struktuuri, ja ümberpöördult.

Sellesarnased seadused annaksid olulise panuse teadvuseteooriasse. Esialgu teame me ainult niipalju, et teadvus kasvab mingil moel välja füüsikalisest maailmast, kuid me ei tea, missuguste füüsikaliste omaduste alusel see toimub — see tähendab, me ei tea, millised omadused on mängus seose füüsikapoolses otsas. Koherentsusseadusi arvestades võime anda osalise vastuse: teadvus tekib tänu teadlikkust kandvale funktsionaalsele organisatsioonile. Struktuurse koherentsuse printsiibi abil võime isegi jõuda üsna konkreetse arusaamiseni selle superventsussuhte komponentidest: teadvus mitte ainult ei sünni teadlikkusest, vaid ka tema struktuuri määrab kindlaks teadlikkuse struktuur.

Muidugi, tõenäoliselt ei ole see seadus fundamentaalne psühhofüüsikaseadus. Fundamentaalseadused ühendavad omadusi, mis on fundamentaalsemad või millel on vähemalt puhtam definitsioon kui säärasel kõrgtaseme konstruktil, nagu seda on teadlikkus. Kuid mitte kõik seadused ei ole fundamentaalsed. Võib isegi selguda, et koherentsusprintsiibid ei ole ranged seadused; äärealadel võib olla erandeid, eriti kui arvestada teadlikkuse mõiste määratluse lünklikkust. Ent isegi kui need seadused ei ole ei fundamentaalsed ega ranged, seavad nad siiski tugevaid kitsendusi kõigile fundamentaalsetele psühhofüüsikaseadustele. Igaüks, kes pakub välja teadvuseteooria, millest ei saa tuletada koherentsusprintsiipe, jääb sellega kindla peale hätta. Ja ümberpöördult: kui pakutav fundamentaalne psühhofüüsikaseadus on lihtne ja hästi põhjendatud ning kui temast saab tuletada koherentsusprintsiibid, siis on põhjust kaaluda tema aktsepteerimist.

Mis alusel võiks siis koherentsusprintsiipe seadusteks tunnistada? Peamise põhjenduse niisuguse otsuse kasuks annavad meile korrelatsioonid tuttavatel juhtudel — minu jaoks viimselt minu enda juhul. Ilmselged korrelatsioonid teadlikkuse ja teadvuse vahel on minu enda juhul nii üksikasjalikud ja tähelepanuväärsed, et siin peab olema midagi enamat kui juhuslik reeglipära. Selle aluseks peab olema mingi seadus. Küsimus on ainult selles, missugune see seadus on. Sellest seadusest peab tulenema, et minu enda juhul kaasneb teadlikkusega alati teadvus ja ümberpöördult, ning et nende struktuurid on omavahel vastavuses. Minu pakutud koherentsusprintsiibid rahuldavad seda nõuet. Kas seda teeks ka mõni teine printsiip?

Tundub olevat väga tõenäoline, et mingit liiki teadlikkus on teadvuseks tarvilik. Igatahes käib kõigi teadvuse juhtudega, mida mina tunnen, kindlalt kaasas teadlikkus. Ei paista olevat palju põhjust uskuda, et võib esineda juhtumeid, kus teadvusega ei kaasne mingeid funktsionaalseid protsesse. Kui sääraseid ongi, siis ei ole meil nende kohta mingeid tõendeid, isegi mitte kaudseid, ja neid ei saaks juba põhimõtteliselt ollagi. Seepärast on mõistlik rakendada Occami habemenuga ja oletada, et kus on teadvus, seal on ka teadlikkus. Kui me selles eksime — kui näiteks paigallseisval elektronil ​**[lk.244]**​ on sama rikkalik teadvuselu nagu mõnel Proustil —, siis ei ole meil kuidagi võimalik sellest midagi teada.

Küsimus, kas teadlikkus on teadvuseks ka piisav, on keerulisem. Kui tunnistada teadlikkus tarvilikuks, siis peab igal alusseaduse kandidaadil olema kuju “Teadvus tekib teadlikkuse pluss veel millegi olemasolul”. Selleks, et reeglipärad oleksid minu enda juhul seletatavad, peab mingi sellekujuline printsiip olema alusseadusest vähemasti tuletatav. Seega tuleb nüüd veel küsida: Mis see teadlikkusele lisanduv “miski” on — või ei olegi peale teadlikkuse midagi tarvis?

Nimetagem seda hüpoteetilist lisakomponenti *X-faktoriks*​. Mul on teadvus kas ainuüksi tänu teadlikkusele, või siis tänu teadlikkusele ja X-faktorile. X-faktoriks võib väga hästi olla ükskõik missugune omadus, kui see on mul praegu olemas, ja eriti, kui see on mulle omane kogu minu elu vältel. Võib-olla tuleb X-faktor kuidagi kodakondsusest, nii et teadlikkusest sünnib teadvus ainult austraallastel; võib-olla asukohast, nii et teadlikkusest sünnib teadvus ainult saja miljoni miili raadiuses mõnest kinnistähest; võib-olla aga hoopis isikuidentsusest, ja sellisel juhul sünnib teadlikkusest teadvus ainult David Chalmersil.

Kõik need seadused oleksid ühildatavad tõenditega, mis mul on, ja seletaksid korrelatsiooni. Miks tunduvad nad siis nii arutud? Sellepärast, et kõigil neil juhtudel on X-faktor täiesti meelevaldne. Ei ole mingit põhjust arvata, et teadvus peaks sõltuma taolistest asjadest; need tunduvad pigem ülearuste ja kõrvaliste lisandustena. Asi pole nii, nagu mängiks X-faktor mingit rolli mõne teadvusega seonduva fenomeni seletamises. Vähemalt võib teadlikkus aidata seletada fenomenaalseid otsustusi, mis on teadvusega tihedas kokkupuutes, ja seega on meil põhjust uskuda, et siin on mingi seos. Kuid ülalloetletud X-faktorid ilmuvad justkui tühjalt kohalt. Miks peaks universum olema selline, et teadlikkusest sünnib teadvus ühes ja ainult selles ühes inimeses? Säärane maailma olemisviis oleks kummaline, meelevaldne.

Sama võib öelda ka tõepärasemate X-faktorite kohta, mida võidakse tõemeeli esitada. Üheks sääraseks loomulikuks X-faktori kandidaadiks on rakupõhine bioloogia või koguni inimlik neurofüsioloogia. Ja tõepoolest ongi mõni autor oletanud, et teadvust saab olla vaid olenditel, kellel on õige bioloogiline struktuur. On ka oletatud, et teadvus sünnib funktsionaalsest organisatsioonist ainult juhul, kui seda organisatsiooni ei ole implementeeritud homunkuluse meetodil nagu ülaltoodud näites Hiina elanikkonnaga. Ent ka niisugused X-faktorid on samavõrra meelevaldsed. Nad teevad seadusi vaid keerulisemaks, andmata midagi vastu. Miks peaks maailm olema niisugune, et teadlikkusest sünnib teadvus ainult mingi kindla bioloogiaga olenditel, või niisugune, et homunkuluse meetod ei tööta? Oma kõrvaliste lisaviguritega jätavad need hüpoteesid pentsiku mulje.

Miks võiks keegi üldse uskuda mingisse X-faktorisse? Ma arvan, et sellistel uskumustel on loomulik, ehkki eksitav põhjus. Mingi väga sügav intuitsioon ütleb meile, et teadvus ​**[lk.245]**​ on midagi enamat ja kõrgemat kui funktsionaalne organisatsioon. Seda ütleb loomulikult ka minu intuitsioon: teadvus on lisafakt, mille seletamiseks mitte mingisugune funktsionaalne organisatsioon ei saa loogiliselt piisav olla. Meil on ka loomulik kalduvus uskuda, et kõik on füüsiline ning et teadvusele peab leiduma nii- või teistsugune füüsikaline seletus. See kahekordne surve vallandab meis loomuliku reaktsiooni: midagi tuleb veel lisada, ja see lisandus peab olema füüsikaline. Üheks loomulikuks kandidaadiks selle lisakomponendi kohale ongi inimese bioloogia. Ja võib tõepoolest arvata, et nii saame üle lõhest funktsionaalse organisatsiooni ja inimese bioloogia vahel.

Kuid see on täiesti ekslik arusaam. Bioloogia lisamine pildile ei aita esialgse probleemi lahendamisele mitte sugugi kaasa. Lünk on sama suur kui enne: teadvus tundub olevat midagi enamat ja kõrgemat ka bioloogiast. Nagu ülal näidatud, ei piisa teadvuse seletamiseks mitte mingitest füüsikalistest faktidest. X-faktor ei saa meid kuidagi aidata; me otsime oma raskusele lahendust valest kohast. Häda tuli sellest, et me eeldasime materialismi kehtivust. Niipea kui me tunnistame materialismi vääraks, saab selgeks, et füüsikalise X-faktori otsimine ei anna teadvuse seletamiseks midagi ja otsida tuleb hoopis “Y-faktorit”, veel midagi lisaks füüsikalistele faktidele. Sellise Y-faktori võime leida, postuleerides redutseerimatute psühhofüüsikaseaduste olemasolu. Niipea kui me need oma mudelile lisame, ei pea me enam kõrvale heitma intuitiivset muljet, mis ütleb, et teadvus on lisafakt, ja raskus kaob.

Soov leida füüsilist X-faktorit on jäänuk soovist saada kaht head korraga: tunnistada teadvust, kuid jääda ka materialistiks. Kui tunnistada, et teadvus on mittefüüsikaline lisafakt ning et on olemas sõltumatud psühhofüüsikaseadused, osutub X-faktor täiesti ülearuseks. Sõltumatu keha–vaimu seose ja sellele lisaks veel​ X-faktori otsimine tähendab nõuda kaht hüvet seal, kus piisab ühestainsast.

Seega ei jää X-faktorile teadvuseteoorias mingit seletavat rolli, vaid ta teeb pilti ainult segasemaks. Iga selline faktor teeb fundamentaalseadusi keerulisemaks kui tarvis. Võrreldes lihtsa pildiga, kus teadvus sünnib teadlikkusest, tundub universum, kus teadvus sõltub mingist omaette X-faktorist, üsna ebamõistlikuna. Sama hästi võiks Newtoni seadustes olla klausel, mis ütleb, et igale jõule leidub võrdne ja vastupidine jõud, kui vastasmõjustuvad objektid ei ole kullast. Lihtsusprintsiibid dikteerivad, et parim hüpotees on see, mis ei nõua X-faktorit ja kus teadlikkusest sünnib teadvus ilma mingite kitsendusteta.

Mõned lugejad kahtlevad kindlasti veel nüüdki funktsionalistlikus järelduses, milleni ma olen jõudnud, olgugi see funktsionalism dualistlik. On tõsi, et X-faktorite argument on mõnevõrra habras ja eeldused, millele ta toetub, lähtuvad suuresti püüdest asja lihtsustada. Järgnevas peatükis esitan selle järelduse toetuseks konkreetsemaid argumente ning näitan mõtteliste katsete abil, et teadvusega olendi täpsel funktsionaalsel koopial oleks ka teadvuselamus ​**[lk.246]**​ täpselt samasugune. Esialgu aga tähendan, et need kaalutlused annavad seesugusele funktsionalismile vähemasti tugeva *prima facie* põhjenduse.

Siinkohal tasub korraks peatuda ning heita ülevaatlik pilk üldisele epistemoloogilisele skeemile. See, millega me siin tegeleme, on põhimõtteliselt abduktsioon: me märkame silmatorkavaid reeglipäraseid vastavusi teadvuse ja teadlikkuse vahel meie enda juhul ning postuleerime kõige lihtsama võimaliku alusseaduse. See on samasugune arutluskäik nagu see, mida me kohaldame füüsikalisi teooriaid formuleerides ning isegi kritiseerides skeptilisi hüpoteese põhjuslikkuse ja välismaailma kohta. Kõigil neil juhtudel on aluseks eeldus, et maailm on lihtne ja mõistlik koht. Selle eelduseta võib mõelda mida tahes, seda eeldust tunnistades aga asetub kõik omale kohale.

Tundub, et minu siinne vaateviis annab parima võimaliku lahenduse ka teiste vaimude probleemile. Me märkame reeglipäraseid vastavusi ühelt poolt elamuste ning teiselt poolt füüsiliste ja funktsionaalsete seisundite vahel meis endis, postuleerime nende seletamiseks lihtsad ja homogeensed alusseadused ning nende seaduste abiga järeldame teadvuse olemasolu teistel. Võib-olla toetubki usk sellesse, et ka teistel on teadvus, mingile selletaolisele mõttekäigule, võib-olla ka mitte, kuid igatahes tundub, et nii omandab see uskumus mõistliku põhjenduse.

Huvitav on spekuleerida selle üle, mida saab meie koherentsusprintsiipidest järeldada teadvuse olemasolu kohta väljaspool inimliiki, ka meist palju lihtsamatel organismidel. Küsimus on ebaselge, sest me oskame oma arusaama teadlikkusest selgesti määratleda ainult inimlikule keerukusele lähenevate juhtude puhuks. Tundub olevat mõistlik öelda, et koeral ja isegi hiirel on olemas teadlikkus. (Eneseteadlikkust võib-olla mitte, aga see on juba teine küsimus). Näiteks tundub olevat mõistlik öelda, et koer on tuletõrjehüdrandist teadlik, kui mõista sõna “teadlik” oma baastähenduses. Koera juhtsüsteemidel on kindlasti ligipääs informatsioonile hüdrandi kohta ja nad võivad kasutada seda käitumise juhtimiseks. Koherentsusprintsiibi põhjal tundub tõenäoline, et koeral on hüdrandist ka elamus, mis ei ole meie visuaalsest maailmaelamusest väga erinev. See on kooskõlas terve mõistusega; siin püüan ma seda tervemõistuslikku mõttekäiku vaid pisut põhjalikumalt sõnastada.

Sama kehtib arvatavasti ka hiirte ja isegi kärbeste kohta. Kärbsel on mingites piirides pertseptuaalne ligipääs informatsioonile oma ümbruse kohta, tajude sisu imbub arvatavasti tema kognitiivsüsteemi ja on käitumise suunamiseks kättesaadav. On vist mõistlik oletada, et seda võib pidada teadlikkuseks ning et koherentsusprintsiibi järgi kaasneb sellega ka midagi elamusetaolist. Umbes siinkohal läheb lugu aga keeruliseks. Ahvatlev oleks laiendada koherentsust infotöötluse skaalal veelgi allapoole, kuid varem või hiljem hakkab teadlikkuse mõiste alt vedama ega suuda oma ebamäärasuse tõttu meid seletamistöös enam aidata. Siin ma selle üle rohkem ei spekuleeri, kuid tulen küsimuse juurde [allpool](#ORG.8.4.1) tagasi.

## ​**[lk.247]**​ 7. peatükk. Kvaalide olematus, kvaalide hääbumine, kvaalide hüplemine

### 1. Organisatsioonilise invariantsuse printsiip

Kui teadvus kasvab välja füüsilisest maailmast, siis milliste füüsikaliste omaduste varal see toimub? Eeldatavasti niisuguste, mis võivad ajus realiseeruda; kuid ei ole ilmne, millised need nimelt on. Ühed pakuvad biokeemilisi, teised kvantfüüsikalisi omadusi; paljud autorid on väljendanud selles küsimuses nõutust. Üks loomulikke oletusi on, et teadvus sünnib aju funktsionaalse organisatsiooni tagajärjena. Sellisel juhul ei ole aju keemiline ja isegi kvantfüüsikaline substraat teadvuse tekkimise seisukohast oluline. Oluline on aju abstraktne põhjuslik organisatsioon, mis võib realiseeruda väga erinevates füüsilistes substraatides.

Funktsionaalset organisatsiooni on kõige parem mõista kui süsteemi erinevate osade ning võib-olla ka süsteemi osade ning väliste sisendite ja väljundite vahelist abstraktset põhjuslikku vastasmõju​. Funktsionaalne organisatsiooni defineerimiseks tuleb teada 1) abstraktsete komponentide arvu, 2) iga komponendi võimalike seisundite arvu ning 3) sõltuvussuhete süsteemi, mis ütleks täpselt, kuidas iga komponendi seisund sõltub kõigi komponentide senistest seisunditest ja süsteemi sisenditest ning kuidas süsteemi väljundid sõltuvad komponentide senistest seisunditest. Kõik muu peale komponentide arvu ja sõltuvussuhete jääb definitsioonist välja.

Füüsiline süsteem on ühe või teise funktsionaalse organisatsiooni *realiseering*​, kui teda saab jagada vajalikuks arvuks füüsilisteks komponentideks, millest igaühel on vajalik arv seisundeid ja millede sisendite ja väljundite vahelised põhjuslikud sõltuvussuhted peegeldavad täpselt ​**[lk.248]**​ asjaomase funktsionaalse organisatsiooni definitsioonis loetletud sõltuvussuhteid. (Sisuliselt samaväärse, kuid formaalsema definitsiooni annab Chalmers (1994a, 1994b), ja sellest leidub kokkuvõte ka käesoleva raamatu [9. peatükis](#PARA.169-1), kuid esialgu piisab ka siintoodud mitteformaalsest sõnastusest).

Üht ja sama funktsionaalset organisatsiooni võivad realiseerida erinevad füüsilised süsteemid. Näiteks organisatsioni, mida aju realiseerib neuronite tasandil, oleks põhimõtteliselt võimalik realiseerida ka süsteemis, mis koosneb ränist. Aju funktsionaalse organisatsiooni abstraktne kirjeldus ei ütle midagi süsteemi osade füüsilise koosseisu ja põhjuslike seoste implementatsiooni kohta. Oluline on vaid osade olemasolu ja nende seisundite vahelised sõltuvussuhted.

Füüsilise süsteemi funktsionaalsel organisatsioonil on mitu eri taset. Tasemete arv oleneb sellest, kui peenelt me individualiseerime süsteemi osi ja kui peenelt me eristame nende osade seisundeid. Näiteks väga jämedalt võttes võib öelda, et aju poolkerade vastastikku sõltuvate seisundite sobiva valiku korral võib aju kaht poolkera vaadelda kui lihtsa kahe komponendiga organisatsiooni realiseeringut. Enamasti aga on mõistlik vaadelda kognitiivseid süsteeme peenemal tasemel. Kui meid huvitab kognitsioon, valime vaatlustaseme tavaliselt küllalt peene, et suuta kindlaks määrata ajuga seonduvad käitumisvõimed, seletamaks käitumist mingi sobiva täpsusastmega. Organisatsiooni liiga jämedatasemelise kirjelduse korral (nagu äsjakirjeldatud kaheosalise organisatsioon) ei ole võimalik käitumisvõimeid kindlaks teha, kuna mehhanismid, mis käitumist juhivad, jäävad kirjelduse pragude vahele; teisel sama lihtsal süsteemil võib olla samasugune organisatsioon, kuid teistsugune käitumine. Ent piisavalt peenel tasandil, võib-olla neurontasandil, hakkab funktsionaalne organisatsioon käitumisvõimeid määrama. Isegi kui asendada meie neuronid ränikiipidega, produtseeriks süsteem samasuguse käitumise, senikaua kui nende kiipide põhjuslike vastasmõjude muster oleks samasugune nagu neuronitel.

Siitpeale vaatleme süsteemide funktsionaalset organisatsiooni alati piisavalt peenel tasemel, et sellest piisaks käitumuslike võimete määratlemiseks. Niisugust organisatsiooni nimetame edaspidi *peeneteraliseks*​ funktsionaalseks organisatsiooniks. Illustratiivsetel eesmärkidel hakkan ma aju organisatsiooni vaatlema enamasti neurontasemel, kuigi võib-olla piisaks ka mingist kõrgemast tasemest; pole ka võimatu, et tuleb laskuda mingile madalamale tasemele. Kuid minu argumendid on üldistatavad ega sõltu vaatlustasemest. Lisaks tuleb edasise käsitluse eesmärkidel postuleerida, et selleks, et kahe süsteemi funktsionaalsed organisatsioonid ühtiksid, peavad nad olema õigetel aegadel teineteisele vastavates seisundites: kuigi minu magava teisiku kohta võib ehk mingis avaras mõttes öelda, et tema organisatsioon ühtib minu omaga, allpool nõutavas ranges mõttes see nii ei ole. Süsteeme, millede funktsionaalsed organisatsioonid selles ranges mõttes ühtivad, nimetan ma edaspidi *funktsionaalseteks isomorfideks*​.

Ma väidan, et teadvuselamus kasvab välja peeneteralisest funktsionaalsest organisatsioonist. Konkreetsemalt öeldes: ma püüan põhjendada *organisatsioonilise* ​**[lk.249]**​ *invariantsuse printsiipi*​, mis ütleb, et igal süsteemil, mille peeneteraline funktsionaalne organisatsioon ühtib mingi teadvuselamustega süsteemi omaga, on ka kvalitatiivselt samasugused elamused. Selle printsiibi kohaselt on teadvus organisatsiooniline invariant — omadus, mis on kõigil funktsionaalsetel isomorfidel konstantne. Pole oluline, kas seda organisatsiooni realiseerivad ränikiibid, Hiina elanikkond või õllepurgid ja pingpongipallid. Kui funktsionaalne organisatsioon on õige, determineerib ta ka teadvuselamuse.

Seda teesi on sageli seostatud reduktiivfunktsionalistliku vaatega teadvusele, näiteks seisukohaga, et teadvuseks ei ole tarvis midagi muud kui sellekohast funktsionaalset seisundit. Niisugusest vaatest tuleneks invariantsusprintsiip loomulikul teel. Kuid seda printsiipi võib pidada tõeseks ka sõltumatult. Nii nagu võib uskuda, et teadvus kasvab välja füüsilisest süsteemist, kuid ei ole ise füüsiline seisund, võib uskuda ka, et teadvus kasvab välja funktsionaalsest organisatsioonist, kuid ei ole ise funktsionaalne seisund. Just niisugust positsiooni ma siin kaitsengi; seda võiks nimetada *mittereduktiivseks funktsionalismiks*​, ja seda võib vaadelda kui üht võimalust funktsionalismi ja omadusdualismi ühendamiseks.

Järgnevas ei huvita mind kuigivõrd äsjakirjeldatud positsiooni mittereduktiivsed aspektid, sest minu peamine ülesanne on põhjendada invariantsusprintsiipi. Minu argumendid on võib-olla isegi reduktiivsetele funktsionalistidele vastuvõetavad. Kuigi nad reduktiivseid järeldusi täielikult ei kinnita, toetavad nad siiski pigem seda positsiooni teiste reduktiivsete seisukohtade vastu, näiteks selle, mis peab teadvust mingiks biokeemiliseks omaduseks. Loomulikult arvan ma, et kõik reduktiivsed seisukohad on viimaks ikkagi ebaõiged, kuid järgnev arutelu ei sõltu sellest küsimusest kuigivõrd.

Põhimõtteliselt olen ma üht invariantsusprintsiibi erikuju juba püüdnud põhjendada, nimelt [6. peatükis](#PARA.131-6) “X-faktori” argumendiga. Käesolevas peatükkis aga püüan ma põhjendada seda printsiipi palju otsesemal teel — mõtteliste eksperimentide abil.

####  Kvaalide olematus ja kvaalide inversioon

Invariantsusprintsiipi ei tunnista kaugeltki kõik. Tema vastu on sõna võtnud paljud nii dualistlike kui ka materialistlike vaadetega autorid. Paljud on uskunud, et selleks, et süsteemil oleks teadvus, peab tal olema õige biokeemiline koostis. See tähendaks, et metallist robotil ja ränikiipidega arvutil ei saa iial olla elamusi, ükskõik missugune oleks nende põhjuslik organisatsioon. Teised on möönnud, et sobiva organisatsiooniga robotil või arvutil võib teadvus ollagi, kuid tema elamused oleksid meie omadest täiesti erinevad.

Vastavalt neile kahele positsioonile on invariantsusprintsiibi vastu esitatud kahesuguseid argumente. Esimesse liiki kuuluvad *olematute kvaalide* [​*absent qualia*​] argumendid, mis kirjeldavad mingi kindla funktsionaalse organisatsiooni eriti veidraid realiseeringuid mõnes süsteemis, mis on nii eriskummaline, ​**[lk.250]**​, et tema puhul on loomulik oletada teadvuselamuse kvaliteetide (kvaalide) puudumist. Laialt tuntud on Blocki (1978) näide, kus meile omane organisatsioon realiseerub mingi riigi elanikkonnas (nagu [3. peatükis](#PARA.64-10)). Niisuguses realiseeringus ei saa ju, väidetakse, ometi sündida teadvuselamus. Ja kui see on nii, siis ei saa teadvus sündida funktsionaalsest organisatsioonist.

Teise argumentide klassi kuuluvad *invertkvaalide* või *invertspektri* argumendid. Need ütlevad, et kui meie funktsionaalne organisatsioon realiseeruks teistsuguses füüsilises substraadis, võiksid süsteemil elamused olla, kuid need oleksid teistsugused elamused. Seal, kus meil on punased elamused, võivad temal olla sinised elamused, jne. Sageli maalivad nende argumentide esitajad meile ette keerulisi stsenaariume, kus me ühel hommikul pärast ajuoperatsiooni ärgates näeme punase asemel sinist, kuigi meie funktsionaalne organisatsioon pole muutunud.

Paljud autorid, kes püüavad põhjendada kvaalide puudumise ja kvaalide inversiooni võimalikkust, on esitanud argumente ainult loogilise võimalikkuse kasuks; funktsionalismi reduktiivse vormi kummutamiseks sellest piisab. [3. peatükis](#ORG.3.1.2) kasutasin ma selliseid argumente ka ise. Siinse peatüki vastuargumentidega ei taha ma nende autoritega polemiseerida; kõne all on funktsionalismi nõrgem vorm, mis ei puuduta loogilise võimalikkuse küsimusi.

Peatüki põhiküsimus on: kas kvaalide puudumine ja kvaalide inversioon on looduslikult (või empiiriliselt) võimalikud? Loogiliselt on võimalik, et taldrik lendab üles, kui lasta ta käest lahti vaakumis või planeedi pinnal, kuid looduslikult on see ikkagi võimatu, sest loodusseadused ei luba seda. Samamoodi ei tõesta kvaalide puudumise ja kvaalide inversiooni loogiline võimalikkus kaugeltki veel seda, et nad oleksid ka looduslikult võimalikud. Invariantsusprintsiip ütleb, et funktsionaalne organisatsioon determineerib teadvuselamuse mingi seaduspärase ühenduslüli kaudu aktuaalses maailmas, ja loogilise võimalikkuse küsimus ei puutu siin asjasse. Käesolevas peatükis tähendab sõna “võimalikkus” üksi, ilma täpsustuseta, alati looduslikku võimalikkust.

Järgnevas asun vaatlema kvaalide puudumise ja kvaalide inversiooni loodusliku võimalikkuse kasuks esitatud argumente, seejärel aga esitan detailsed argumendid nende võimalikkuste vastu. Neis viimastes argumentides on määrav koht mõttelistel eksperimentidel. Kvaalide puudumise võimalikkuse kummutamiseks vaatlen mõttelist eksperimenti *hääbuvate kvaalidega*​, kvaalide inversiooni võimalikkuse kummutamiseks mõttelist eksperimenti *hüplevate kvaalidega*​.

Argumendid, mida me neist mõttelistest eksperimentidest tuletada saame, on vaid tõepärased, kuid minu arvates kaunis tugevad. Kinnitamaks nende mõtteliste eksperimentide kiuste, et kvaalide puudumine ja kvaalide inversioon on looduslikult võimalikud, tuleks aktsepteerida mõningaid väheusutavaid teese teadvuselamuse olemuse ning eelkõige teadvuse ja kognitsiooni vahelise suhte kohta. Mõnedel loomulikel eeldustel selle suhte kohta ​**[lk.251]**​ osutub invariantsusprintsiip palju tõepärasemaks kui kõik teised hüpoteesid.

Võib-olla on kasulik vaadelda nende mõtteliste eksperimentide rolli analoogilisena Schrödingeri kassi mõttelise eksperimendi rolliga kvantmehhaanika interpreteerimisel. Schrödingeri kassi katse ei langeta lõplikku otsust ühe või teise interpretatsiooni kasuks, kuid toob esile erinevate interpretatsioonide tõepärasemaid ja vähem tõepäraseid aspekte, millega iga interpretatsioon peab nii või teisiti arvestama. Samamoodi peab iga teadvuseteooria nii või teisiti arvestama ka hääbuvate ja hüplevate kvaalide eksperimentidega, ja mõned neist tulevad sellega paremini toime kui teised. Nii selguvad mitmesuguste olemasolevate teooriate voorused ja puudused.

### 2. Kvaalide puudumine

Kvaalide puudumise loodusliku võimalikkuse kasuks ei ole positiivseid argumente esitatud nii sageli kui kvaalide inversiooni võimalikkuse kasuks, kuid neid siiski leidub. Kõige põhjalikumalt on neid argumente käsitlenud Block (1978).

Peaaegu kõigil neil argumentidel on ühesugune kuju: kõigepealt pakub esitaja välja meie funktsionaalse organisatsiooni realiseeringu mingil tavatul kandjal ning apelleerib siis intuitsioonile. Näiteks osutatakse, et meie aju organisatsiooni võiks simuleerida Hiina elanikkond, või koguni, et seda võiks teha Boliivia majandus. Kui me paneksime iga Hiina elaniku simuleerima üht neuronit (selleks peaks rahvaarv olema kümme või sada korda suurem, aga see ei puutu praegu asjasse) ja korraldaksime nende vahel raadioühendused, mis simuleeriksid sünaptilisi ühendusi, siis oleks funktsionaalne organisatsioon paigas. Aga — ütleb argument — sel eriskummalisel süsteemil ei oleks ju ometi teadvust!

Argumendil ei puudu intuitiivne veenvus. Paljudele inimestele ütleb tunne, et see lihtsalt ei oleks niisugune süsteem, millel võiks olla teadvuselamusi. Säärane “kollektiivne teadvus” oleks justkui midagi, mis kuuluks pigem ulmejuttu kui reaalsusse. Kuid peale intuitiivse veenvuse ei ole siin midagi. Otsustav see argument kindlasti ei ole. Paljud autorid on osutanud[[295]](#footnote-296), et kui elamuse teke säärases süsteemis ongi intuitiivselt ebausutav, siis samavõrd on intuitiivselt ebausutav, et elamus võiks sündida ajus! Sest kes võiks arvata, et see kämp halli ollust võiks produtseerida erksaid subjektiivseid elamusi? Ja ometi teeb ta seda. See muidugi ei tõesta, et ühe riigi elanikkond võiks produtseerida teadvuse, kuid on tugevaks vastukaaluks intuitiivsele argumendile, et see ei ole võimalik.

Muidugi ei näeks me sellises süsteemis mitte mingisugust teadvuselamust. Aga selles pole midagi üllatavat — ei näe me teadvuselamust ju üldse mitte kellelgi. Võiks tunduda, et säärases ​**[lk.252]**​ süsteemis ei ole teadvuselamusele “kohta”, kuid teisiti ei tundu see olevat ka ajuga. Kolmandaks, me võime seletada kõnealuse süsteemi toimimist ilma teadvuselamusele viitamata, kuid seegi on ka tavalistel juhtudel nii. Kui loogilise superventsuse puudumise tegelik mõte meile kohale jõuab, ei pane mõte, et mõne riigi elanikkonna pinnalt võiks sündida teadvuselamus, meid rohkem imestama kui mõte, et seda võiks teha aju[[296]](#footnote-297) [[297]](#footnote-298) [[298]](#footnote-299).

Invariantsusprintsiibile on vastuseisu avaldatud ka põhjendusega, et funktsionaalne organisatsioon võib põhimõtteliselt tekkida ka juhuslikult, näiteks Boliivia majanduses või isegi ämbris (Hinckfuss, tsitaat raamatust Lycan 1987, lk 32). Kuid see nõuaks ennekuulmatut kokkusattumust.[[299]](#footnote-300) Süsteemil peaks olema rohkem kui miljard osa, millest igal on kindel arv seisundeid (ütleme, kümme), ja nende seisundite vahel peaks olema tohutu suur ja keerukas täpselt õigete põhjuslike seoste süsteem, nii et üks kindel seisundimuster annaks tulemuseks just selle seisundimustri, teine just tolle jne. Vajaliku funktsionaalse organisatsiooni realiseerumiseks ei piisa sellest, et need konditsionaalid oleksid lihtsad reeglipärad (s.t. sellele seisundimustrile juhtus seekord järgnema too seisundimuster); nad peaksid olema usaldusväärsed seosed, mis kehtiksid ka kontrafaktiliselt, s.t. niisugused, et sellele kindlale seisundimustrile järgneb mingi kindel teine seisundimuster alati.[[300]](#footnote-301).

Pole raske näha, et kui süsteem koosneb miljardist (10⁹) osast, siis peab ta sobiva funktsionaalse realiseerimiseks rahuldama umbes 1010⁹ säärast konditsionaali. Tõenäosus, et suvaline süsteem suudab niisuguse osade ja seisundite jaotuse korral realiseerida kõik need konditsionaalid, on suurusjärgus ${1}/{\left(10^{10⁹}\right)^{}}$[[301]](#footnote-302). (Tegelikult teeb nõue, et kõik konditsionaalid oleksid usaldusväärsed, tõenäosuse veel sellestki arvust palju väiksemaks.) Hoolimata meie vabadusest süsteemi osadeksjagamisel on ikkagi ülimalt ebatõenäoline, et selline organisatsioon realiseeruks suvalises süsteemis või üldse üheski süsteemis, mida pole vorminud loodusliku valiku äärmiselt meelevaldsusetud mehhanismid.

Kui me mõistame, kui rangeid kitsendusi seab süsteemi struktuurile funktsionaalse organisatsiooni kirjeldus, siis ei pea me enam nii uskumatuks mõtet, et isegi Hiina elanikkonnast võib saada teadvuselamuse kandja, kui teda õigesti organiseerida. Kiirendades selle rahvahulga käitumist mõtteliselt umbes miljon korda ja vähendades teda inimese pea suuruseks, saame midagi, mis sarnaneb tublisti ajuga, ainult et neuronite asemel on homunkulused — tibatillukesed inimesed —, ja meile ei jää enam kuigi palju põhjust arvata, et neuronid saavad elamuse kandmisega paremini hakkama kui homunkulused.

Muidugi — tähendab Block —, me teame, et neuronid saavad sellega hakkama, aga meil ei ole mingit teadmist homunkuluste kohta. Küsimus jääb seega lahtiseks. Oluline on see, et säärased mõttekäigud on väga nõrgad argumendid kvaalide puudumise loodusliku võimalikkuse vastu. Selleks, et küsimus üht- või teistpidi otsustada, on tarvis midagi veenvamat. Pole võimatu, ​**[lk.253]**​ et õigus on Blockil, kui ta ütleb, et meie intuitsioon asetab tõendamiskohustuse sellele, kes usub, et kvaalid on organisatsiooniliselt invariantsed. Minule pole see sugugi ilmne, kuid järgnevas püüan ma seda kohustust täita.

Mõnikord tuuakse kvaalide puudumise vastu välja argument hoopis teisest vallast: see argument toetub pimenägemise fenomenile ja kinnitab, et pimenägemisega patsiendid on meiega igas olulises mõttes funktsionaalselt sarnased — nad suudavad objekte eristada, nende kohta ütlusi anda jne. —, kuid neil puudub nägemiselamus. Seega ei sõltu elamuse olemasolu või puudumine visuaalsete protsesside funktsionaalsest organisatsioonist.

Kuid nagu me [6. peatükis](#PARA.125-6) nägime, toimub infotöötlus pimenägijatel oluliselt teisiti kui normaalsetel katseisikutel. Nendel puudub meile tuttav otsene ligipääs visuaalsele informatsioonile. Kui neil informatsioonile üldse ligipääsu on, siis on see kaudne, ja igatahes ei ole informatsioon kättesaadav käitumise juhtimiseks tavapärasel viisil. Tõepoolest, just nende käitumises avalduvate erinevuste põhjal infotöötluse organisatsioonis me ju märkamegi nende juures midagi ebatavalist ning peame seepõhjal postuleerima elamuse puudumise. Seega ei sobi pimenägemise juhtumid tõenduseks invariantsusprintsiibi vastu.

### 3. Kvaalide hääbumine

Minu positiivne argument kvaalide puudumise võimalikkuse vastu toetub mõttelisele eksperimendile, kus aju osi vahetatakse järk-järgult välja, näiteks ränikiipide vastu. Taoliste mõtteliste eksperimentidega on sageli vastatud kvaalide puudumise argumentidele tehisintellekti uurijate folklooris ja mõnikord ka trükisõnas. Järkjärgulise asendamise stsenaariumi on pooldanud Pylyshyn (1980), kuigi tema ei sidunud sellega ühtki süstemaatilist argumenti. Argumente, milledel ei puudu sarnasus nendega, mida mina kohe esitama hakkan, on esitanud Savitt (1982) ja Cuda (1985), kuigi nendel on mõttearenduse kulg teistsugune ja ka asja moraal ei ole neil päris sama[[302]](#footnote-303).

See hääbuvate kvaalide argument ei ole kõige tugevam ja peamine argument kvaalide puudumise võimalikkuse vastu; see roll kuulub on “hüplevate kvaalide” argumendile, mis tuleb jutuks [5. jaos](#ORG.7.5) ja mis on argument ka kvaalide inversiooni võimalikkuse vastu. Kuid hääbuvate kvaalide argument on ka iseenesest tugev ning on ühtlasti heaks toeks ja taustaks järgmisele[[303]](#footnote-304), tugevamale argumendile.

See argument on oma vormilt *reductio ad absurdum*​. Oletame, et kvaalide puudumine on looduslikult võimalik. Sellisel juhul võib eksisteerida süsteem, mille funktsionaalne organisatsioon ühtib mõne teadvusega süsteemi (näiteks minu) omaga, kuid millel teadvuselamus täielikult puudub. Üldisust minetamata​[[304]](#footnote-305) võime oletada, et see on nii sellepärast, et süsteem ei koosne neuronitest, vaid ränikiipidest. (Allpool näitan, ​**[lk.254]**​ kuidas laiendada seda argumenti muudliiki isomorfidele.) Olgu selle funktsionaalse isomorfi nimeks Robot. Roboti kognitiivse süsteemi põhjuslikkusmustrid ühtivad minu omadega, kuid tal on sombi teadvus.

Neil eeldustel võime konstrueerida rea minu ja Roboti vahelisi vaheastmeid, kus iga aste erineb eelmisest väga vähe ja funktsionaalne organisatsioon jääb kogu aeg samaks. Võime näiteks ette kujutada, et mingi arv minu neuroneid asendatakse ränikiipidega. Esialgu asendatakse üksainus neuron. Asendajaks on ränikiip, mille lokaalne funktsioon ühtib neuroni omaga. Seal, kus neuronil oli ühendus teiste neuronitega, on nüüd samade neuronitega ühenduses kiip. Ränikiibi seisund on tundlik samadele elektrilistele ja keemilistele signaalidele nagu enne neuroni seisund. Me võime ette kujutada, et kiibil on tillukesed andurid, mis detekteerivad elektrisignaale ja ioone, edastades kiibile vastava digitaalse signaali. Olukordades, kus neuron produtseerib elektrilisi ja keemilisi väljundsignaale, teeb sama ka kiip. (Võime kujutleda, et tal on tillukesed täiturid, mis produtseerivad elektrisignaale ja ioone vastavalt kiibi seesmisele seisundile.) Oluline on, et kiibi seesmised seisundid on niisugused, et tema sisend- ja väljundfunktsioonid ühtivad täpselt neuroni omadega. Pole tähtis, kuidas kiip seda saavutab — näiteks otsingutabeli abil, mis assotsieerib iga sisendsignaaliga talle vastuseks sobiva väljundtoime, või neuronisiseste protsesside arvutusliku simuleerimise teel —, peaasi, et tema sisendite ja väljundite vahel oleks õige seos. Kui lokaalsed vastasmõjud on õiged, siis ei mõjuta asendus süsteemi üldist funktsioneerimist.

Teisel etapil asendame ränikiipidega kaks neuronit. Kõige hõlpsam on võtta nendeks kaks naaberneuronit. Nii pole meil tarvis ebamugavaid andureid ja täitureid, mis nende kahe kiibi vahel ühendust peaksid. Me võime asendada need ükskõik millise ühendusega, mis on tundlik esimese kiibi seesmisele seisundile ja mõjutab õigel viisil teise kiibi seesmist seisundit. (Ühendused võivad muidugi olla mõlemasuunalised.) Me hoolitseme selle eest, et see ühendus — milleks võib olla näiteks mingi elektrooniline signaal — ei erineks vastavast ühendusest Robotil.

Edasised vaheastmed konstrueerime analoogiliselt. Igal järgneval astmel asendame ränikiipidega järjest suurema naaberneuronite rühma. Selle kiibirühma sees puudub biokeemiline substraat täielikult. Biokeemilisi mehhanisme on ainult ülejäänud süsteemis ning rühma piirkiipide ja nende naaberneuronite vahelistes ühendustes. Viimases etapis on kogu süsteemi kõik neuronid asendatud kiipidega, nii et biokeemilised mehhanismid ei mängi enam olulist rolli. (Ma ei hakka siinkohal arutama spetsiifilisemaid küsimusi, näiteks ​**[lk.255]**​ kas mingi mittetriviaalne roll on ka gliiarakkudel; kui jah, siis on nad süsteemi funktsionaalse organisatsiooni komponendid ja tuleb samuti asendada.)

Me võime ette kujutada, et see seesmine süsteem on kogu aeg ühenduses kehaga, on tundlik kehasse saabuvatele stiimulitele ning produtseerib andurite ja täiturite abil õigeid motoorseid reaktsioone. Iga süsteem selles reas on minuga funktsionaalselt isomorfne nii peenel tasemel, et ka tema käitumuslikud soodumused ühtivad minu omadega. Ent kui spektri ühes otsas olen mina, siis süsteem, mis jääb teise otsa, on sisuliselt Roboti koopia.

Et pilti konkreetsemaks teha, kujutlegem, et mina olen esimene süsteem ja mul on parajasti väga külluslik teadvuselamus. Näiteks: ma olen spordihallis, kus käib korvpallimatš, näen enda ümber erksavärvilistes riietes pealtvaatajaid ja kuulen nende ergutushüüdeid, haistan nende kaasavõetud suupistete isuäratavat lõhna, tunnen tuikavat peavalu jne. Koondagem tähelepanu erepunastele ja -kollastele värvuselamustele mängijate särkide​[[305]](#footnote-306) nägemisest. Viimane süsteem meie konstrueeritud reas — Robot — on samas olukorras, töötleb samu sisendeid ja produtseerib sarnast käitumist, kuid hüpoteesi kohaselt ei ole tal seejuures mitte mingisuguseid elamusi.

Minu ja Roboti vahel on palju vaheastmeid. Küsime: Mis tunne on olla mõni nendest? Missugused on nende elamused, kui neil neid on? Kuidas teadvuselamus mööda spektrit liikudes varieerub? Eeldatavasti on kõige minupoolsematel astmetel elamused üpris sarnased minu omadega, teises otsas aga on elamusi vähe või üldse mitte; aga vahepealsetel astmetel?

Arvestades, et süsteemil, mis jääb spektri teise, Roboti-poolsesse otsa, ei ole teadvust, näib, et kusagil vahepeal peab toimuma üks kahest: kas (1) teadvus hääbub mööda rida edasi liikudes järk-järgult, kuni lõpuks kaob, või (2) kusagil vahepeal kaob teadvus ühekorraga täielikult, kuigi eelmisel astmel olid külluslikud teadvuselamused veel täiesti olemas. Nimetagem esimest võimalust *kvaalide hääbumiseks* ja teist *kvaalide hüppeliseks kadumiseks*​.

Kvaalide hüppelise kadumise varianti ei ole raske välistada. Selle hüpoteesi kohaselt võib ühe neuroni väljavahetamise tõttu (kõike muud samaks jättes) kaduda kogu teadvuselamuste väli. See tundub äärmiselt ebausutava, et mitte öelda pentsiku mõttena. Kui see oleks võimalik, tähendaks see niivõrd järske katkestusi loodusseadustes, nagu me kusagil mujal ei näe. Iga konkreetne punkt, kus kvaalid korraga kaoksid (50 protsenti neuroneid? 25 protsenti?), oleks täiesti meelevaldne. Me võime teha oma katse mõttes veelgi peeneteralisemaks ja minna neuroni sisse, nii et viimaks selgub, et mõne üksiku molekuli asendamine põhjustaks terve elamuste valla kadumise. (Vastasel korral jõuaksime jälle hääbuvate kvaalide stsenaariumi juurde.) Nagu neis asjus ikka, nii ka seda hüpoteesi ei ole võimalik ümber lükata, kuid ta on sellegipoolest väga ebausutav[[306]](#footnote-307).

(Siin annaks vastu vaielda, et ka mittelineaarses dünaamikas on olukordi, kus ühe suuruse sõltuvus teisest on väga tundlik, s.t. väikesed muutused ühes põhjustavad suuri muutusi teises. Kuid kõigil neist juhtudest on sõltuvus ​**[lk.256]**​ sellegipoolest pidev[[307]](#footnote-308): on olemas vahepealsed juhtumid, kus sõltuv suurus omandab ka vahepealseid väärtusi, ja seega annab analoogia mittelineaarsete süsteemidega meile tulemuseks kvaalide hääbumise. Ja igal juhul tuleneb tundlik sõltuvus neil juhtudel enamasti mitme fundamentaalsema pideva sõltuvuse koosmõjust. Kõigis seni teada olevates fundamentaalsetes seadustes on ühe suuruse sõltuvus teisest pidevast suurusest selles mõttes pidev, ja pidevuste liitmisest ei saa kuidagi tekkida mittepidevust. Kvaalide hüppeline kadumine eeldaks seega, erinevalt mittelineaarsest dünaamikast, järske pidevuse katkemisi fundamentaalseadustes).

Kvaalide hüppeline kadumine välistatud, jääb veel vaadelda kvaalide hääbumise hüpoteesi. Et seda stsenaariumi paremini mõista, vaadelgem süsteemi, mis jääb täpselt minu ja Roboti vahelise spektri keskele, kus teadvus on oluliselt tuhmunud, kuid mitte veel täielikult kadunud. Paneme sellele süsteemile nimeks Joe. Mis tunne on olla Joe? Mõistagi on Joe minuga funktsionaalselt isomorfne. Ta ütleb oma elamuste kohta kõike sedasama, mida mina ütlen enda elamuste kohta. Korvpallimatši vaadates ei suuda ta mängijate erepunaseid ja -kollaseid särke nähes hüüatusi tagasi hoida.

Hüpoteesi kohaselt ei ole Joel aga üldse eredaid punaseid ja kollaseid elamusi, vaid võib-olla näiteks tuhmroosad ja määrdunudpruunid, võib-olla äärmuseni ähmastunud või peaaegu mustaks tumenenud punased ja kollased — on palju mõeldavaid viise, kuidas punased elamused võiksid järk-järgult üle minna elamuste täielikuks puudumiseks, ja arvatavasti veelgi rohkem viise, mis ei tule meile pähegi. Kuid eeldatavasti peavad elamused enne kadumist kaotama oma ereduse — muidu jõuaksime tagasi kvaalide hüppelise kadumise hüpoteesi juurde. Samamoodi peaks spektris leiduma punkt, kus vahepealse süsteemi elamuses pole enam peeni erisusi, mis on olemas minu elamuses; kui oletada, et kõik minu elamuse erisused on olemas kuni hetkeni, mil nad ühekorraga kaovad, jõuame taas veel ühe versiooni juurde kvaalide hüppelise kadumise variandist.

Kujutagem konkreetsuse mõttes ette, et seal, kus mina näen erepunast, näeb Joe tuhmroosat, kusjuures paljusid toonidevahelisi erinevusi, mis leiduvad minu elamuses, tema elamuses enam ei ole. Seal, kus minu elamuses on vali müra, on Joe elamuses võib-olla ainult kauge kõmin. Mõnes asjas on Joel vedanudki: kui minul on põletav peavalu, tunneb tema ainult kerget kipitust.

Oluline punkt on siin see, et Joe arusaamine kõigist oma elamustest on süstemaatiliselt vigane. Muidugi ütleb ta, et tema elamuses on korvpallurite särgid erepunased ja -kollased, kuigi tema elamuses on nad vaid tuhmroosad ja määrdunudpruunid.[[308]](#footnote-309) [[309]](#footnote-310) Kui teda küsitleda, ütleb ta, et tema elamuses on kõikvõimalikke väga peeni punaseid varjundeid, kuid tegelikult on paljud neist varjunditest tema jaoks täiesti eristamatud. Ta kaebab valju müra üle, kuigi tema kuulmiselamus on tegelikult väga nõrk. Mis veelgi hullem, kui tõlgendada sõna “uskumus” funktsionaalselt, siis võib öelda, et Joe isegi usub, ​**[lk.257]**​ et tal on kõik need keerukad elamused, mida tal tegelikult ei ole. Ühesõnaga, Joe teadvuselamus on Joest endast täiesti irdus ja ta on võimetu sellega ühendusse astuma.

See kõlab küll üsna ebausutavalt: me kõneleme olendist, kellel ratsionaalsed protsessid toimivad ja teadvus on faktiliselt olemas, kuid kes omaenda teadvuselamustes täielikult eksib. Vahest ehk viimases äärmuses, kui kõik on seestpidi pime, ongi mõistlik oletada, et süsteem oma ütlustes ja otsustustes nii põhjalikut eksib — mingis mõttes ei olegi süsteemis ju kedagi eksimas. Kuid vahepealsetel juhtudel on see palju vähem tõepärane. Kõigil tuttavatel juhtudel on teadvusega olendid, kui miski nende tähelepanu kõrvale ei kalluta ja nad on täie mõistuse juures, üldiselt suutelised moodustama oma elamuse kohta paikapidavaid otsustusi. Kui tunnete ja mõistusega olend, kes ei kannata ühegi funktsionaalse häire all, on nii süstemaatiliselt irdus iseenda elamustest, siis järelikult on seos tema teadvuse ja kognitsiooni vahel väga nõrk. Meil pole kaalukat põhjust uskuda, et teadvus on nii anomaalne fenomen, ja on küllalt põhjust uskuda vastupidist.

Muidugi, puht loogiliselt on kvaalide hääbumine võimalik. Iseenda elamustes nii sügavalt eksiva süsteemi kirjelduses ei ole mingit vastuolu.[[310]](#footnote-311) [[311]](#footnote-312) Kuid loogiline võimalikkus ja looduslik võimalikkus on kaks ise asja. Meil ei ole põhjust uskuda, et midagi niisugust võiks tegelikkuses juhtuda, kuid on kõigiti põhjust uskuda vastupidist. Üks silmatorkavamaid empiirilisi fakte teadvuse kohta paistab olevat see, et kui teadvusega ja piisavalt keerukaid mõisteid valdaval olendil on mingi elamus, siis on ta võimeline moodustama oma elamuse kohta ka otsustusi. Võib ju ette tulla ebaharilikke juhtumeid, kus otsustuste moodustamise mehhanismides on ratsionaalsete protsesside tugeva kahjustuse tõttu talitlushäired, kuid juhtum, millest meie kõneleme, ei ole niisugune. Joel toimivad ratsionaalsed protsessid sama hästi kui minul — me eeldasime ju, et ta on minuga funktsionaalselt isomorfne. Ta lihtsalt mõistab oma elamusi täiesti vääriti.

Igapäevases elus tuleb muidugi ette väga mitmesuguseid kvaalide hääbumise juhtumeid. Mõelgem näiteks sellele, mis toimub meiega uinudes, või liikugem evolutsiooniahelat mööda mõtteliselt tagasi trilobiitideni. Kõigil neil juhtudel hääbub teadvuselamus spektrit mööda liikudes järk-järgult. Kuid iga kord kaasnevad elamuse hääbumisega ka sellekohased funktsionaalsed muutused. Uinumise eel ma ei usu, nagu oleksin ma täiesti ärkvel ja nagu mul oleks intensiivseid elamusi (senikaua kui ma pole hakanud juba und nägema — siis on mul neid väga tõenäoliselt küll). Koera värvuselamuste nappusega käib käsikäes eristusvõime nappus tema visuaalsetes mehhanismides. Need juhtumid on täiesti erinevad juhtumist, mis meid siin huvitab — juhtumist, kus elamus hääbub, funktsioneerimine jääb aga samaks. Joe visuaalsed mehhanismid teevad valguse lainepikkustel sama peeni vahesid ja tema otsustuste järgi kajastuvad ​**[lk.258]**​ need erinevused ka tema elamuses, meie aga peaksime uskuma, et see pole üldse nii.

Sarnase mõttelise eksperimendi üle arutledes pakub Searle (1992) välja järgmise võimaluse:

Sedamööda, kuidas sinu hääbuvasse ajju implanteeritakse järjest uusi ja uusi ränikiipe, märkad sa, et sinu teadvuselamus tõmbub aina enam koomale, ilma et see avaldaks mingit mõju sinu välisele käitumisele. Sa avastad oma ülimaks hämmastuseks, et sa hakkad tõepoolest kaotama meelevalda oma välise käitumise üle. Näiteks paned sa tähele, et kui arstid testivad sinu nägemist, kuuled sa neid ütlevat: “Me näitame sulle nüüd punast eset; palun ütle meile, mida sa enda ees näed.” Sa tahaksid hüüda: “Ma ei näe midagi. Ma olen täiesti pimedaks jäänud,” kuid kuuled oma häält täiesti sinu tahtest sõltumatult ütlemas: “Ma näen enda ees punast eset.” Seda mõttelist katset lõpuni läbi tehes saame palju masendavama tulemuse kui eelmisel korral. Me kujutame ette, kuidas sinu teadvuselamus aegamööda olematuks kahaneb, ilma et sinu väliselt jälgitavas käitumises midagi muutuks. (lk. 66–67)

Siin peab Searle kvaalide hääbumist võimalikuks, kuid lisab, et süsteem ei pea tingimata eksima oma uskumustes oma elamuse kohta. Süsteemi uskumused oma elamuse kohta võivad olla tõesed, ainult et nad ei suuda mõjutada tema käitumist.

Selle võimaluse võib vist siiski välistada. Süsteemis ei ole lihtsalt ruumi uute uskumuste tekkimiseks. Kui just mitte tunnistada mingit väga tugevat dualismi, siis tuleb arvata, et uskumuste erinevus kajastub süsteemi funktsioneerimises — kui ka mitte käitumises, siis vähemalt mõneski protsessis. Ent vaadeldav süsteem on ju peensusteni identne esialgse süsteemiga, see tähendab minuga. Temas pole lihtsalt ruumi uutele uskumustele nagu “ma ei näe midagi”, uutele soovidele nagu soov hüüda, ja muudele uutele kognitiivsetele seisunditele nagu hämming. Füüsilises süsteemis ei saa olla midagi sellele hämmingule vastavat. Selle jaoks ei ole ruumi neuronites, mis on ju identsed tavalisi uskumusi kandvate neuronitega; ja Searle ei arva ju ometi, nagu võiks neuronite lihtne asendamine ränikiipidega sünnitada uusi uskumusi! Kui neuronite ja räni vahel ei ole mingit lausa hämmastavat, maagilist vastasmõjuefekti — mis pealegi infotöötluses mitte kuidagi ei avaldu, kuna organisatsioon jääb kogu aeg ühesuguseks —, siis selliseid uusi uskumusi tekkida ei saa.

Organisatsiooni säilitav üleminek neuronitelt ränile lihtsalt ei muuda süsteemi piisavalt selleks, et põhjustada nii märkimisväärset muutust tema kognitiivsete seisundite sisus ja struktuuris. Üks asi on punase elamuse asendumine sinisega, kuid hoopis teist suurusjärku oleks uskumuse “Suurepärane mäng!” asendumine uskumusega “Appi! Ma olen sattunud halba õudusfilmi!”. Kui kognitiivse sisu nii põhjalik muutus ei kajastuks mitte kuidagi funktsionaalses organisatsioonis, siis peaks kognitsioon hõljuma tühjuses süsteemi seesmisest funktsioneerimisest vabana otsekui kehatu kartesiaanlik vaim; kui kognitiivsete seisundite sisu oleks üldse füüsikaliste seisundite suhtes superventne, siis vaid kõige ​**[lk.259]**​ meelevaldsemate ja kapriissemate reeglite järgi. (Nagu: kui organisatsioon on realiseeritud neuronites, siis “Kenad värvid!”, kui ränis, siis “Oh õudust!”)

Niisiis eeldab kvaalide hääbumise võimalikkus kas mingit väga veidrat seost uskumuste sisu ja füüsiliste seisundite vahel või siis täiesti ratsionaalsete, kuid omaenda teadvuselamustes rängalt eksivate olendite võimalikkust. Mõlemad hüpoteesid on tunduvalt vähem usutavad kui hüpotees, et ratsionaalsete teadvusega olendite otsustused oma elamuste kohta on üldiselt õiged. Palju mõistlikum hüpotees on seega, et neuronite asendamisel kvaalid ei hääbu sugugi. Praktikas on Joega sarnaneval süsteemil sama külluslikud teadvuselamused nagu minul. Kui see on nii, siis oli meie esialgne oletus ekslik ja meie esialgsel isomorfil, Robotil, on teadvuselamused olemas.

Seda argumenti saab hõlpsasti laiendada teistele funktsionaalsetele isomorfidele. Selleks, et käsitleda juhtumit, kus minu organisatsiooni implementeerib mingi maa elanikkond, võime konstrueerida sarnase spektri juhtumitest, mille ühes otsas on minu ränist isomorf ja teises otsas kogu see rahvahulk. Selleks laotame ränist süsteemi järk-järgult laiali, kuni tema pindala ulatub paljude ruutmiilideni. Ühtlasti teeme teda niipalju aeglasemaks, et kiirus, millega sisendsignaalid kiipidesse saabuvad, oleks võimetekohane ka inimesele. Seejärel hakkame kiipe ükshaaval inimestega asendama, jälgides, et nad väljastaksid igale sisendsignaalile vastuseks õiged väljundsignaalid. Nii saamegi lõpptulemuseks olukorra, kus kogu elanikkonna organisatsioon vastab minu neuronite omale ning võib-olla isegi juhib raadio teel midagi kehataolist. Igal etapil on süsteem minuga funktsionaalselt isomorfne ja talle saab rakendada kõiki seniseid argumente. Teadvuselamus jääb kas alles, hääbub sujuvalt või kaob hüppeliselt. Kaks viimast võimalust on sama väheusutavad nagu ennegi. Seega võime järeldada, et süsteem, mida realiseerib maa elanikkond, võib olla teadvuselamuste kandja täpselt samuti, nagu seda on aju.

Sedasama võib teha suvalise funktsionaalselt isomorfse süsteemiga, ka kujult, suuruselt, kiiruselt, füüsiliselt koostiselt jne. väga erinevatega. Kõigil juhtudel on järeldus sama. Kui vaadeldaval süsteemil ei ole teadvust, siis peab eksisteerima vahepealne isomorfne teadvusega süsteem, millel on tuhmunud elamused ja mis mõistab omaenda elamusi täiesti vääriti. Kui me ei ole valmis aktsepteerima seda tohutut lõhet teadvuse ja kognitsiooni vahel, siis pidi esialgsel süsteemil teadvus ikkagi olema.

Kui kvaalide puudumine on võimalik, siis on võimalik ka kvaalide hääbumine. Kuid kvaalide hääbumine on, nagu ma olen püüdnud näidata, peaaegu kindlasti võimatu. Sellest järeldub, et ka kvaalide puudumine on peaaegu kindlasti võimatu.

Asun nüüd vaatlema mitmesuguseid sellele argumendile esitatud vastuväiteid.

#### Esimene vastuväide: Neuronite asendamine ei ole praktikas võimalik

Praktilise mõttelaadiga lugejale ei tarvitse kogu see mõtteliste eksperimentide metodoloogia muljet avaldada. Ta võib vaielda, et neuronite asendamine ränikiipidega ​**[lk.260]**​ kuulub ulmekirjanduse valdkonda, mitte tegelikkusesse; igatahes oleks säärane asendamine praktikas võimatu, nii et kõik järeldused, mida meie mõttelisest eksperimendist teha saab, ei kajasta tegelikku olukorda.

Kui oponent peab silmas lihtsalt tehnilist võimatust, siis pole häda veel nii suur. Küsimus on selles, missugused oleksid selliste süsteemide elamused, kui nad oleksid olemas, ükskõik kas me suudame neid tegelikult konstrueerida või mitte. Iseasi on loodusliku võimatusega. Võib-olla ränikiipide võimalustest lihtsalt ei jätku selleks, et täita funktsioone, mida täidab ajus neuron. Ei ole selge, kas sellel vastuväitel on põhimõttelist alust; käe- ja jalaproteesid meil juba on, silmaproteesid on tulekul, miks siis ei peaks võimalikud olema ka neuroniproteesid? Ent isegi kui ränipõhine funktsionaalne isomorf olekski võimatu (neuronfunktsiooni mittearvutuvuse tõttu?), ei riivaks see invariantsusprintsiipi toetavat argumenti sugugi. Invariantsusprintsiip ütleb ainult, et kui teadvusega süsteemi funktsionaalne isomorf on olemas, siis on tal samasugused teadvuselamused. Kui ränipõhised isomorfid on võimatud, siis ei puutu meie arutelud ränisüsteemide kohta lihtsalt asjasse.

Oponent võib viia jutu raskustele, mida valmistab räni–neuron-liides: ta ütleb, et puhas neuronsüsteem ja puhas ränisüsteem võivad ju mõlemad täiesti võimalikud olla, kuid vahepealsete süsteemidega pole asi nii kindel. Võib-olla ei ole kiibi tibatillukeses ruumalas piisavalt ruumi anduritele ja täituritele? Peavad ju täituritel olema keemiliste ainete reservuaarid, et neid vajadusel vallandada. Aga reservuaar ei tarvitse üldsegi suur olla: meie mõttekäik nõuab isomorfismi ainult mõneks sekundiks! Pealegi saab meie mõttelises eksperimendis alati muudatusi teha ja süsteemi mõtteliselt laiendada. Igatahes on raske näha, kuidas saaks niisugusele argumendile ehitada kuigivõrd soliidset ja põhimõttelist vastuväidet invariantsusprintsiibile. Süsteeme, millede vahel saaks järkjärgulise asendamise teel üleminekuid luua, on arvatavasti võimalik leida; kas vastuväitja tahab öelda, et invariantsusprintsiip kehtib ainult nende, aga mitte teiste süsteemide kohta? Kui jah, siis on tema arusaamine asjast üpris meelevaldne; kui ei, siis peab ta otsima vastuväiteid, mis ulatuksid sügavamale.

#### Teine vastuväide: Süsteem võib mõista oma elamusi täiesti valesti

Teine vastuväide toetub asjaolule, et mõnes olukorras eksivad subjektid oma elamustes täielikult. Näiteks nägemiskaosalgamise​ korral usuvad katsealused, et neil on nägemiselamusi, mida neil vististi tegelikult ei ole. Kuid niisugused subjektid ei ole enam täiesti ratsionaalsed süsteemid. Süsteemis, mille uskumuste kujunemise mehhanismides on ​**[lk.261]**​ häireid, võib kõike juhtuda. Selline süsteem võib uskuda, et ta on Napoleon või et Kuu on roosa. Minu “hääbunud” elamustega isomorf Joe on seevastu täiesti ratsionaalne süsteem, mille kognitiivsed mehhanismid toimivad sama hästi nagu minu omad. Tema jutt jätab täiesti aruka mulje. Me ei leia ühtki silmatorkavalt suurt loogilist vastuolu tema uskumuste vahel ega ühtki süstemaatilist psühhiaatrilist häiret, mis kallutaks tema mõtlemisprotsesse ekslikele järeldustele. Joe on täiesti mõtlev ja arukas inimene, kes ei kaldu sugugi konfabuleerima, nagu seda juhtub nägemiskaosalgajatega​. Nende kahe juhu vahel pole vähimatki analoogiat. Tõepärane pole mitte väide, et ükski süsteem ei saa oma elamustes põhjalikult eksida, vaid et seda ei saa teha ükski ratsionaalne süsteem, mille kognitiivsed mehhanismid töötavad veatult. Joe on kindlasti ratsionaalne süsteem, mille mehhanismid talitlevad sama hästi kui minu omad, ja seepärast tema juhtum meie argumenti ei mõjuta.

#### Kolmas vastuväide: Soriitargumendid on kahtlased

Mõned oponendid ütlevad, et see argument on vormilt *soriit* ehk “libeda tee” argument, ning et niisugused argumendid on tavaliselt kahtlased. Soriidi abil saab “näidata”, et isegi üks liivatera on hunnik: miljon liivatera on ju hunnik, ja kui me võtame hunnikust ühe liivatera ära, jääb hunnik ikka hunnikuks. Ent see, kes nii vastab, pole argumendisse süvenenud. Soriit tuleb välja siis, kui ignoreeritakse asjaolu, et mingi näiline dihhotoomia on tegelikult kontiinum: liivahunniku ja liivatera vahele jääb ka mitmesuguseid raskesti määratletavaid vahepealseid juhtumeid. Minu argument seevastu kõneleb selge sõnaga kontiinumist ning väidab, et vahepealsed juhtumid on võimatud sõltumatutel põhjustel.

Argument oleks soriit, kui tema vorm oleks järgmine: mul on teadvus; kui asendada teadvusega süsteemis üks neuron ränikiibiga, siis on tal ikkagi teadvus; seega on teadvus ka üleni ränipõhisel süsteemil. Kuid tema vorm ei ole niisugune. On tõsi, et kvaalide hüppelise kadumise vastase argumendi aluseks on järsu muutuse võimatus, kuid asi on just selles, et siin on jutt järsust ja suurest muutusest, kus külluslikud teadvuselamused asenduvad ühel hoobil teadvuselamuste täieliku puudumisega. See on ebatõenäoline põhjustel, millel pole midagi ühist soriitidega.[[312]](#footnote-313).

#### Neljas vastuväide: Sarnaste argumentidega saaks tõestada ka käitumuslikku invariantsust

Neljas vastuväide ütleb, et meie argument püüab tõestada liiga palju. Kui tema abil oleks võimalik tõestada organisatsioonilise invariantsuse printsiipi, siis saaks analoogilise argumendiga tõestada ​**[lk.262]**​ ka käitumusliku invariantsuse printsiipi. Selleks tuleks konstrueerida süsteemide kontiinum, mille ühes otsas oleksin mina ja teises suvaline minuga käitumuslikult samaväärne süsteem. Meie argumendiga analoogiline arutluskäik viiks järeldusele, et sellisel süsteemil peab olema teadvus. Kuid võib vist uskuda, et leidub minuga käitumuslikult ekvivalentseid süsteeme, mille ei ole teadvust — näiteks Blocki (1981) hiiglaslik otsingutabel, millesse on salvestatud õige väljundtoime iga sisendsignaalide kombinatsiooni jaoks. Seega peab meie argumendis olema viga.

Vastuväide ei ole korrektne kahel põhjusel. Esiteks toetus minu argument osaliselt asjaolule, et minuga funktsionaalselt isomorfsel süsteemil on samasugune kognitiivne struktuur nagu minul, mis tähendab muu seas ka seda, et ta moodustab samasuguseid otsustusi. See viiski meid järelduseni, et hääbunud kvaalidega süsteemi Joe otsustused peavad olema radikaalsed ekslikud. Käitumuslikult samaväärsete süsteemide kohta seda öelda ei saa. Näitleja, kes mind täiuslikult kehastab, ei pea moodustama samasuguseid otsustusi nagu mina. Sama kehtib ka otsingutabelite ja vahepealsete süsteemide kohta — nende töömehhanismid on hoopis teised.

Teiseks ei ole üldsegi selge, kuidas oleks võimalik üleminek minust minu suvalise käitumusliku isomorfini väikeste sammude haaval, rikkumata käitumuslikku samaväärsusgt. Kuidas saaks seda teha näiteks otsingutabeliga? Suurte sammudega oleks see ehk kuidagi võimalik, kuid meie argumendiks sellest ei piisa: kui järjestikuste süsteemide vahelised erinevused on suured, siis ei ole kvaalide hüppeline kadumine enam nii ebausutav. Funktsionaalsete isomorfide puhul leidus väga väikeste sammude haaval liikumiseks loomulik viis, kuid käitumuslike isomorfide puhuks säärast loomulikku meetodit ei ole. Seetõttu tundub ebatõenäoline, et kõnealusel argumendil oleks vähimatki edulootust.

.

Kõike kokku võttes arvan ma, et ainus mõistlik viis, kuidas organisatsioonilise invariantsuse oponent meie argumendile vastata saaks, on teha süda kõvaks ja leppida kvaalide hääbumise võimalikkusega ning sellest tulenevalt ka võimalusega, et ratsionaalsel teadvusega süsteemil võib olla oma elamusest radikaalselt vale ettekujutus — või siis teha süda veel kõvemaks ja tunnistada kvaalide hüppelise kadumise võimalikkust ja sellest tulenevaid järske katkestusi. See positsioon on väheahvatlev, kuna ta eeldab teadvuse ja kognitsiooni lahusust, ja alternatiiv tundub palju usutavam; kuid erinevalt teistest vastuväidetest ei ole ta nii ilmselgelt väär. Ja [5. jaos](#ORG.7.5) esitatud hüplevate kvaalide argument annab kvaalide puudumise võimalikkusesse mitteuskumiseks veelgi enam põhjust. Nii et invariantsusprintsiibi vastased ei saa niipeagi rahu.

Tähendan möödaminnes, et sarnase argumendiga saaks tõestada ka teadvuselamuste olemasolu süsteemidel, mille funktsionaalne organisatsioon ei ole ühe või teise teadvusega süsteemi omaga identne, vaid ainult sarnane. Invariantsusprintsiip on omaette võttes ühildatav solipsistliku teesiga, et elamus sünnib minu ja ainult minu organisatsioonist. Kuid samamoodi, nagu me ülal kujutlesime minu füüsilise koosseisu järkjärgulist muutumist, võib ette kujutada ka minu organisatsiooni järkjärgulist muutumist, mille korral minu uskumused oma elamuse kohta jääksid ​**[lk.263]**​ kogu aeg suuremalt jaolt samaks, ma oleksin kogu aeg ratsionaalne süsteem jne. Ülaltoodutega sarnastel põhjustel tundub väga tõenäoline, et teadvuselamus jääb sellise ülemineku korral alles.

### 4. Kvaalide inversioon

Hääbuvate kvaalide argument viib järeldusele, et minu funktsionaalsetel isomorfidel on teadvuselamus, kuid ei luba järeldada, et nende teadvuselamus on samasugune nagu minul. See tähendab, et funktsionaalne organisatsioon määrab teadvuselamuse olemasolu või puudumise, kuid ei pruugi määrata elamuse iseloomu. Tõestamaks, et funktsionaalne organisatsioon määrab elamuse iseloomu, peame tõestama invertkvaalidega funktsionaalsete isomorfide võimatuse.

Kvaalide inversiooni idee on tuttav enamikule meist. Pole vist palju neid, kes ei oleks kunagi mõelnud, kas see, mida üks inimesele näeb punasena, ei paista teise inimese silmis ehk hoopis sinisena, ja ümberpöördult. See on üks neist filosoofilistest mõistatustest, mille puhul ei ole algusest peale selge, kas tal üldse mõtetki on, ja mis võib jätta nõutuks isegi pärast pikki mõtisklusi.

Kvaalide inversiooni võimalikkuse üle arutles esimesena vististi John Locke oma teoses “Essee inimarust” [​*Essay Concerning Human Understanding*​]:

​*Kuid sinise värvuse idee ühel inimesel võib erineda samast ideest teisel inimesel.*​ Ja meie lihtsatele ideedele ei saaks ette heita väärust ka siis, kui *üks ja seesama objekt sünnitaks eri inimeste vaimudes*​ meie organite erineva ehituse tõttu ühel ja samal ajal *erinevaid ideid*​; näiteks kui kannikese idee, mille loovad silmad ühe inimese vaimus, oleks samasugune nagu saialille idee, mis sünnib teise inimese vaimus, ja ümberpöördult. Sest kuna ühe inimese vaim ei saa minna teise inimese kehasse, et tajuda, missuguseid ettekujutusi [​*appearances*​] need organid esile kutsuvad, mistõttu seda ei ole iial võimalik teada saada, siis ei saa ei ideed ega nimed kunagi segi minna ja kummaski ei saa olla väärust. Sest kui kõik asjad, millel on kannikese tekstuur, sünnitavad alati ühtviisi idee, mida ta nimetab siniseks, ja asjad, millel on saialille tekstuur, sünnitavad alati ühtviisi idee, mida nimetab kollaseks, siis olgu need ettekujutused tema vaimus missugused tahes, ikka on ta võimeline asjade vahel, millega tal tegemist on, tegema nende ettekujutuste põhjal sama reeglipäraselt vahet ning mõistma ja tähistama erinevusi, mida märgivad nimed “sinine” ja “kollane”, just nii, nagu oleksid ettekujutused või ideed, mida tema vaim neist kahest lillest saab, täpselt samasugused nagu ideed teiste inimeste vaimudes. (2. rmt., 32. ptk., 15. jagu.)

Siin pakub kvaalide inversioon Locke’ile huvi pigem sarnase käitumisega süsteemide kui täpsete funktsionaalsete isomorfide juhul. Pole vist kahtlust, et ta kõneleb siin kontseptuaalsest võimalikkusest. Meid aga huvitab küsimus, kas ta kõneleb ka empiirilisest võimalikkusest.

​**[lk.264]**​ Isegi need, kes peavad end materialistideks, on sageli väljendanud oletust, et funktsionaalsetel isomorfidel võivad olla erinevad teadvuselamused. Näiteks peetakse sageli looduslikult võimalikuks, et minu teistsuguse füüsilise ehitusega funktsionaalsel isomorfil võivad olla sinised elamused seal, kus minul on punased elamused, või midagi sellesarnast. See ongi invertkvaalide hüpotees. Kui see hüpotees vastab tõele, siis võib teadvuselamuse olemasolu sõltuda ainult funktsionaalsest organisatsioonist, kuid elamuste iseloom sõltuks füsioloogilisest ehitusest või mõnest teisest mittefunktsionaalsest tegurist.[[313]](#footnote-314).

Me oleme juba näinud​, et seda seisukohta on raske materialismiga ühitada. Kui on looduslikult võimalik, et funktsionaalsel isomorfil on invertkvaalid, siis on see ka loogiliselt võimalik. Seega on loogiliselt samavõrd võimalik, et invertkvaalid on ka füüsilisel isomorfil, sest neuronitel ei ole üht või teist liiki kvaalidega kontseptuaalselt rohkem ühist kui ränil. Sellest järeldub, et spetsiifiliste elamuste olemus on lisafakt, mis ulatub füüsikaliste faktide ringist väljapoole ja kõrgemale, ning et materialism peab olema väär (kui mitte ajada “tugeva metafüüsilise paratamatuse” liini). Järgnevas ma aga selle punktiga enam ei tegele. Minu siinsed mõttekäigud ei sõltu materialismi või dualismi tõesusest.

Invertkvaalide või, nagu mõnikord öeldakse, invertspektri võimalikkuse vastu on mõnikord (näiteks Schlick 1932) vaieldud verifikatisionistlikel põhjustel: meil ei ole mitte kuidagi võimalik teada saada, et toimumas on midagi muud, ja seega ei saa olla[[314]](#footnote-315) ka mingit tõelist erinevust. Neid argumente ma mõistagi ei tunnista: pelgast faktist, et meil pole võimalik öelda, millised kvaalid on ühe või teise süsteemi elamuses, ei piisa selleks, et järeldada, et see pole faktiküsimus, sest kvaalide olemusel ei ole kontseptuaalset seost käitumisega; seepärast jätan seesugused vastuväited siinkohal kõrvale. Nagu ma [3. peatükis](#PARA.65-4) mainisin, on sellele hüpoteesi vastu mõnikord ette toodud ka asjaolu, et meie värvusruum on asümmeetriline, mistõttu mitte mingisugune inversioon ei võimalda õigeid vastavusi (näiteks Hardin 1987; Harrison 1967, 1973). Mõned minu [juba esitanud](#PARA.65-5) vastused on kohased ka siin, kuigi seekord on küsimuse all looduslik võimalikkus; näiteks võime me endiselt apelleerida asümmeetrilise värvusruumiga olendi võimalikkusele ja küsida, kas tema funktsionaalsel isomorfil saaks olla invertspekter. Igatahes jätan ma selle raskuse praegu kõrvale ja eeldan siinse argumendi piires, et meie värvusruum on sümmeetriline, ning püüan näidata, et kvaalide inversioon on ka sellisel juhul võimatu.

Arutlustes kvaalide inverteerumise üle on palju segast. Kui ma ütlen “sinine elamus”, kas ma mõtlen selle all (1) seda, mida nimetab siniseks elamuseks subjekt, (2) elamust, mida põhjustab sinine objekt, või (3) seda, mida nimetan siniseks elamuseks mina? Mina valin siin kolmanda võimaluse. Kogu siinses arutelus tähendab “sinine elamus” elamust, mida mina nimetan siniseks, elamust, mis tavaliselt tekib minul siniseid asju, näiteks taevas või merd nähes, jne. Sellise sõnatarvituse korral on mõeldav, ​**[lk.265]**​ et teistel (või kunagi tulevikus koguni minul endal) võivad siniseid elamusi esile kutsuda kollased objektid, objektid, mida nad nimetavad punasteks jne.

.

Argumendid kvaalide inversiooni kasuks seisnevad — nagu see oli ka kvaalide puudumisega — sageli lihtsalt väites, et kvaalide inverteerumine on mõeldav. Kuid on ilmne, et säärased väited jätavad lahtiseks küsimuse looduslikust võimalikkusest. Paar argumenti on esitatud ka kvaalide mitmesuguste inversioonide loodusliku võimalikkuse kasuks, kuid ükski neist ei ohusta organisatsioonilise invariantsuse printsiipi.

Esimene neist argumentidest kinnitab kvaalide inverteerumise võimalikkust funktsionaalse organisatsiooni konstantsena püsides (vt. Gert 1965, Lycan 1973 ja Wittgenstein 1968). Kõigepealt märgime, et kvaale saaks inverteerida subjektisiseselt, kasvõi näiteks minu võrkkesta ja minu aju tsentraalsete piirkondade vahelisi ühendusi ümber lülitades, sellal kui ma ise magan. Ärgates hakkan ma korraga rääkima, et taevas oleks justkui punane, rohi kollane jne., ning on igati põhjust arvata, et minu kvaalid on inverteerunud. Järgmisena vaatleme võimalikku olukorda, kus inimese aju on sel viisil ümber korraldatud juba sündides. Niisugusel inimesel võivad olla kvaalid, mis on normiga võrreldes süstemaatilises inversioonis, kuid loomulikult on ta õppinud nimetama taevast siniseks, rohtu roheliseks jne., nii et invertsus ei pruugi tema käitumises mitte kuidagi avalduda.[[315]](#footnote-316) [[316]](#footnote-317)

See ei luba siiski järeldada, nagu oleks kvaalide inverteerumine funktsionaalse organisatsiooni samaks jäädes võimalik. Et seda mõista, piisab, kui panna tähele, et minu ühendusi ümber korraldades on deemon​[[317]](#footnote-318) minu funktsionaalset organisatsiooni oluliselt muutnud. Funktsionaalne organisatsioon on teistsugune ka sündimisest saadik ümber korraldatud ühendustega subjektil, mistõttu niisugune inimene ei ole minu funktsionaalne teisik. Laias laastus võib tal olla minuga ühiseid funktsionaalseid omadusi, kuid kindlasti ei ühti tema funktsionaalne organisatsioon minu omaga peensustes. Seega need juhud organisatsioonilise invariantsuse printsiipi ei kõiguta.[[318]](#footnote-319).

Veel üks sarnane argument, mille on esitanud Block (1990), kõneleb “invert-Maast”, kus taevas on kollane, rohi on punane, jne.[[319]](#footnote-320) Kujutame ette, et mind röövitakse ja viiakse invert-Maale, kuid ühtlasi pannakse mulle silma värvusi inverteerivad kontaktläätsed, nii et kõik tundub mulle normaalsena. Block argumenteerib selle stsenaariumi abil kvaalide representatsionalistliku käsitluse vastu, mis samastab sinise elamuse tajuseisundiga, mille objektiks on sinised asjad. (Invert-Maal seevastu põhjustavad siniseid elamusi kollased asjad.[[320]](#footnote-321)) Samuti vaidleb ta selle argumendi abil funktsionalistliku vaate vastu, mis võrdsustab sinised elamused seisunditega, mida põhjustavad sinised objektid.

Ka see juhtum ei riiva organisatsioonilise invariantsuse printsiipi mitte kuidagi. Kui ma näen läbi inverteerivate läätsede invert-Maa kollast taevast, on minu seesmine funktsionaalne organisatsioon ju täpselt samasugune nagu siis, kui ma näen päris-Maa sinist taevast, ja ka minu elamus on, just nagu printsiip ennustabki, täpselt samasugune. Mõtteline eksperiment lubab parimal juhul irrutada elamused ​**[lk.266]**​ subjekti keskkonna omadustest ja “laiadest” funktsionaalsetest omadustest, mis hõlmavad tema keskkonda​[[321]](#footnote-322); mind aga huvitab siin ainuüksi seesmine funktsionaalne organisatsioon. Meie mõtteline katse ei anna meile juhtu, kus kahel seesmiselt isomorfsel süsteemil oleksid erinevad elamused, ega puuduta seega invariantsusprintsiipi[[322]](#footnote-323).

### 5. Kvaalide hüplemine

Mõni arvab võib-olla, et hääbuvate kvaalide argumendisse vastavaid muudatusi tehes võib saada argumendi kvaalide inversiooni võimalikkuse vastu. See paraku ei tööta. Kujutage ette, kuidas säärane argument võiks käia. Alustuseks vaatleme mind, kui mul on punane elamus, ja invertkvaalidega süsteemi, kellel on sinine elamus. Järkjärgulise asendamise teel konstrueerime rea juhtumeid, millest igaühel on mingi vahepealne elamus. Kuid selles pole ju midagi valesti! Vahepealsed süsteemid on lihtsalt kvaalide inversiooni kergemad juhtumid ning seega mitte problemaatilisemad kui äärmuslik juhtum.

Kindlasti ei pruugi olla ilmne, missugused nimelt on vahepealsete süsteemide elamused. Võib-olla pole meie tavalises värvusruumis ühtki värvust, mis sellesse rolli sobiks, sattumata vastuollu süsteemi kategoriseerimis- ja eristamismustritega. Võib-olla on nende värvuselamused sootuks teistsugused, niisugused, mida minul olla ei saagi, kuid mis moodustavad sellegipoolest kontiinumi punasest siniseni. See oleks kummaline, kuid mitte väga uskumatu. Oluline on see, et raskust, mida me nägime hääbuvate kvaalide juhtumi puhul, ei teki siin üldse. Need süsteemid ei mõista oma elamuse omadusi süstemaatiliselt vääriti. Kui nad ütlevad, et nende elamustes on erinevused, võivad nende elamustes ollagi erinevused; kui nad ütlevad, et neil on intensiivne elamus, võib neil ollagi intensiivne elamus; jne. Kindlasti on värvused, mida nad nimetavad “punaseks”, erinevad sellest värvusest, mida nimetan punaseks mina. Kuid see ei valmista meile mingit raskust — nii on see juba tavalise inversiooni puhulgi. Oluline on see, et erinevalt hääbuvate kvaalide juhust jäävad nende süsteemide elamuste struktuursed omadused täpselt samadeks.

Sellegipoolest võib nende mõttekäikude abiga leida hea argumendi kvaalide inversiooni võimalikkuse vastu.[[323]](#footnote-324) Oletame jällegi *reductio* eesmärgil, et kvaalide inverteerumine on empiiriliselt võimalik. Siis on võimalik kahe funktsionaalselt isomorfse ja ühesuguses funktsionaalses seisundis, kuid erinevate elamustega süsteemi olemasolu. Oletame näitlikkuse huvides, et need süsteemid oleme mina oma punase elamusega ja minu ränipõhine isomorf oma sinise elamusega. (Üht üldistatavusse puutuvat peensust, mis siin tekib, vaatlen allpool​.)

Konstrueerime, nagu ennegi, vahepealsete süsteemide jada, mille ühes otsas olen mina ja teises minu isomorf. Edasi aga kulgeb argumentatsioon teisiti. Me ei pea muretsema selle pärast, kuidas elamused jada mööda liikudes muutuvad. Võib-olla toimub see järsku, võib-olla pöördub kogu pilt pahupidi — kuigi ​**[lk.267]**​ kõige mõistlikum on vist arvata, et muutused on sujuvad. Oluline on vaid see, et jadas leidub kaks niisugust punkti A ja B, et (1) süsteemide A ja B vahel ei vahetu rohkem kui kümnendik ajust, ja et (2) süsteemi A elamused erinevad oluliselt süsteemi B elamustest. Et näha, et see peab nii olema, tarvitseb vaadelda väiksemat jada, mis koosneb ainult esialgsest süsteemist ning süsteemidest, milles on asenda­tud 10 prot­senti, 20 protsenti, …, 90 protsenti, 100 protsenti ajust[[324]](#footnote-325). Punane elamus ja sinine elamus on niivõrd erinevad, et selles väiksemas jadas peab leiduma kaks naabrit, mille elamused on oluliselt erinevad (s.t. piisavalt erinevad, et erinevus oleks märgatav, kui mõlemad elamused oleksid samal isikul) — ei ole ju võimalik jõuda punasest siniseni kümne märkamatu hüppega.

On tõsi, et elamuste vahel võib olla märkamatuid erinevusi. Kui punase värvuse varjund muutub küllalt vähe, ei suuda ma erinevust märgata. Võiks arvata, et see on nii sellepärast, et erinevus on ainult maailmas, aga mitte elamuses; ent kui asi piirduks sellega, siis võiks teha sellise muutuse tuhat korda ja viimaks näidatagi, et punane ja sinine värvus kutsuvad esile samu elamusi — mis oleks naeruväärne. Seega võib elamuses võib olla ka märkamatuid erinevusi. Seda fenomeni võib täheldada, kui vaadata suurt pinda, mille värvus on eri kohtades väga nõrgalt erinev; kohati on äärmiselt raske öelda, kas pinna eri osad kutsuvad esile samu või erinevaid värvuselamusi. Peaasi on, et need märkamatud erinevused on väga väikesed. Parimal juhul võivad kümme sellist hüpet viia meid punase värvuse ühest varjundist sama värvuse väga pisut teistsuguse varjundini. (See punkt toob argumendi üldisusse väikese lünga; tulen selle juurde hiljem tagasi.)

Seega peab punase ja sinise süsteemi vahel leiduma kaks oluliselt erinevate elamustega süsteemi, mis aga ei erine oma seesmise ehituse poolest rohkem kui 10 protsendi ulatuses. Näitlikkuse mõttes võtame nendeks süsteemideks minu ja Billi. Seal, kus minul on punane elamus, on Billi elamus veidi teistsugune. Me võime väga hästi oletada, et Bill näeb sinist; võib-olla on tema elamus minu omaga sarnasem, kuid mõttekäiku see ei muuda. Need kaks süsteemi erinevad ka selle poolest, et seal, kus mingis minu aju väikeses piirkonnas on neuronid, on Billi ajus ränikiibid. See ränikontuuri asendus neuronkontuuriga on ainus füüsiline erinevus minu ja Billi vahel.

Meie mõttelise katse otsustav samm on võtta samasugune ränist kontuur nagu see, mis on Billil, ja paigaldada see dubleeriva kontuurina minu pähe. See kontuur on funktsionaalselt isomorfne ühega nendest kontuuridest, mis minu peas juba on. Me varustame kontuuri andurite ja täituritega, et ta saaks suhelda minu aju teiste osadega. Kuid me ei ühenda teda ajuga otse, vaid lüliti kaudu, mis võib tööle panna kord neuronitest, kord ränikiipidest kontuuri. Nupule vajutades lülitub neuronkontuur välja ja tema töö võtab üle ränikontuur. Võime kujutleda, et lüliti on ühenduses liitumispunktidega, millede kaudu ​**[lk.268]**​ kõnealune kontuur suhtleb aju teiste osadega. Lüliti katkestab neuronkontuuri ühendused ja ühendab nende asemele ränikontuuri täiturid. (Mõlema kontuuri andurid võivad jääda paigale, nii et nende seesmised seisundid muutuvad samamoodi nagu enne, kuid aju protsesse mõjutab ainult üks kontuur korraga. Ent katset võib teha ka nii, et lüliti ühendab ja lahutab nii andurid kui täiturid, tagamaks, et dubleerival kontuuril puuduks igasugune ühendus ülejäänud süsteemiga. See muudaks mõningaid üksikasju, kuid mitte loo moraali.)

Nupulevajutus suunab juhtimisahela otsekohe ümber, nii et infotöötlus, mis toimus enne närvikontuuris, hakkab nüüd toimuma ränikontuuris. Kuid minu funktsionaalne organisatsioon jääb täpselt samasuguseks nagu enne ümberlülitamist. Kaks süsteemi erinevad teineteisest põhimõtteliselt ainult ühe kontuuri füüsiliselt ehituselt. Ka teise, “ripneva” kontuuri füüsilises ehituses on erinevusi, kuid need ei ole funktsionaalse organisatsiooni seisukohalt olulised, kuna neil pole rolli süsteemi teiste komponentide mõjutamises ja käitumise suunamises.

Mis juhtub nupule vajutades minu elamusega? Enne uue kontuuri paigaldamist oli elamus punane. Pärast paigaldamist, kuid enne ümberlülitamist on ta arvatavasti endiselt punane, sest erinevus on ainult ühes lisandunud kontuuris, mis ei võta mitte kuidagi ajuprotsessidest osa ega mõjuta neid rohkem kui siis, kui ma oleksin ta lihtsalt ära söönud. Pärast ümberlülitamist olen ma aga enam-vähem samasugune süsteem nagu Bill. Ainus erinevus Billi ja minu vahel on nüüd see, et minu süsteemis on tegevusetu ja põhjuslikult irrelevantne neuronkontuur. (Võime koguni kujutleda, et uue kontuuri sisselülitamine hävitab selle.) Meie eelduse kohaselt oli Billi elamus sinine. Seega on pärast ümberlülitamist ka minul sinine elamus.

Juhtub niisiis see, et minu elamus muutub “otse minu silma all”. Seal, kus ma varem aistisin punast, aistin ma nüüd sinist. Elamus, mida kutsub esile õun minu laual, on nüüd järsku sinine. Võime mõtteliselt isegi vajutada lülitit mitu korda edasi-tagasi, nii et elamus hakkab minu silme ees “hüplema”.

Esimesel pilgul tundub, et selle pildiga on kõik korras — ja tal on ju iseäralikku võlugi —, kuid ometi toimub siin midagi väga kummalist. Minu elamused muutuvad ühe hüppega punastest sinisteks, kuid mina ei märka mingit muutust! Isegi kui lülitit mitu korda vajutada, nii et minu kvaalid hakkavad siia-sinna hüplema, teen mina edasi seda, mida ma seni tegin, ega märka midagi ebatavalist. Hüpoteesi kohaselt jääb minu funktsionaalne organisatsioon kogu aeg normaalseks. Ja see tähendab, et minu funktsionaalse organisatsiooni ajalised muutused on pärast ümberlülitamist täpselt samasugused, nagu nad oleksid olnud ilma selleta. Minu käitumissoodumustes ​**[lk.269]**​ ei muutu eriti midagi. Mul ei ole korraga tahtmist öelda: “Hmm! Midagi imelikku on toimumas!” Ei ole ruumi võpatuseks, hüüatuseks ega isegi tähelepanu hajumiseks. Igasugune ebatavaline reaktsioon tähendaks, et kahe kontuuri vahel on ikkagi funktsionaalne erinevus, ja see on vastuolus eeldatud isomorfismiga. Minu kognitiivne organisatsioon jääb katseplaani järgi täpselt samasuguseks, nagu ta tavaliselt ikka on, ja seega ka samasuguseks, nagu ta oleks olnud ilma nupuvajutuseta.

On selge, et ükskõik kuidas uskumusi ka funktsionaalselt tõlgendada, kontuuride ümberlülitamise käigus ma uusi uskumusi omandada ei saa. Isegi kui funktsionaalse käsitlusega mitte nõustuda, on ikkagi äärmiselt ebatõenäoline, et lihtne neuronkontuuri asendamine ränikontuuriga üldise organisatsiooni samaks jäädes võiks põhjustada oluliste uute uskumuste tekkimist, nagu näiteks “Minu kvaal muutus äsja teistsuguseks.” Nagu hääbuvate kvaalide juhul, nii ka siin ei ole niisuguseks muutuseks lihtsalt ruumi, kui see ei toimu just mingis kehatus kartesiaanlikus vaimus[[325]](#footnote-326).

Seega oleme taas jõudnud *ad absurdum*​. On vist täiesti ebausutav, et minu elamused, millele ma täit tähelepanu pööran, võiksid nii põhjalikult muutuda, ilma et ma muutust märkaksin. See tähendaks jällegi radikaalset lõhet teadvuse ja kognitsiooni vahel. Kui niisugune asi oleks võimalik, siis tähendaks see, et psühholoogia ja fenomenoloogia käivad teine täiesti teist sammu — veel palju enamgi kui hääbuvate kvaalide stsenaariumis.

See “hüplevate kvaalide” stsenaarium on võib-olla loogiliselt võimalik (kuigi see on nii äärmuslik juhtum, et tundubki võimalikuna ainult loogiliselt), kuid see ei tee teda veel empiirilise võimalusena usutavaks — sama vähe, nagu on usutav, et maailm loodi viis minutit tagasi. Empiirilise hüpoteesina tundub olevat palju tõepärasem, et kui ratsionaalse ja ärksa inimese elamused oluliselt muutuvad, siis peaks ta olema võimeline muutust märkama. Vastasel korral peaks teadvuse ja kognitsiooni vaheline ühendusniit olema küll ülipeenike.

Tõepoolest, oletus, et kvaalide hüplemine on looduslikult võimalik, viib meid ärevale mõttele: võib-olla on nad seda tõepoolest, ja võib-olla toimub midagi niisugust meiega kogu aeg. Meie funktsionaalsete mehhanismide füsioloogilised omadused muutuvad ju pidevalt. Nende funktsionaalsed omadused on küllaltki robustsed; võiks eeldada, et nende robustsuse on taganud evolutsioon. Kuid mittefunktsionaalsetel omadustel pole mingit adaptiivset põhjust muutumatuks jääda. Madaltasemel, molekulaarsetes omadustes toimub muutusi kahtlemata iga hetk. Asukoht, aatomkoostis jt. omadused võivad muutuda, ilma et muutuks funktsionaalne roll, ja selliseid muutusi toimub peaaegu kindlasti pidevalt.

Kui möönda võimalust, et kvaalid ei sõltu ainult funktsionaalsest organisatsioonist, vaid ka implementatsioonilistest seikadest, siis on võimalik, et kvaalid hüplevad meie silme ees tegelikult kogu aeg. Ei tundu olevat mingit põhimõttelist põhjust, ​**[lk.270]**​ miks neuronite asendamine ränikiipidega peaks põhjustama erinevusi, muutused neuronrealiseeringus aga mitte[[326]](#footnote-327); põhimõttelise piiri saab tõmmata ainult funktsionaalsel tasemel.[[327]](#footnote-328) Põhjus, miks me ei taha uskuda, et selline hüplemine toimuks meie endi juhul, on selles, et me tunnistame põhimõtet, mis ütleb, et kui meie elamused oluliselt muutuvad, siis me oleme suutelised seda muutust märkama. Kui me aga aktsepteeriksime kvaalide hüplemise võimalikkust esialgsel juhul, siis tuleks see põhimõte kõrvale heita ja selle abil ei saaks keegi enam kaitsta end skeptitsismi eest tavalisematelgi juhtudel.

Ei saa välistada, et taolist katset on tõepoolest võimalik sooritada ka tegelikkuses. Praktilised raskused oleksid muidugi tohutud, kuid vähemalt põhimõtteliselt on võimalik, et keegi paigaldaks minsse niisuguse kontuuri ja ma võiksin ise näha, mis seepeale juhtub, ning sellest maailmale teada anda. Kuid muidugi puudub säärasel eksperimendil tegelikult igasugune mõte, sest me teame, missugune on tulemus: ma raporteerin, et minu elamus — kogu aeg üks ja sama punane varjud — ei muutunud ja et ma ei märganud midagi imelikku. Minu seisukoht, et kvaale determineerib funktsionaalne organisatsioon, muutub veelgi kindlamaks. See ei ole muidugi veel tõestus, kuid säärasele tõendile on raske tõsimeeli vastu vaielda.

Kõige selle põhjal pean ma kõige usutavamaks hüpoteesi, et neuronite asendamisel funktsionaalse organisatsiooni samaks jäädes jäävad ka kvaalid samasugusteks. Kui hüplevate kvaalide stsenaarium tundub uskumatuna, siis võib selles süüdistada algset eeldust, et funktsionaalselt isomorfsel ränipõhisel süsteemil võib olla sinine elamus seal, kus minul on punane. Ja kõige mõistlikum, mida selles olukorras teha saab, on loobuda sellest eeldusest ja järeldada, et elamuse determineerib ainuüksi funktsionaalne organisatsioon.

Tuleb märkida, et see mõtteline eksperiment töötab kvaalide puudumise võimalikkuse vastu sama hästi kui kvaalide inversiooni võimalikkuse vastu. Me lihtsalt võtame süsteemide astmikus, mille teises otsas kvaalid puuduvad, kaks süsteemi, millede elamused on oluliselt erinevad, ja paigaldame ühele neist, nagu ennegi, dubleeriva kontuuri. Edasi läheb kõik samamoodi: kui kvaalide puudumine on võimalik, siis paneb ümberlülitamine kvaalid minu silme ees vilkuma, nii et nad on kord eredad, kord tuhmid, mina ei märka aga midagi. Ja jällegi on palju usutavam, et selline märkamata jääv hüplemine on võimatu, mistõttu on võimatu ka kvaalide puudumine.

Mina isiklikult leian, et see argument kvaalide puudumise vastu on veenvamgi kui see, mille ma esitasin [3. jaos](#ORG.7.3), kuigi mõlemal on oma roll. Oponent võib lihtsalt südame kõvaks teha ja nõustuda kvaalide hääbumise võimalikkusega, aga tunnistada kvaalide hüplemise võimalikkust on küllap suurusjärgu võrra raskem. Oluline erinevus on nähtavasti just selles, et ümberlülitamine toimub ainsa hetkega​[[328]](#footnote-329), samuti asjaolus, et nii dünaamiline ja rabav fenomen jääb subjektil märkamata. Kvaalide hääbumine tähendaks, et mõnel süsteemil pole iseenda teadvuselamusega mingit kokkupuudet, kvaalide hüplemine osutaks aga veelgi kummalisemale lüngale.

​**[lk.271]**​ Hüplevate kvaalide argumendil on oma struktuuri tõttu rohkem nõrku kohti kui hääbuvate kvaalide argumendil. Siiski tundub, et ükski neist pole niisugune, millest oponent saaks kasu lõigata. Järgnevalt arutan neid nõrkusi ning lisaks ka veel üht vastuväidet sellele argumendile. Hüplevate kvaalide argumendile saab esitada taas kõiki neidsamu vastuväiteid, mis on rakendatavad hääbuvate kvaalide argumendilegi, ja ka vastused on enam-vähem samad, mistõttu ma ei hakka neid enam kordama.

#### Esimene vastuväide: lüngad seoses kiiruse ja ajalooga

Argumenti, mille ma siin esitasin, saab loomulikul viisil laiendada neuronite​/​räni juhtumilt paljudele teistele funktsionaalsete isomorfide näidetele. Kuid on ka paar erandit, mis puudutavad kiirust ja ajalugu. Kui isomorf talitleb palju kiiremini või aeglasemalt kui originaal, siis pole mõtet loota, et kõik hakkaks normaalselt toimima, kui me lihtsalt paigaldame ühe süsteemi kontuuri teise süsteemi. Seepärast ei välista minu argument võimalust, et kiiruse muutus, mis jätab funktsionaalse organisatsiooni endiseks, põhjustab kvaalide inversiooni. Sarnane võimalus jääb ka füüsiliste isomorfide puhul, mis erinevad ainult oma ajaloo poolest: kui ma olen sündinud Lõuna-Hampshire’is, on minu elamus roheline, aga minu Põhja-Hampshire’is sündinud teisiku elamus on võib-olla punane. Hüplevate kvaalide stsenaariumis (erinevalt hääbuvate kvaalide stsenaariumist) ei saa ajalugu varieeruda, ja seega ei riiva see argument hüpoteesi, et kvaalid on mineviku suhtes superventsed.

Ent kumbki neist hüpoteesidest ei olnud ju nagunii kuigi usutav. On mõistlik oletada, et ajalugu peaks avaldama meie kvaalidele mõju meie füüsilise struktuuri kaudu. Ent sõltuvus ajaloost, millest oli juttu ülal, oleks palju tugevam: põhimõtteliselt tähendaks see, et praegustele kvaalidele avaldab “mittelokaalselt” mõju kauge ajalugu, ilma et see kuidagi kajastuks füüsilises struktuuris või lähemas aegruumilises ümbruses. Mis puudutab kiirust, siis oletus, et kiiruse muutus suudaks kvaale inverteerida, samas kui mitte miski muu seda ei suudaks, tundub üsna meelevaldne. Neis hüpoteesides pole vastuolu, nii et põhimõtteliselt võiks oponent nendega nõustuda, ent selleks pole kuigi palju põhjust. Kui me oleme juba endale selgeks teinud, et kõigis teistes muutustes, mis jätavad organisatsiooni samaks, jäävad samadeks ka kvaalid, siis pole mõte, et kiirus ja ajalugu oleksid ainsad, mis neid muuta suudavad, enam kuigi ahvatlev.

#### Teine vastuväide: Aga nõrgad inversioonid?

Teine selle argumendi pisipuudus on, et ta ei välista väga nõrkade spektriinversioonide võimalikkust. Näiteks võib tumepunase ja veel veidi tumedama punase vahel olla üheksa niisugust vahepealset varjundit, et kõik naabervarjundite paarid on eristamatud. Hüplevate kvaalide ​**[lk.272]**​ stsenaariumi see ei kõiguta: seda, et süsteem ei märka ümberlülitamisel mingit erinevust, oligi oodata.

Loomulikult ei ole arvus ⅒ kahe naabersüsteemi erinevuse määrana midagi erilist. Kuid see arv ei tohi olla liiga suur. Kui me annaksime talle väärtuseks ½, siis tekiksid kahtlused seoses isikulise identsusega: meile võidakse õigusega öelda, et nupulevajutus loob uue isiku, ja pole ju midagi imelikku selles, kui uus isik muutust ei märka. Me võiksime minna sellele raskusele põrkumata 20 või 25 protsendini; kuid nii jääks ikka veel alles väga nõrkade inversioonide võimalus, mis tekiksid nelja või viie märkamatu muutuse liitudes. Mure jääb siiski väiksemaks, kui panna tähele, et on väga ebatõenäoline, et elamus sõltuks kõigist ajupiirkondadest võrdselt. Näiteks kui värvuselamus sõltub peamiselt mingist väikesest nägemiskorteksi piirkonnast, siis tarvitseb selleks, et ühe hoobiga realiseerida suvaline kvaalide inversioon, asendada vaid väike osa süsteemist, ja argument toimiks ka kvaalide kõige nõrgema, vaevu märgatava inversiooni vastu.

Igal juhul ei ole lüngad siin ohtlikud. Parimal​ juhul jääb püsima võimalus, et organisatsioon ei determineeri elamust päris täielikult, kuigi puudujääk on äärmiselt väike. Determinismi puudujääk on võib-olla nii väike, et ei paku mingit huvi, kuid igal juhul ei saa positsiooni, milleni ta välja viib, ahvatlevaks pidada. Tundub mõistlik oletada, et elamused on kas kõigiti inverteeritavad või siis täiesti mitteinverteeritavad; miks peaks maailm olema niisugune, et inversioon võib olla ainult nõrk? See tunduks üsna meelevaldne. Me ei saa seda võimalust välistada, ent kuigi palju algset usutavust​[[329]](#footnote-330) sellel hüpoteesil ei ole.

#### Kolmas vastuväide: märkamata jäänud kvaalid

Kõnealune argument jätab ruumi ka võimalusele, et inverteeruvad need kvaalid, mida me ei märka. Näiteks kui me ei pööra tähelepanu oma nägemisvälja perifeeriale, võib seal toimuda kvaalide inversioon, ilma et me seda märkaksime. Tõepoolest, hiljutised katsed (Rensink, O’Regan ja Clark 1995) näitavad, et kujutises võib muuta väga olulisi tunnuseid, ilma et vaataja seda märkaks, kui ta ei pööra erilist tähelepanu just nendele tunnustele. (Neis katsetes on kahe pildi kuvamise vahel tavaliselt lühike ajavahe, nii et olukord pole täpselt samasugune nagu hüplevate kvaalide stsenaariumis, kuid on siiski sarnane.) Niisiis ei välista need argumendid märkamata jäänud kvaalide inverteerumise võimalust.

Ent kõigis neis mõttekäikudes pole midagi, mis viitaks võimalusele, et inverteeruda võiksid need kvaalid, mida subjekt märkas. Nii et kui keegi püüab seda lünka poleemikas ära kasutada, siis peab ta leppima väheahvatleva positsiooniga, mille järgi kvaalid on organisatsiooniliselt invariantsed, kui nad on inimese tähelepanu fookusele küllalt lähedal, vastasel korral aga ​**[lk.273]**​ sõltuvad muudest teguritest. (Ja inverteerunud roheline elamus nägemisvälja perifeerias pöördub tagasi punaseks, kui talle tähelepanu pöörata?) Võimalus sellest positsioonist midagi teoreetiliselt rahuldavat välja võtta tundub väga ebatõenäoline. Nagu ka teisedki meie argumendi lüngad, nii ka see lünk võimaldab jõuda ainult positsioonidele, millel puudub märkimisväärne usutavus​[[330]](#footnote-331) juba ette.

#### Neljas vastuväide: Kahekordne ümberlülitamine

Veel üks vastuväide[[331]](#footnote-332): kujutame ette sarnast eksperimenti, kus me lülitame punaste ja siniste sisendite ühendused tsentraalsetesse ajupiirkondadesse ümber nii, et sinised stiimulid hakkavad mängivad rolli, mida mängisid varem punased stiimulid, ja samas lülitame kompensatsiooniks süstemaatiliselt ümber ka tsentraalsest piirkonnast perifeeriasse suunduvad ühendused. Kui sinine sisendstiimul viib tsentrumi seisundisse, mis varem seostus punase värvusega, siis nüüd on ühendused tsentrumi ja aju ülejäänud osade vahel ümber suunatud nii, et aju teised osad funktsioneerivad täpselt samamoodi, nagu enne kõiki neid ümberkorraldusi. Nii muutub minu elamus peaaegu kindlasti punasest siniseks, kuid minu käitumissoodumused jäävad kõigiti samaks. Sel juhul oleks peaaegu kindel, et korduvate ümberlülitamiste toimel hakkaksid kvaalid hüplema. Kas ei tule seega lõpuks välja, et hüplevate kvaalide idee on ikkagi mõistlik?

Esiteks tuleb märkida, et niisugune ühenduste ümbersuunamine oleks tohutult mahukam töö kui kõik teised, mida ma olen seni kirjeldanud. Aju tsentraalne osa mõjutab teisi osi väga mitmesugustes kohtades. Kõik need ühendused tuleb ümber korraldada, kusjuures tuleb tähele panna, et ühelgi juhul ei piisa lihtsast ühenduste ümbervahetamisest. Me ei saa niisama lihtsalt asendada “punaseid väljundeid” “siniste väljunditega”, nagu see oli võimalik sisendite puhul, kuna tsentraalsüsteemi väljundandmed võivad representeerida väga mitmesuguseid asju — mälestusi, motoorseid käske jne. —, nii et siin ei piirdu erinevus sinise ja punase väljundi vahel lihtsalt “polaarsusega”. Sobiva “sinise väljundi” kindlaksmääramiseks tuleks tõenäoliselt simuleerida kõiki protsesse kogu tsentraalses ajupiirkonnas, arvestades selle algseisundit ja sisendstiimuleid, et näha, mis sealt välja tuleb. Ja sellisel juhul ei tee kogu põhjuslikku tööd enam tsentrum, vaid simulatsioon, ja stsenaarium minetab oma jõu argumendina.

Teiseks paneme tähele, et isegi kui peaks selguma, et väljundeid saab mingil lihtsal viisil ümber kommuteerida, säilivad seejuures ainult käitumissoodumused, ent mitte funktsionaalne organisatsioon. Mis tunne katseisikul siis olla võiks? Ma kujutan ette, et sellisel juhul ma märkaksin muutust ja püüaksin tegutseda sellele vastavalt, kuid mulle tunduks, otsekui sekkuks minu tegevusse mingi nukujuht. Erinevalt eelmisest juhtumist leiduks nendele uutele uskumustele ja muudele kognitiivsetele seisunditele seekord koht; nende aluseks oleksid mitmesugused tsentraalpiirkonna seisundid. Ja me võime ette kujutada, et kui tsentrumi sisendandmed tagasiside korras näitavad, et motoorsed liigutused on olnud hoopis teistsugused kui ​**[lk.274]**​ kavandatud, põhjustab see tsentraalpiirkonna seisundis tugeva vapustuse. Tegelikult viib see meid tagasi esimese vastuväite juurde, sest neid tagasisideefekte näib olevat peaaegu võimatu süstemaatiliselt kompenseerida. Igal juhul järeldub funktsionaalse organisatsiooni olulisest erinevusest, et need juhtumid ei ole analoogilised.

### 6. Mittereduktiivne funktsionalism

Teeme kokkuvõtte. Kui kvaalide puudumine on võimalik, siis on võimalik ka kvaalide hääbumine; kui kvaalide inversioon on võimalik, siis on võimalik ka kvaalide hüplemine; ja kui kvaalide puudumine on võimalik, siis on võimalik ka kvaalide hüplemine. Kuid kvaalide hääbumise võimalikkus on ebausutav, ja kvaalide hüplemise võimalikkus on äärmiselt ebausutav. Seega on ka kvaalide puudumine ja kvaalide inversioon äärmiselt ebausutavad. Kõigest sellest järeldub, et meil on igati põhjust uskuda, et organisatsioonilise invariantsuse printsiip on tõene ning et funktsionaalne organisatsioon determineerib teadvuselamuse täielikult.

Tuleb märkida, et need argumendid ei sea jalule funktsionalismi oma kõige tugevamas mõttes, sest parimal juhul näitavad nad vaid seda, et kvaalide puudumine ja inverteerumine on empiiriliselt (või looduslikult) võimatud. Neid argumente ei saa laiendada argumentideks ka kvaalide puudumise ja kvaalide inversiooni loogilise võimatuse kasuks, nagu sooviksid mõned funktsionalistid. Selleks on kaks põhjust. Esiteks näib, et ei hääbuvate kvaalide ega hüplevate kvaalide hüpoteesis ei ole kogu nende ebausutavusest hoolimata mingit vastuolu. Nende hüpoteeside loogilise võimalikkuse vastu saaks vaielda, öeldes näiteks, et kvaalide mõistet konstitueeribki muu hulgas see, et me suudame märgata nende vahel erinevusi. See intuitsioonipärane kontseptuaalne on vaieldav, kuid nii või teisiti leidub ka teine põhjus, miks need argumendid ei suuda tõestada, et funktsionaalne organisatsioon elamusi loogiliselt determineerib.

Et seda teist põhjust näha, paneme tähele, et need argumendid lähtuvad empiirilise eeldusena mõnedest faktidest funktsionaalse organisatsiooni leviku kohta füüsikalistes süsteemides: sellest, et mul on mingit kindlat liiki teadvuselamusi, või et neid on mingitel bioloogilistel süsteemidel. Kui me suudaksime tõestada, et kvaalide hääbumine ja kvaalide hüplemine on loogiliselt võimatud, siis saaks selle põhjal näidata järgmise tingimuslause loogilist paratamatust: kui mingil süsteemil on mingit kindlat liiki teadvuselamusi, siis on samasugused elamused igal süsteemil, mille peeneteraline funktsionaalne organisatsioon ühtib tema omaga.[[332]](#footnote-333) Kuid me ei saa tõestada selle tingimuslause väite​[[333]](#footnote-334) loogilist paratamatust, tõestamata, et tingimus on loogiliselt paratamatu, tingimus ise on aga empiiriline. Tõestamaks, et elamus tuleneb loogilise paratamatusega funktsionaalsest organisatsioonist, tuleks kõigepealt tõestada, et elamus on füüsika suhtes loogiliselt superventne, ent nagu ma olen püüdnud näidata, pole seda võimalik teha. Ent isegi kui me seda suudaksime, peaks see arvatavasti toimuma funktsionaalse definitsiooni abil, ​**[lk.275]**​ aga kui selline definitsioon oleks olemas, siis järelduks kvaalide puudumise ja kvaalide inverteerumise võimatus otsekohe, ilma vigurlike argumentideta. Nii et kui me soovime tõestada kvaalide puudumise või inverteerumise loogilist ja metafüüsilist võimatust, siis pole hääbuvate ja hüplevate kvaalide argumentidest meile nagunii kuigivõrd kasu.

Seega ei suuda need argumendid tõestada funktsionalismi tugevat vormi, mis ütleb, et funktsionaalne organisatsioon on teadvuselamusele konstitutiivne. Kuid nende varal saab põhjendada funktsionalismi nõrgemat vormi, mida ma olen nimetanud *mittereduktiivseks funktsionalismiks* ja mille järgi funktsionaalne organisatsioon on teadvuselamuseks piisav loodusliku paratamatusega. Selle vaate järgi determineerib funktsionaalne organisatsioon teadvuselamuse, kuid teadvuselamus ei pruugi olla taandatav funktsionaalsele organisatsioonile.

Igal juhul on järeldus ikka veel tugev. Invariantsusprintsiip ütleb meile, et teadvusega kognitiivsüsteemid võivad põhimõtteliselt realiseeruda väga erinevates meediumides. Eelkõige annab see järeldus palju tuge tehisintellektiteadlaste ambitsioonidele; sellest tuleb lähemalt juttu [9. peatükis](#ORG.9). Kui mittereduktiivne funktsionalism peab paika, siis ei takista teadvuse mitteredutseeritavus teadvusega arvutisüsteemi konstrueerimist.

Peaasi on, et meie püüdlustel kitsendada printsiipe, mille varal teadvus on füüsika suhtes looduslikult superventne, on olnud edu. Me oleme ahendanud superventsusbaasis leiduvate relevantsete omaduste hulga organisatsiooniliste omaduste hulgani. Mingis mõttes võime öelda, et teadvus ei ole superventne mitte ainult füüsika, vaid ka organisatsiooni suhtes. Siin tuleb sõnu hoolega valida, kuna iga süsteem realiseerib hulgaliselt väga mitmesuguseid funktsionaalseid organisatsioone. Kuid me võime öelda järgmist: igas füüsikalises süsteemis, milles sünnib teadvuselamus, realiseerub mingi niisugune funktsionaalne organisatsioon, et igal süsteemil, milles realiseerub samasugune organisatsioon, on loodusliku paratamatusega samasugused teadvuselamused[[334]](#footnote-335). Sobiva organisatsiooni leidmiseks tuleb vaadelda süsteemi küllalt peeneteraliselt, et oleks võimalik fikseerida kognitiivseid seisundeid, näiteks otsustusi. Seda võib omakorda saavutada, kui nõuda, et see organisatsioon oleks piisavalt peeneteraline, et fikseerida käitumist produtseerivaid mehhanisme ja käitumissoodumusi. Mitte midagi muud hääbuvate ja hüplevate kvaalide argumendid ei nõua, ja seepärast ei ole midagi enamat tarvis ka organisatsiooniliseks invariantsuseks.

Seega kehtib seadus, mis ütleb, et mingite kindlate funktsionaalsete organisatsioonide realiseerumisega kaasneb mingit kindlat liiki teadvuselamus[[335]](#footnote-336). See ei tähenda, nagu tuleks seda pidada fundamentaalseaduseks. Oleks kummaline, kui universumis oleks fundamentaalseadusi, mis seostaksid keerukaid funktsionaalseid organisatsioone teadvuselamustega. Pigem võiks eeldada, et see seadus tuleneb lihtsamatest, fundamentaalsematest psühhofüüsikaseadustest. Esialgu aga toimib organisatsioonilise invariantsuse printsiip tugeva kitsendusena lõplikule teooriale.

## ​**[lk.276]**​ 8. peatükk. Teadvus ja informatsioon: mõned spekulatsioonid

### 1. Teel fundamentaalse teooria poole

Siiani on meil õnnestunud tuua välja teadvuse ja füüsikaliste protsesside vahel mõned seosed, mis väärivad psühhofüüsika seaduste nime: esiteks koherentsusprintsiip, mis ühendab teadvust teadlikkusega või globaalse kättesaadavusega; teiseks spetsiifilisem struktuurse koherentsuse printsiip, mis seostab teadvuse struktuuri teadlikkuse struktuuriga; ja kolmandaks organisatsioonilise invariantsuse printsiip. Neist printsiipidest võivad saada valmiva teadvuseteooria komponendid. Nad lubavad meil ennustada ja isegi seletada mõningaid teadvuselamuse fakte füüsikaliste faktide abil. Ja igatahes seavad nad tulevase teadvuseteooria vormile kitsendusi: teooria, mis ei oleks nende seadustega ühitatav, oleks üsna tõenäoliselt väär. Aga tarvis on veel midagi. Need printsiibid üksi ei moodusta ühtekokku veel kaugeltki valmis teooriat.

Häda on selles, et ükski neist printsiipidest ei ole sobiv kandidaat teadvuseteooria fundamentaalseaduse kohale. Kõik nad väljendavad üsna kõrgtasemelisi reeglipärasid. Näiteks teadlikkuse (või globaalse kättesaadavuse) mõiste on kõrgtaseme mõiste, kusjuures mõnevõrra ebamääraste piirjoontega; on väga ebatõenäoline, et niisugune mõiste figureeriks mõnes fundamentaalses seaduses. Organisatsioonilise invariantsuse printsiip väljendab teatavat seaduspärasust võib-olla vähem ebamääraselt, kuid kaugeltki mitte fundamentaalsel tasemel. Teine häda on, et need printsiibid ütlevad äärmiselt vähe keha–vaimu seose olemuse kohta. Väga palju mitmesuguseid küsimusi selle seose kohta on jäänud vastuseta. Näiteks: Missugusest organisatsioonist võib sündida teadvuselamus? Kui lihtsaks saab organisatsiooni teha, enne kui elamus kaob? Ja kuidas ennustada elamuse spetsiifilist iseloomu (mitte ainult struktuuri​**[lk.277]**​) tema füüsilise aluse põhjal? Täielik teadvuseteooria peaks andma nendele küsimustele vastused, kuid meie senistest printsiipidest selleks ei piisa.

Teooria lõpuleviimiseks on meil tarvis fundamentaalsete füüsikaseadustega analoogilist psühhofüüsikaseaduste süsteemi. Need fundamentaalseadused (ehk *alusseadused*​) tuleb sõnastada tasemel, mis ühendaks elamuse alusomadusi füüsikalise maailma elementaarsete omadustega. Seadused peavad olema täpsed ja ei tohi jätta ruumi ebamäärasusele. Ühenduses mingi süsteemi füüsikaliste faktidega peavad nad võimaldama täielikult ennustada kõiki selle süsteemi fenomenaalseid fakte. Veelgi enam, nii nagu füüsika alusseadustest saab (​vähemalt kui lisada neile rajatingimused) tuletada kõik kõrgtaseme füüsikaseadused ja reeglipärad, samamoodi peavad teadvuse alusseadustest olema tuletatavad ja nende põhjal seletatavad kõik väga mitmesugused kõrgemate tasemete seadused, näiteks koherentsusprintsiibid ja organisatsioonilise invariantsuse printsiip. Kui me oleme saanud kätte fundamentaalsete seaduste, nii füüsika- kui ka psühhofüüsikaseaduste süsteemi, on meil mingis mõttes võimalik mõista universumi alus­struktuuri.

See on kõrge latt, milleni me ei hakka veel niipeagi küündima. Aga me võime vähemalt selles suunas liikuda. Organisatsioonilise invariantsuse printsiip ja struktuurse koherentsuse printsiip seavad fundamentaalse teooria vormile juba praegu rangeid kitsendusi, mistõttu põhiliste konstruktide kohale, millest võiksid saada teooria fundamentaalsed komponendid, ei ole kuigi palju kandidaate. Käesolevas peatükis pakun ma välja mõned ideed, mis võiksid meid fundamentaalsele teooriale lähemale viia. Ma ei esita täisväärtuslikku teooriat koos ammendava alusseaduste süsteemiga, kuid teen ettepanekuid konstruktide kohta, mida need seadused kasutada võiksid, ja seaduste üldise kuju kohta. Seda kõike võiks võtta kui *prototeooriat* — skeletti, mille ümber saaks ehitada teooria.

Selles peatükis on ideed märksa visandlikumad ja spekulatiivsemad kui raamatu teistes osades, ja küsimusi, mida nad tõstatavad, on sama palju kui vastuseid. Ka võivad eriti just selles peatükis esitatud ideed väga hästi osutuda täiesti ekslikeks. Need on uitmõtted, mida esitades ma ei ole püüdnud luua mudelit, mis peaks vastu rangele filosoofilisele teimimisele, vaid ainult käia välja mõnedki ideed. Me peame hakkama mõtlema fundamentaalsetele teadvuseteooriatele, ja võib-olla on alljärgnevas jutus midagi, millest võib edaspidi kasu olla.

### 2. Informatsiooni aspektid

Peamine mõiste, millest ma siin peatükis juttu teen, on *informatsioon*​. Praegusaegses ideederuumis on hõljumas palju erinevaid arusaamu informatsioonikäsitlusi, ja see tähendab, et kõigepealt tuleb teha selgeks, millest me õieti räägime. Informatsiooni mõistel, mis mind siin huvitab, on palju ühist sellega, mida käsitles Shannon (1948). Pakun käesolevas peatükis välja ühe võimaluse selle idee kohandamiseks ja edasiarendamiseks.​**[lk.278]**​ Ma ei hakka oma mõtteid väga formaalselt sõnastama, vaid kasutan formalismi parajasti niipalju, nagu on tarvis, et tabada informatsiooni mõiste kõige relevantsemaid aspekte. Käesolevas jaos on palju tehnilisi peensusi, järgnevad jaod on lihtsamad.

Shannonit ei huvitanud *semantiline*​ arusaam informatsioonist, mille kohaselt informatsioon on alati informatsioon millegi kohta. Teda huvitas formaalne või süntaktiline arusaam informatsioonist, mille aluseks on ühe võimaluse[[336]](#footnote-337) valik mingist hulgast. Kõige elementaarsem informatsioonihulk on *bitt*​, mis kujutab valikut kahe võimaluse vahel: me ütleme, et ühe võimaluse valik ruumis, mis koosneb kahest võimalikust väärtusest (0 ja 1), kannab informatsiooni üks bitt.[[337]](#footnote-338) Keerulisematel juhtudel kannab informatsiooni teade​[[338]](#footnote-339), mis on valitud võimalike kahendteadete ruumist, näiteks “0110010101”. Shannoni käsitluses ei ole oluline mitte nende valikute sisuline külg, vaid ainult valitud väärtuse spetsiifiline koht võimaluste ruumis.

Seda ideed võib formaliseerida *inforuumi*​ mõiste abil. Inforuum on abstraktne ruum, mis koosneb mitmest seisundist — mida ma nimetan *infoseisunditeks*​ — ja nende seisundite vaheliste *erinevussuhete*​ alusstruktuurist. Lihtsaim mittetriviaalne inforuum koosneb kahest seisundist, mille vahel on elementaarne erinevus. Me võime vaadelda neid seisundeid kui kaht *bitti* väärtustega 0 ja 1. Kogu nende olemuse ammendabki nende erinevus teineteisest. See tähendab, et seda inforuumi iseloomustab täielikult tema erinevusstruktuur.

Teistel inforuumidel on keerukam kuju. Keerukus võib suureneda kahel viisil: esiteks võib muutuda keerukamaks väärtuste vaheliste erinevuste struktuur, ja teiseks võib seesmine struktuur olla väärtustel endil. Näiteks esimese võimaluse kohta võib tuua ruumi, mis koosneb neljast võimalikust väärtusest 0, 1, 2 ja 3. Teise võimaluse näiteks on ruum, mille väärtuste hulka kuulub, ütleme, “110010101”. Neid kaht võimalust saab muidugi ka kombineerida, nii et tulemuseks on ruum, mille seisundite hulka kuulub näiteks teade “233102032”. Vaatleme nüüd neid kaht keerukuse liiki lähemalt.

Alustame esimesest. Kõige ilmsemad on siin kolme väärtusega ruum, nelja väärtusega ruum jne., mille erinevusstruktuur on kahest väärtusest koosneva ruumi erinevusstruktuuri loomulik laiendus. Näiteks neljaelemendilisest ruumist valitud väärtus A, B, C või D kannab informatsiooni samal viisil, nagu seda teeb bitt. Tähiste “A”, “B” jne. iseloomul pole siin loomulikult mingit tähtsust; ka siin on ruumi juures oluline ainult tema struktuur.

Veelgi olulisem on, et inforuum võib olla ka pidev. Sellisel juhul moodustavad väärtused kontiinumi, mis on analoogiline reaalarvude kontiinumiga 0 ja 1 vahel. Sellisel ruumil on lõpmata palju seisundeid. Tal on palju keerukam erinevusstruktuur kui eelmistel: ​**[lk.279]**​ see struktuur vastab otseselt kontiinumi topoloogiale, kus väärtused võivad jääda teiste väärtuste vahepeale, nende omavaheline kaugus võib olla suurem või väiksem jne. Ent endiselt võib pidada informatsiooni kandjaks suvalist kontinuumist valitud punkti.

Inforuum võib struktuurilt olla ka mingi *n*​-mõõtmelise ruumipiirkonnaga analoogne kahe- või enamamõõtmeline kontiinum. Informatsiooni võib kanda näiteks üksainus punkt kolmemõõtmelisest ruumipiirkonnast. Kõige üldisemal juhul võib inforuumi struktuuriks olla suvalise topoloogilise ruumi struktuur, mis tähendab põhimõtteliselt mingit hulka koos tema elementide vaheliste naabrus​[[339]](#footnote-340)- või ümbrussuhetega. Täpsemad üksikasjad ei ole aga alljärgnevas käsitluses kuigi olulised, sest struktuurid, mida ma vaatlema hakkan, näiteks kontiinum, on meile intuitiivselt tuttavad.

Teist liiki keerukuse puhul on väärtustel endil seesmine struktuur: nad koosnevad mitmest lihtsamast väärtusest, mida ma nimetan *elementideks*​. Näiteks võib tuua kümnebitiste väärtusruumi, kus väärtuseks on näiteks “teade” 1001101000. Selles ruumis koosneb iga väärtus kümnest elemendist ja iga elementi võib vaadelda kuuluvana omaenda kahest väärtusest koosnevasse alamruumi, mis vastab esialgsele kaheelemendilisele väärtusruumile. Ja kogu seda ruumi võib vaadelda kui kümne iseseisva alamruumi otsekorrutist.

Seesmine struktuur võib olla ka huvitavam. Näiteks võib infoseisundi sisestruktuur olla pidev, moodustades omamoodi pideva analoogi ülalkirjeldatud kümneelemendilisele struktuurile. Niisugusel seisundil oleks lõpmatu hulk elemente, millest igaüks kuuluks omaenda alamruumi. Me võime kujutleda vastavat inforuumi kui funktsioonide ruumi üle kontiinumi (kus igale väärtusele vastaks omaenda alamruum) või mingi keerukama pideva ruumi.

Ka alamruumid võivad olla keerukad esimeses ülalöeldud mõttes: näiteks võivad iga alamruumi elemendid kuuluda mingisse kontiinumisse, nii et keerukusi on kahel tasemel. Näiteks võib iga väärtus koosneda pidevast struktuurist, mille iga element võib omandada väärtusi mingist pidevast alamruumist. Sellesse ruumi kuuluvat infoseisundit võib vaadelda kui lainekuju või mingit muud pideva määramis- ja muutumispiirkonnaga funktsiooni; see on ülalkirjeldatud diskreetsete “teadete” analoog pidevas ruumis.

Kõige üldisemal juhul võib inforuumil olla kahetasemeline struktuur: igal liitväärtusel võib olla seesmine struktuur, ja iga selle väärtuse element võib kuuluda mingisse alamruumi, millel on omaenda topoloogiline erinevusstruktuur. Esimest võiks nimetada ruumi *kombinatoorseks*​ struktuuriks, teist alamruumide *relatsiooniliseks*​ struktuuriks. Paljudel juhtudel on kõigil alamruumidel ühesugune relatsiooniline struktuur, nii et võib kõnelda lihtsalt ruumi enda relatsioonilisest struktuurist. ​**[lk.280]**​ Ruumi *täielik struktuur* on kombinatoorse struktuuri ja relatsioonilise struktuuri ühend. Allpool vaatlen ma sageli inforuume, millel on ainult relatsiooniline struktuur ilma kombinatoorse struktuurita — s.t. juhtu, kus infoseisundil on üksainus element —, sest nii läheb arutluskäik palju lihtsamaks.

See mudel vaatab täiesti mööda informatsiooni semantilisest sisust, mistõttu informatsioonil, nagu mina seda siin käsitlen, on parimal juhul vaid kaudne seos informatsiooni semantilise erikujuga, mida on vaadelnud filosoofidest näiteks Dretske (1981) ning Barwise ja Perry (1983). Mudelit saaks laiendada, nii et ta sisaldaks ka semantilist elementi, seostades iga infoseisundiga mingi semantilise sisu. Kuid praegusel kujul on mudel semantilistest kaalutlustest sõltumatu.

See formaliseering väljendab Shannoni ideed, mille järgi informatsioon tähendab sisuliselt ühe väärtuse valikut (ruumi relatsioonilises struktuuris) mitmest võimalikust ning ideed, et keeruka informatsiooni saab (ruumi kombinatoorses struktuuris) kokku seada lihtsast informatsioonist. Shannoni järgi võib informatsiooni moodustada nii üksainus bitt kui ka pikk teade, näiteks “10011010”. Shannon vaatleb ka juhtu, kus informatsioon sisaldub pidevas ruumis või pideva määramispiirkonnaga funktsioonide ruumis. Kõigil juhtudel on oluline ühe väärtuse valik üksteist vastastikku välistavate võimaluste ruumist.

Shannonit ennast huvitab sageli infoseisundis sisalduva *informatsiooni hulk*​, s.t. infoseisundi *spetsiifilisuse* mõõt inforuumis. Kui võimalike väärtuste arv inforuumis on kaks, siis sisaldab valitud väärtus ühe biti informatsiooni, kui neli, siis kaks bitti, kui ​*n*​, siis log₂ ​*n*​ bitti. Kui inforuum on alamruumide kombinatsioon, siis sisaldab iga väärtus informatsiooni, mis võrdub tema elementides sisalduvate informatsioonihulkade summaga: kümnekohaline kahendteade sisaldab kümme bitti informatsiooni. See meetod sobib diskreetsete ruumide puhul; pidevates ruumides on informatsiooni hulga määratlemine keerukam. Mind ei huvita siinkohal kuigivõrd informatsiooni hulk, vaid infoseisundid ise; nende suhet informatsiooni hulgasse võib kujutleda analoogilisena aine suhtega massisse.

#### Füüsiliselt realiseeruv informatsioon

Minu määratluse järgi on inforuumid abstraktsed ruumid, ja ka infoseisundid on abstraktsed seisundid. Nad ei kuulu konkreetsesse füüsilisse või fenomenaalsesse maailma. Ent õigel viisil vaadates võime leida informatsiooni nii füüsilises kui ka fenomenaalses maailmas. Selleks peame arutama läbi mitmesugused võimalused inforuumide ja -seisundite realiseerumiseks maailmas. Võtan vaatluse alla kõigepealt füüsilised ja seejärel fenomenaalsed realiseeringud.

​**[lk.281]**​ Tundub olevat intuitiivselt selge, et inforuumide ja -seisundite realiseeringuid leidub kõikjal kogu füüsilises maailmas. Näiteks minu laelambi lülitit võib vaadelda kui kahe väärtusega inforuumi realiseeringut — neid kaht väärtust realiseerivad tema seisundid “ülal” ja “all”. Me võime vaadelda CD-plaati kui keerukast bittstruktuurist moodustunud kombinaatorse inforuumi realiseeringut. Sarnaselt võime näha informatsiooni realiseeringuid termostaadis, raamatus ja telefoniliinis. Kuidas neid intuitiivseid ettekujutusi mõtestada?

Infoseisundeid saab seostada füüsikaliste süsteemidega loomulikul viisil, vaadeldes informatsiooni füüsilist realiseerumist Batesoni (1972) juhtmõtte “infomatsioon on eristav erinevus” [​*difference that makes a difference*​] vaimus. Kuigi minu lambilüliti asendeid võib olla pidevas vahemikus lõpmata palju, ei too suurem osa sellest varieeruvusest minu toa valgustuse mõttes kaasa mingit erinevust. Olgu lülitinupp täiesti ülal või veerandi võrra allpool, lamp põleb ühtviisi. Kui aga nihutada lülitit rohkem kui kolmandiku võrra allapoole, siis lamp enam ei põle. Valguse mõttes on lülitil ainult kaks relevantset seisundit, mida võib nimetada “ülal” ja “all”. Nende kahe seisundi erinevus on valguse seisukohalt ainus eristav erinevus. Seega võime vaadelda lülitit kui süsteemi, mis realiseerib kahe väärtusega inforuumi ja mille ühed füüsilised seisundid vastavad ühele, teised teisele väärtusele.

Üldiselt on füüsilise objektiga seostuv inforuum alati määratletav põhjusliku trajektooriga (milleks on meie juhul trajektoor lülitist valguse süttimise või kustumiseni) ja võimalike tagajärgede ruumiga selle trajektoori lõpul (meie juhul valguse sisse- või väljalülitumisega). Füüsikalised seisundid on vastavuses infoseisunditega olenevalt nende mõjust põhjuslikule trajektoorile. Kui kahel füüsikalisel seisundil on trajektoorile ühesugune mõju — näiteks valguslüliti kahel asendil, mis mõlemad lülitavad valguse sisse —, siis vastab neile üks ja sama infoseisund. Sel viisil füüsikalisi seisundeid jaotades jõuame eristavate füüsikaliste erinevuste hulgani, mis moodustavad inforuumi füüsilise realiseeringu.

Inforuumi struktuur vastab otseselt tagajärgede ruumi struktuurile, mis võib jällegi olla kas diskreetne või pidev. Näiteks laelambi puhul võib põhjuslik trajektoor viia kahe relevantse tagajärjeni — valguse sisse- või väljalülitumiseni. Seega võib lülitit vaadelda kui kahe väärtusega inforuumi realiseeringut.

Sarnaselt võime vaadelda ka pidevaid inforuume. Dimmeriga lambil muudab dimmerinupu keeramine valgusvoogu pidevas vahemikus. (Praktikas võib see vahemik olla diskreetne, kuid ma vaatlen ideaalset juhtu.) Nupukeeramise mõju valgusvoole defineerib pideva inforuumi, mille realiseeringuks on valguslüliti. Lüliti füüsilistele seisunditele, mis annavad ühesuguse valgusvoo (seisundid vahemikes, kus nupu keeramine ​**[lk.282]**​ ei mõju, või seisundid, mis erinevad üksteisest ebaoluliste parameetrite, näiteks lüliti värvuse poolest), vastab üksainus infoseisund. Infoseisundite ruum on topoloogiliselt struktuurilt kontiinum, kus seisundite erinevuste struktuurile vastab erinevuste struktuur nupukeeramise toimetes valgusvoo muutumisele.

Samamoodi võib analüüsida ka informatsiooni, mida realiseerib CD-plaat. Plaadil on lõpmatu arv võimalikke füüsikalisi seisundeid, ent oma toime seisukohalt CD-mängijale realiseerib ta lõpliku hulga võimalikke infoseisundeid. Paljud muutused kettal — mikroskoopilised muutused, mida optiline lugemisseade oma lahutusvõime tõttu ei erista, väike kriimustus tööpinnal või suur pilt tagaküljel — ei mõjuta lugeja tööd. Ketta infoseisundile on relevantsed ainult niisugused erinevused, mis kajastuvad lugemisseadme väljundis. Need on erinevused, mis tulenevad lohkudest ketta põhipinnal ja nende vahelistest põhipinna lõikudest, mis üheskoos moodustavad meie mõistes bittide järgnevuse.[[340]](#footnote-341) Igale ketta seisundile vastab üks väärtus teatavas suures inforuumis. Ühe ja sama salvestise kõigi eksemplaride füüsilistele seisunditele vastab, kui plaadi kirjutamine toimus vigadeta, üks ja sama infoseisund. Erinevate salvestistega CD-plaatidele, või ka ühe ja sama salvestisega defektsetele plaatidele, vastavad erinevad infoseisundid, kuna nende mõju plaadimängijale on erinev.

Inforuumil, mida realiseerib CD-plaat, on kombinatoorne struktuur. Iga bitt mõjutab CD-seadet teistest bittidest sõltumatult, nii et plaadi iga punkti võib vaadelda kui iseseisva kahest seisundist koosneva alamruumi realiseeringut. Kõigist neist sõltumatutest mõjudest moodustub plaadi kõigi mõjude ruumis kombinatoorne struktuur, ja seega leidub samasugune kombinatoorne struktuur ka inforuumis, mille realiseering see plaat on. Seda inforuumi võib vaadelda kui suure hulga kaheseisundiliste alamruumide otsekorrutist.[[341]](#footnote-342)

Pange tähele, et informatsioon oma füüsilises realiseeringus on informatsioon ainult niivõrd, kui teda on võimalik töödelda. Mackay (1969) ütleb selle kohta: “Informatsioon on see, mida informatsioon teeb [​*information is as information does*​]”. See on kooskõlas Shannoni enda arusaamaga informatsioonist. Shannoni “informatsioonihulk” tähendab informatsiooni spetsiifilisuse määra edastatavate seisundite ruumis, s.t. niisuguste seisundite ruumis, mis võivad mängida erinevaid rolle mõnel [[342]](#footnote-343) põhjuslikul trajektooril (ehk, nagu Shannon ütleb, mõnes sidekanalis). Shannoni jaoks on informatsioon seisund, mida on võimalik edastada, ja kaudselt defineeribki inforuumi mahu ju edastava agendi funktsioon. Informatsioon on erinevus, mis võib luua erinevusi [​*difference that can make a difference*​] edastamisel.

Seda kõike selgitab Shannoni üldtuntud diagramm (joonis 8.1) ja selle juurde kuuluvad arutlused. Infoallikaks on mingi “teadete” (infoseisundite) hulk; eeldatavasti kodeerib saatja ​**[lk.283]**​ erinevad teated erinevateks signaalideks, ning erinevad edastatud signaalid vastavad erinevatele teadetele. Tõepoolest, just selle omaduse alusel me ju otsustamegi, mis tingimustel saab erinevaid seisundeid pidada erinevateks teadeteks. Kui süsteemi kaks erinevat füüsikalist seisundit konverteeruvad samaks signaaliks, siis realiseerivad nad sama teadet. Protsessi käigus võib hiljem toimuda informatsiooni moondumine ja kadu, nagu kujutab diagrammi parempoolne osa, kuid infokanali oluline omadus on, et erinevad infoseisundid avaldavad saatja kaudu erisugust mõju. Shannon, kes füüsikaliste seisundite ja infoseisundite vahelist suhet oma selgituses otseselt ei vaatle, ütleb kõike seda pigem vihjamisi kui sõnaselgelt, ent tähelepanelikumalt vaadeldes on ilmne, et infoseisundid on eristatavad just tänu edastatavusprintsiibile.

Joonis 8.1. Infokanal Shannoni diagrammi järgi. (1. diagramm raamatust Claude E. Shannon ja Warren Weaver, *The Mathematical Theory of Communication* [Matemaatiline sideteooria], 1963. Copyright 1949 by the Board of Trustees of the University of Illinois Press. Used with permission of the University of Illinois Press.​).

Ma ei püüa anda täpseid kriteeriume inforuumi realiseeringule füüsikalises süsteemis, vaid jään mitteformaalsele “eristava erinevuse” printsiibi tasemele. Selle mitteformaalse idee formaliseerimiseks on mitu võimalust, millest mõned seavad realiseeringutele rangemaid kitsendusi kui teised. Siinkohal oleks enneaegne jääda lõplikult pidama ühele neist. Mitteformaalsele tasemele jäädes jätame endale esialgu manööverdamisruumi üksikasjades, mis hakkavad võib-olla selguma, kui meil on parem ettekujutus mõneks konkreetseks rakenduseks sobivatest tingimustest. Nende üksikasjade täpsustamine on üks osa teadvuseteooria üksikasjaliku väljatöötamise protsessist.

#### Fenomenaalselt realiseeruv informatsioon

Maailmas leiduvale informatsioonile mõeldes peetakse tavaliselt silmas informatsiooni füüsilisi realiseeringuid, kuid see ei ole ainus võimalus maailmast informatsiooni leidmiseks. ​**[lk.284]**​ Informatsiooni võib leida ka meie fenomenoloogilisest maailmast. Elamusseisunditel on loomulikke ja otseseid vastavusmustreid inforuumidega. Fenomenaalsetel seisunditel on loomulikud omavahelised sarnasuste ja erinevuste mustrid, milledest moodustub inforuumi erinevusstruktuur. Seega võime vaadelda fenomenaalseid seisundeid kui neis ruumides realiseeruvaid infoseisundeid.

Näiteks lihtsate värvuselamuste ruumil on kolmemõõtmeline relatsiooniline struktuur, millest meil on [juba](#PARA.124-5) juttu olnud. Nende elamuste sarnasus- ja erinevusmustritest saame abstraheerimise teel kolmemõõtmelise relatsioonilise struktuuriga abstraktse inforuumi, mida realiseerib fenomenaalne ruum. Iga lihtne värvuselamus vastab ühele kindlale asukohale selles ruumis. Mingi konkreetne punane elamus on üks meie fenomenaalses maailmas realiseeruvaid infoseisundeid, mingi konkreetne roheline elamus on teine seisund.

Keerulisemad elamused, näiteks elamus, mida pakub vaateväli tervikuna, on keeruka kombinatoorse struktuuriga inforuumid. Kui ma vaatan näiteks mõnd pilti, paikneb minu elamus ruumis, mille kombinatoorne struktuur vastab (vähemalt) kahemõõtmelise kontiinumi kombinatoorsele struktuurile, kusjuures selle kontiinumi igal elemendil on lihtsa värvusruumi (vähemalt) kolmemõõtmeline relatsiooniline struktuur. Värvilaikude struktuur nägemisväljas ei erine kuigivõrd kahendnumbrite struktuurist kümnekohalises teates​[[343]](#footnote-344), kuigi nii nende kombinatoorne kui ka relatsiooniline struktuur on palju keerulisemad.

Inforuumide fenomenaalseid realiseeringuid otsides ei toetu me põhjuslikule “eristava erinevuse” printsiibile, mille varal me otsisime inforuumide füüsilisi realiseeringuid, vaid elamuste seesmistele omadustele ja nende omavaheliste suhete struktuurile — nende sarnasus- ja erinevussuhetele ja nende seesmisele kombinaatorsele struktuurile. Igal elamusel on teiste elamustega loomulikud sarnasus- ja erinevussuhted ning seega on meil alati võimalik leida inforuumid, millesse nad kuuluvad.

#### Topeltaspektiprintsiip

Selline arusaam informatsioonist toob esile olulise seose füüsilise ja fenomenaalse maailma vahel: kõikjal, kus me näeme inforuumi fenomenaalset realiseeringut, näeme ka sama inforuumi füüsilist realiseeringut. Ja kui elamus realiseerib mingit infoseisundit, siis realiseerib sama infoseisundit ka elamuse füüsiline substraat.

Võtame lihtsa värvuselamuse. See elamus on mingi infoseisundi realiseering kolmemõõtmelises inforuumis. Sama ruumi realiseeringu leiame elamust kandvates ajuprotsessides: selleks on neuronites kodeeruvate representatsioonide kolmemõõtmeline ruum nägemiskorteksis. Selle ​**[lk.285]**​ kolmemõõtmelise ruumi elemendid on otseses vastavuses fenomenaalse inforuumi elementidega.

Me ei tea täpselt, missugune on nende seisundite kodeering, ja seega ei tea me täpselt, missugune on selle inforuumi füüsiline realiseering. Kuid me teame, et see peab olemas olema, kuna hilisemates protsessides ilmnevad süstemaatiliselt kõik informatsiooni realiseerumise tagajärjed. Meie ütlused võivad süstemaatiliselt varieeruda olenevalt asukohast värvusruumis, näiteks kui me hindame värvusi suhteliselt tumedamateks ja heledamateks; ja me oskame leida objektide vastavusi teiste objektidega olenevalt sarnasustest ja erinevustest nende värvustes. Seega teame, et nägemiskorteksis peavad olema mingid asjakohased erinevused, mis on edastatavad teistesse ajupiirkondadesse, andes tulemuseks kolmemõõtmelise võimalike tagajärgede ruumi. Analoogiliselt väikeste erinevustega valguslüliti seisundis on ka kaht teineteisest eristamatut elamust kandvatel seisunditel alati ühesugused relevantsed tagajärjed, ja kahe sarnaseid elamusi kandva seisundi tagajärjed on sarnased. Seega võime vaadelda nägemiskorteksit kui infoseisundite realiseeringut kolmemõõtmelises ruumis.

Sama kehtib ka keerulisemate, näiteks tervet vaatevälja hõlmavate elamuste kohta. Need realiseeruvad mingis kombinatoorses inforuumis, ja sama ruumi peavad füüsiliselt realiseerima ka elamust kandvad ajuprotsessid. Me teame, et nägemisvälja igas kohas vastab lihtelamuste erinevus erinevustele mitmesugustes hilisemates tagajärgedes, kusjuures neid hilisemaid erinevusi saab üksteisest lahutada nende asukoha põhjal nägemisväljas. Me võime näiteks anda küsimustele värvuse kohta nägemisvälja eri punktides​[[344]](#footnote-345) erinevaid vastuseid; seega vastab igale punktile omaette tagajärgede ruum​[[345]](#footnote-346). Seega peab kusagil nägemiskorteksis leiduma mingi kombinatoorse inforuumi kodeering, et kõiki relevantseid erinevusi oleks võimalik edastada järgmistele protsessidele. Võimalike relevantsete seisundite ruum on siin isomorfne võimalike elamuste ruumiga; seega näeme siin ühe ja sama infoseisundi kaht realiseeringut, nii füüsilist kui fenomenaalset.

Informatsioon ei pea tingimata kodeeruma lokaalselt, näiteks väikeses, mõnest üksikust neuronist koosnevas struktuuris. Informatsiooni realiseering võib vabalt olla holistlik, nagu mõnede holograafiliste infosalvestusvormide​[[346]](#footnote-347) puhul. Erinevused, mis vastavad relevantsetele erinevustele nägemiskorteksi seisundites, võivad paikneda hajusalt üle kogu korteksi. Ent senikaua kui need erinevused on edastatavad ja neil on relevantsed tagajärjed, on see ikkagi realiseering.

On loomulik oletada, et sellele inforuumide kaksikelule vastab mingi duaalsus ka sügavamal tasemel. Võiksime isegi mõelda, et see kahetine realiseerumine annabki võtme füüsikaliste protsesside ja teadvuselamuse ​**[lk.286]**​ vahelise fundamentaalse seose mõistmiseks. Neid siduva ühenduslüli loomiseks on tarvis mingisugustki konstrukti, ja informatsioon tundub selleks täiesti sobiva kandidaadina. Võib-olla saaks printsiipidest, millede kohaselt see informatsiooni kahekordne realiseerumine toimub, välja arendada füüsilist maailma fenomenaalse maailmaga ühendava alusseaduste süsteemi.

Selleks võiks pakkuda välja alusprintsiibi, mille järgi informatsioonil on (aktuaalses maailmas) on kaks aspekti, füüsiline ja fenomenaalne. Iga fenomenaalne seisund on mingi infoseisundi realiseering — sellesama, mida realiseerib ka aju kognitiivsüsteem. Ja ümberpöördult, vähemalt mõnede inforuumide puhul käib füüsilise realiseeringuga iga kord kaasas ka fenomenaalne realiseering.[[347]](#footnote-348)

See printsiip üksi ei vii meid täielikule psühhofüüsikateooriale veel ligilähedalegi. Ta annab sellele vaid midagi malli taolist — üldise skeemi, mille raames saab formuleerida konkreetseid seadusi. Selle printsiibi alusel teooriat välja töötades tuleb vastata väga mitmesugustele küsimustele. Näiteks: missuguste füüsilise realiseeringuga inforuumide kohta täpsemalt see alusprintsiip kehtib? Võtan selle küsimuse lähema uurimise alla [4. jaos](#ORG.8.4), esialgu aga jätan lahtiseks. Ebamäärasust on ka informatsiooni füüsilise realiseeringu definitsioonis: et viia psühhofüüsikateooria täie konkreetsuseni, tuleb täpselt teada, mis tingimustel saab kõnelda inforuumi füüsilisest realiseerumisest. Kuid need raskused kuuluvadki teooria väljatöötamise protsessi juurde.

Veel mõned olulised küsimused puudutavad selle vaate ontoloogiat. Kui tõsiselt kogu seda topeltaspektsuse juttu üldse võtta? Mil määral see mudel informatsiooni asjastab, käsitleb seda kui midagi reaalset? Kas ta eeldab, et füüsiline või fenomenaalne maailm või mõlemad on ontoloogiliselt sõltuvad informatsioonilisest maailmast? Ma jätan kõik need küsimused esialgu lahtiseks. Siinsamas peatükis [allpool](#ORG.8.5) võtan ma vaatluse alla selle vaate ontoloogia mitmesugused võimalikud tõlgendused. Mõnede jaoks neist on informatsiooni lihtsalt kasulik konstrukt psühhofüüsika seaduste iseloomustamiseks; teised annavad talle fundamentaalsema ontoloogilise rolli. Ka topeltaspektsuse ideed võtavad mõned tõlgendused tõsisemalt kui teised.

Esialgu aga jätan ma need metafüüsilised probleemid kõrvale. Kõnealust printsiipi võib käsitada lihtsalt kui füüsilist ja fenomenaalset maailma ühendavat seadust, mis oma ontoloogilistelt järelmitelt ei erine kuigivõrd siiani juba vaadeldud seadustest. Me teame juba, et elamus kasvab mingite mõnedele maailma füüsikalistele omadustele rakenduvate seaduste alusel välja füüsilisest maailmast. Ja peamine idee seisneb siin selles, et alustasemeks, millel need seadused füüsilisele maailmale rakenduvad, on informatsiooni füüsiliste realiseeringute tase.

Muidugi ei pruugi informatsioon olla füüsikalise maailma elementaarne omadus, nagu seda on mass ja laeng. Kuid selleks, et mängida rolli fundamentaalsetes psühhofüüsikaseadustes, ​**[lk.287]**​ ei peagi ta elementaarne olema. Kõik vajalikud fundamentaalsed omadused on meil juba olemas — nendeks on füüsikalised ja fenomenaalsed alusomadused. Tarvis on veel vaid konstrukti, mis need kaks valda omavahel ühendaks. Informatsioon tundub olevat lihtne ja sirgjooneline konstrukt, mis sobib seesuguseks ühenduslüliks hästi ja mille abil on lootust jõuda lihtsate ja ammendavate seadusteni. Kui meil õnnestuks leida selliste seaduste süsteem, siis oleks meil tõepoolest käes ka fundamentaalne teadvuseteooria.

Võib-olla aga selgub, et fundamentaalsena võib näha informatsiooni ennast. Mõtet, et füüsika uurimisobjektiks on viimselt informatsioon, on mõned füüsikud juba avaldanud. Kui sellest ideest asja saab, võib selguda, et füüsiline maailm ongi informatsioonilise maailma tuletis, ja sellele positsioonile annaks välja töötada juba väga korraliku ontoloogia. Mõnest siia puutuvast ideest tuleb juttu [siinsamas peatükis allpool](#ORG.8.5.1).

### 3. Mõned tugiargumendid

Mul ei ole argumente, mis tõestaksid lõplikult, et füüsikaliste protsesside ja teadvuselamuse vahelise seose võtmeks on informatsioon. Kuid mõningat kaudset kinnitust võib sellele ideele leida. Esimest seesugust kinnitust olen ma [juba](#ORG.8.2.3) käsitlenud: nimelt tähelepanekut, et ühed ja samad inforuumid realiseeruvad nii füüsiliselt kui fenomenaalselt. Järgnevas mainin veel üht suuremat ja kaht vähemat toetusallikat.

Need kaks vähemat toetusallikat tulenevad asjaolust, et kaheaspektiline vaade informatsioonile on kooskõlas juba varem välja töötatud psühhofüüsikaprintsiipidega, eelkõige struktuurse koherentsuse printsiibiga ja organisatsioonilise invariantsuse printsiibiga. Need printsiibid seavad fundamentaalsele teooriale tugevaid kitsendusi, mille rahuldamise võimalused ei ole ilma pikemata ilmsed, ja seega kõneleb informatsioonilise kontseptsiooni ühitatavus mõlemaga tugevasti tema kasuks.

Eriti kerge on näha ühitatavust struktuurse koherentsuse printsiibiga: mingis mõttes ongi informatsiooniline kontseptsioon spetsiaalselt loodud just selle kitsenduse rahuldamiseks. Elamuse struktuur pole midagi muud kui fenomenaalselt realiseeruva inforuumi struktuur, ja teadlikkuse struktuur pole midagi muud kui füüsiliselt realiseeruva inforuumi struktuur. Esimese punkti õigsuse näitamiseks paneme tähele, et see, mida ma olen nimetanud elamuse [implitsiitseks](#PARA.124-5) struktuuriks, vastab inforuumi relatsioonilisele struktuurile, ja see, mida ma nimetasin elamuse eksplitsiitseks struktuuriks[[348]](#footnote-349), vastab selle ruumi kombinatoorsele struktuurile. Teise punkti mõistmiseks paneme tähele, et erinevused teadlikkuse struktuuri üksikasjades on definitsiooni järgi erinevused, mis toovad kaasa erinevusi [​*differences that make a difference*​] hilisemates protsessides, kuna nad on otseselt kättesaadavad globaalseks juhtimiseks ning osutuvad seega mingi inforuumi füüsiliseks realiseeringuks. Asjaolust, et neis kahes punktis on ​**[lk.288]**​ faktiliselt juttu ühe ja sama inforuumi realiseeringutest, tulenebki struktuurse koherentsuse printsiip.

Tuleb märkida, et topeltaspektiprintsiip üksi ei garanteeri teadlikkuse struktuuri projitseerumist elamusse. Et näha, et niisugune projitseerumine tõesti toimub, tuleb näidata, et füüsikaline inforuum[[349]](#footnote-350) kuulub ruumide hulka, mille kohta topeltaspekti printsiip kehtib. Selleks oleks tarvis topeltaspektiprintsiibi detailsemat versiooni, mis asjassepuutuvate inforuumide hulka sobival viisil kitsendaks, nii et need sisaldaksid vähemasti informatsiooni, mis on meile tuttavatel juhtudel globaalseks juhtimiseks kättesaadav. Praegusel kujul ei ennusta topeltaspektiprintsiip veel täielikku struktuurse koherentsuse printsiipi, kuid ta on sellega vähemalt ühitatav.

Pole raske näha ka seda, et topeltaspektiprintsiip on ühitatav organisatsioonilise invariantsuse printsiibiga. Selleks paneme tähele, et kui süsteem realiseerib mingit inforuumi, siis teeb ta seda oma funktsionaalse organisatsiooni varal. Igal teisel süsteemil, mis on temaga küllalt peeneteraliselt vaadates funktsionaalselt isomorfne, on samasugune eristavate erinevuste muster, ja seega realiseerib ta sama inforuumi. Ja see tähendab, et kui minu elamused sünnivad inforuumide realiseerumise tagajärjel minu ajus, siis realiseeruvad samad inforuumid ka minu funktsionaalses isomorfis, ja temas sünnivad samasugused elamused — nagu invariantsusprintsiip ennustabki.

Fundamentaalne teadvuseteooria peab opereerima füüsikalistele omadustega, mis on organisatsiooniliselt invariantsed ja samas piisavalt lihtsad, et figureerida fundamentaalseadustes. Enamik organisatsiooniliselt invariantseid tunnuseid ei ole eriti lihtsad, enamik lihtsaid tunnuseid aga ei ole organisatsiooniliselt invariantsed. Füüsikaliselt realiseeruv informatsioon on võib-olla kõige loomulikum tunnus, mis rahuldab mõlemat kriteeriumi. Ja see asjaolu kõneleb psühhofüüsika seaduste informatsioonilise käsitluse kasuks.

#### Fenomenaalsete otsustuste seletamine

[Ülal](#PARA.106-1) nägime, et kuigi teadvusele pole võimalik leida reduktiivset seletust, on vähemalt põhimõtteliselt võimalik seletada fenomenaalseid otsustusi — otsustusi kujul “mul on teadvus”, “kas pole teadvus üks kummaline asi?”. Kuigi see asjaolu ei lükka mittereduktiivset teadvuseteooriat ümber, on ta siiski mõnevõrra ebamugav. See, et neile otsustustele saab anda seletuse, milles pole teadvusest endast juttugi, on intuitsioonivastane, kuid sellega võib leppida. Me võime siiski loota, et fenomenaalsete otsustuste seletamisel on mingi sügav seos teadvuse enda seletamisega. Tunduks ebamõistlik ja juhuslik, kui need kaks seletust oleksid teineteisest täiesti sõltumatud.

Selle mõtte saaks formuleerida seletusliku koherentsuse nõudena teadvuseteooriale. Ammendav vaimuteooria peab andma nii (mittereduktiivse) seletuse teadvusele kui ka (reduktiivse) seletuse ​**[lk.289]**​ meie otsustusele, et meil on teadvus, ning on mõistlik oodata, et need kaks seletust on teineteisega kooskõlas. Eelkõige võiksime eeldada, et need infotöötluse iseärasused, mis kannavad fenomenaalsete otsustuste moodustumisel põhikoormust, on needsamad, mis kannavad põhikoormust ka teadvuse enda tekkes. Kui see on nii, siis tuleb välja, et kuigi fenomenaalsete otsustuste seletuses ei figureeri teadvus ise, teevad seda tema juured​[[350]](#footnote-351).

Muidugi ei ole võimalik tõestada, et teadvuseteooria peab seda nõuet rahuldama. Kuid igal teadvuseteoorial, mis seda teeb, on üks tugev külg, mis teistel teooriatel puudub. Kui teooria näitab, mil viisil on fenomenaalsete otsustuste seletuses kesksel kohal samad tegurid, mis on kesksel kohal ka teadvuse enda seletuslikus aluses, siis on ta suutnud põimida need kaks ühtseks pildiks vaimust ja kaotada mingi osa muljest, et nende kahe paralleelsus on mingi pöörane juhus[[351]](#footnote-352).

Ma olen sageli mõelnud, et siin võibki peituda teadvuseteooria võti.[[352]](#footnote-353) Kõigepealt peame me andma endast parima, et mõista, miks moodustuvad teadvuse kohta otsustused. See võib osutuda keeruliseks, kuid mingeid sügavaid metafüüsilisi müsteeriume siin olla ei tohiks; põhimõtteliselt kuulub see küsimus kognitiivteaduse valdkonda. Seejärel tuleb leida selle seletuse põhijooned ja mõtlema, missugust rolli need võiksid mängida teadvuse enda teoorias. Ei ole mingit garantiid, et sel teel võib jõuda rahuldava teadvuseteooriani, kuid strateegia on lootustandev.

Reduktiivse seletuse otsimine meie otsustustele teadvuse kohta toob asjasse arvatavasti igal juhul selgust ning on üks reduktiivsete meetodite kõige väärtuslikumaid rakendusi teadvuseteooria väljatöötamisel. Tähelepanu võiks seejuures koondada küsimusele, miks peaks infotöötlussüsteem moodustama otsustuse, et tal on teadvus, ja eriti, et teadvus on nii kummaline fenomen. [5. peatükis](#ORG.5.2) ma juba ütlesin sel teemal mõne sõna, siin vaatlen küsimust põhjalikumalt. “Seletus”, mille ma kavatsen anda, on vaid usutavalt kõlav pajatus [​*just-so story*​], kuid võib loota, et empiiriliste uuringute abil saaks sellest välja arendada detailse teooria. Nende fenomenide käsitlemisest võib oodata suuri edusamme kognitiiv- ja neuroteaduses.

Jätame siis teadvuse enda hetkeks kõrvale ja koondame tähelepanu kognitiivsele infotöötlusüsteemile kolmanda isiku seisukohalt. Soovi korral kujutagem ette, et me rakendame seda seletust sombile. Mis põhjust võiks meil olla oodata, et infotöötlussüsteem moodustab selliseid otsustusi? Missugused protsessid võiksid olla näiteks otsuse taga, et süsteemil on hetkel mingi värvusaisting? Et selle üle lähemalt mõelda, kujutagem ette, mis võiks toimuda, kui me tajume värvusi.

Kui madaltaseme üksikasjusse mitte süüvida, on lugu umbkaudu järgmine. Meie silma satub mingisuguse spektriga valgusimpulss, mis aktiveerib võrkkestas mitut liiki rakke. Kolmesugused kolvikesed võtavad sellest välja informatsiooni valgushulga kohta kolmes osaliselt kattuvas lainepikkuste vahemikus ​**[lk.290]**​. Paljud esialgse valguslaine erisused kaovad kohe. Nägemisnärv edastab informatsiooni nägemiskorteksisse, kus edasised neuronprotsessid muundavad ta informatsiooniks, mis vastab väärtustele kolmel teljel — võib-olla​[[353]](#footnote-354) punase-rohelise, kollase-sinise ja akromaatilisel teljel. Pole täpselt teada, mis toimub edasi, kuid paistab, et informatsioon, mis vastab värvuse asukohale selles kolmemõõtmelises ruumis, jääb alles ning liigitub lõpuks tuttavate kategooriate “punane”, “roheline”, “pruun” jne. alla. Süsteem lisab neile tähistele verbaalsed kategooriad ja formuleerib ütluse, näiteks “ma näen praegu midagi punast”.

Vaatame nüüd olukorda süsteemi “seisukohalt”. Milliseid otsustusi ta moodustab? Muidugi moodustab ta mingi otsustuse, nagu näiteks “seal on punane objekt”. Ent kui süsteem on ratsionaalne ja refleksioonivõimeline, siis võiks oodata, et ta suudab reflekteerida ka tajuprotsessi enda üle. Võiksime küsida: missugust mõju avaldab süsteemile tajumine?

Oluline on siin see, et kui süsteem tajub punast objekti, ei ole tsentraalsetel protsessidel otsest ligipääsu ei objektile endale ega ka füüsikalistele tajumisprotsessidele, vaid ainult värvusinformatsioonile, mis on lihtsalt mingi asukoht kolmemõõtmelises inforuumis. Kui süsteemil on tarvis anda keeleline ütlus, ei saa ta öelda: “Sellelt täpilt saabub rohkesti peegelduvat valgust lainepikkusega 500–600-nanomeetrit”, sest igasugune informatsioon esialgsete lainepikkuste kohta on kadunud. Samamoodi ei saa ta raporteerida neuronstruktuuri kohta: “Impulsside sagedus on hetkel 50 hertsi”, kuna tal puudub neuronstruktuuridele otsene ligipääs. Süsteemile on kättesaadav ainult värvuse asukoht inforuumis.

Tõepoolest, mis puutub tsentraalsesse infotöötlusse, siis see lihtsalt leiab end olevat ühes selle ruumi punktis. Süsteemil on võime teha eristusi ja ta teab, et tal see võime on, kuid tal pole aimugi, kuidas ta seda teeb. Võib oodata, et ta suudab mõne aja pärast tähistada mitmesuguseid inforuumi punkte, kuhu ta satub, sõnadega “punane”, “roheline” jms., ning teada, missuguses neist seisunditest ta hetkel on. Aga kui temalt küsitakse, kuidas ta seda teab, siis ei oska ta öelda midagi muud kui “ma lihtsalt tean seda, tean otseselt”. Kui temalt küsida: “Milles seisneb erinevus nende seisundite vahel?”, siis ei oska ta anda vastuseks midagi enamat kui “Nad on lihtsalt erinevad” või “See on niisugune, too on naasugune”[[354]](#footnote-355) või “See on punane, too on roheline”. Kui süsteemilt küsitakse, mida see tähendab, ei ole tal öelda muud kui “Nad on lihtsalt erinevad, kvalitatiivselt erinevad”. Mida muud ta öelda saakski?

On loomulik oletada, et süsteem, mis võib otseselt teada oma asukohta inforuumis, ilma et tal oleks ligipääsu mingitele muudele teadmistele, nimetab neid seisundeid lihtviisil või elementaarselt erinevateks oma “kvaliteedilt”. Ja mõistagi tuleks oodata, et need erinevused ​**[lk.291]**​ mõjutavad süsteemi “vahetult”: süsteem satub nendesse seisunditesse, mis on omakorda otseselt kättesaadavad edasiste protsesside suunamiseks. Näiteks ei toetu tema teadmine oma seisundist mingile loogikale. Ja ootuspärane ongi, et need seisundid on täiesti “sõnulseletamatud”: süsteemil puudub ligipääs igasugusele lisainformatsioonile, mis võiks neisse selgust tuua, ja seega ei ole tal võimalik öelda nende seisundite kohta midagi peale osutuse nende omavahelistele sarnasustele ja erinevustele ning mitmesugustele võimalikele assotsiatsioonidele. Päris kindlasti pole mõtet oodata, et ta suudaks selgitada seda “kvaliteeti” mingite elementaarsemate mõistete kaudu.

Sellele saaks vastu vaielda: süsteemi ehitus võiks olla niisugune, et ta saab informatsiooni kätte “aimustena”, umbes nagu pimenägija. Ta võiks öelda näiteks: “Otsustus “punane” tuli mulle lihtsalt pähe”, ütlemata midagi “kvaliteedi” kohta. Kuid asjakorraldus, mis nõuaks, et süsteem jääks aimust ootama, oleks tõenäoliselt ebaefektiivne. Ja kuidas oleks lugu olukordadega, kus inimene peab reageerima visuaalsele informatsioonile otsustusi kujundamata, näiteks tennist mängides? Süsteem ütleks arvatavasti: “Ma lihtsalt leidsin end teadvat, kus pall on, ja tegin, mida vaja, ilma et mul oleks olnud selle kohta mingeid elamusi”? Võib-olla ei olegi niisuguse stsenaariumi loogikal midagi viga, kuid säärane ehitusplaan ei tundu kognitiivse süsteemi jaoks loomulikuna. Palju loomulikum oleks projekteerida süsteem niisuguseks, et ta lihtsalt “näeks” punase ja rohelise erinevust kui niisugust, võtaks tajutud erinevused otseselt oma käitumise aluseks ja annaks küsimustele kindlad ja sirgjoonelised vastused. Vähemasti on see üks mõistlik viis sellise süsteemi kujundamiseks, ja sellest meile praegu piisab.

Seesuguse otsese ligipääsu korral infoseisunditele on loomulik oodata, et süsteem kirjeldab oma kognitiivset arusaama oma tajude kohta elamuste ja kvaliteetide keeles. Ja pole üllatav, et see kõik — vahetud teadmised sõnulseletamatutest seisunditest, mis näivad talle tema suhtluses maailmaga nii olulistena, kuid mida on nii raske täpselt kätte näidata — tundub süsteemile tõepoolest täiesti ootuspäraselt üsna kummalisena, umbes samamoodi, nagu meile tundub kummalisena teadvus.

Siit võikski saada alguse üks võimalikke reduktiivseid seletusi meie otsustustele teadvuse kohta: võiksime öelda, et need otsustused moodustuvad sellepärast, et süsteem leiab end olevat mingites inforuumi punktides nii, et tal on otsene ligipääs nendele punktidele, kuid mitte millelegi muule. See otsene teadmine mõjutab süsteemi kui toores “kvaliteet”: ta teab, et seisundid on erinevad, kuid tal pole võimalik selle väljendamiseks öelda midagi enamat kui “see on niisugune ja too on naasugune”​[[355]](#footnote-356). Vahetu ligipääs elementaarsetele erinevustele viibki otsustuseni, et neil omadustel on mingi salapärane elementaarne olemus, et nad pole väljendatavad fundamentaalsemates mõistetes, ning paljud teised otsustused, mida me teadvuselamuse kohta sageli moodustame.

​**[lk.292]**​ Kõiges selles mängib otsustavat rolli informatsioon. Mitmesugused otsustused elementaarsete “kvaliteetide” kohta moodustuvad just sellepärast, et süsteemil on ligipääs ainult infoseisunditele. Süsteem lihtsalt leiab end kord ühest, kord teisest seisundist, hilisematel protsessidel on aga ligipääs ainult nende varasemate seisundite erinevuste struktuurile, ent mitte millelegi konkreetsele. Kogu töö teeb ära eristavate erinevuste süsteem. Just informatsioon ja meie ligipääs informatsioonile annavadki reduktiivse seletuse meie otsustustele teadvuse kohta.

Mõni lugeja panekski siinkohal võib-olla punkti ja teataks, et teadvuse saladus on lahendatud ja seletus leitud. Muidugi ei pea mina seda õigeks: me oleme andnud seletuse vaid mõnedele otsustustele, mis on palju lihtsam ülesanne. Kuid nüüd saame me liikuda seletusliku koherentsuse printsiibi toel edasi ka teadvuseteoorias. Kui kirjeldatud protsesside käigus realiseeruvad infoseisundid kannavad teadvust puudutavate otsustuste moodustamisel põhiraskust, siis võib on neil olla mingi osa ka teadvuses endas.

Tegelikult just sedakaudu ma algselt jõudsingi informatsioonilise vaateni teadvusele. Kui meie fenomenaalsete otsustuste seletuslik alus peitub eristavate erinevuste struktuuris, siis on loomulik otsida sealtsamast ka teadvuse seletuslikku alust. See seletaks, miks sobivad meie otsustused nii hästi kokku aktuaalsete teadvusseisunditega. Teadvuselamus on mingi infoseisundi realiseering; fenomenaalne otsustus on seletatav sama infoseisundi teise realiseeringuga. Ja mingis mõttes tarvitsebki nende otsustuste õigsuse tagamiseks vaid postuleerida, et informatsioonil on ka fenomenaalne aspekt: sellel[[356]](#footnote-357) informatsioonil on tõepoolest kvalitatiivne aspekt, mis ei avaldu ainult otsustuste süsteemis, vaid ka otseselt fenomenoloogias. See võimaldab siduda teadvust väga kenasti kognitiivse struktuuriga ja saada niiviisi vaimust terviklikum pilt.

Me võime täheldada ka infoseisundite kognitiivse rolli kena ühilduvust elamuse epistemoloogiaga. Elamustele, millega me oleme otseselt tuttavad, vastavad füüsiliselt realiseeruvad infoseisundid, millele süsteemil on otsene (kognitiivne) ligipääs. Süsteem kujundab oma fenomenaalseid otsustusi otsese ligipääsu varal infoseisunditele; see põhjuslik seos haakub kenasti väitega, et elamus — sama infoseisundi fenomenaalne realiseering — annabki põhjenduse moodustuvatele fenomenaalsetele uskumustele. Mõlemal poolel mängib otsustavat rolli üks ja sama infoseisund — ühel juhul selle seisundi füüsiline, teisel tema fenomenaalne realiseering.

See kõik ei tõesta veel lõplikult, et informatsioonipõhine vaade teadvusele on õige, kuid annab sellisele arvamusele tuge juurde.

### ​**[lk.293]**​ 4. Kas elamust võib leida kõikjalt?

Küllap on nüüdseks juba moodustunud järjekord lugejatest, kes tahavad vaielda vastu väitele, et informatsiooni leidub igal pool. Me leiame informatsiooni eest mujalgi kui süsteemides, mida me oleme harjunud teadvuslikeks pidama. Minu CD-mängija on informatsiooni realiseering; minu auto mootor on informatsiooni realiseering; isegi termostaat on informatsiooni realiseering. Tegelikult võib informatsiooni, nagu mina olen seda mõistet siin tõlgendanud, leida kõikjalt, kust võib leida põhjuslikkust. Me leiame põhjuslikkust kõikjalt, seega leiame kõikjalt ka informatsiooni. Aga elamust me ju ei leia kõikjalt?

Infopõhise vaateviisi pooldajal on sellises olukorras kaks võimalust. Esimene ja ilmsem võimalus on oletada, et elamus ei kaasne mitte suvalise füüsiliselt realiseeruva inforuumiga, vaid ainult niisugustega, millel on mingid kindlad omadused, ning asuda otsima edasisi kitsendusi, et eraldada kõigist informatsiooni liikidest välja need, mis on elamusele relevantsed. See nõuaks hoolikat kaalumist, milles need lisakitsendused seisneda võiksid ja kuidas nad sobituksid fundamentaalseadustesse. Võtan taolised strateegiad vaatluse alla [allpool](#PARA.159-5), kuid kõigepealt tahaksin kõnelda teisest võimalusest: võtta süda rindu ja tunnistada, et elamus kaasneb igasuguse informatsiooniga. See tähendaks, et mitte ainult informatsiooni ei leidu kõikjal, vaid ka elamust.

Kui see oletus on õige, siis on elamus omane isegi väga lihtsatele süsteemidele. Seda ideed peetakse sageli veidraks või isegi hullumeelseks. Kuid mina arvan, et ta väärib põhjalikku uurimist. Minu jaoks ei ole see idee ilmselt ekslik, ja tal on omamoodi võlugi. Niisiis asun nüüd vaatlema põhjusi, mille alusel võidaks see vaade kõrvale heita, ning uurima, kui veenvad nad on, ning ühtlasi kaaluma mitmesuguseid positiivseid põhjusi, miks seda vaadet tasuks tõsiselt võtta.

#### Mis tunne on olla termostaat?

Et pilti fookusesse tuua, vaadelgem infotöötlussüsteemi, mis on peaaegu äärmuslikult lihtne — termostaati. Kui vaadelda termostaati infotöötlusseadmena, on tal vaid kolm infoseisundit (üks seisund paneb ta jahutama, teine soojendama, kolmas mitte midagi tegema). Ja meie väide on, et igale neist infoseisunditest vastab mingi fenomenaalne seisund. Need kolm fenomenaalset seisundit on kõik erinevad, ja infoseisundi muutmine muudab fenomenaalset seisundit. Võiksime küsida: milline on nende fenomenaalsete seisundite iseloom? Või teisiti: mis tunne on olla termostaat?

Olla termostaat ei ole vististi kuigi huvitav  — infotöötlus on nii lihtne, et ka vastavad fenomenaalsed seisundid on arvatavasti väga primitiivsed. Termostaadil on kolm elementaarselt erinevat, ilma igasuguse sisestruktuurita fenomenaalset​**[lk.294]**​ seisundit. Võiksime ehk kujutleda neid seisundeid analoogiliselt meie mustade, valgete ja hallide elamustega: termostaadi fenomenaalne väli võib olla üleni must, üleni valge või üleni hall. Kuid isegi nii oleme omistanud termostaadi elamustele ülemäära keeruka struktuuri, kuna visuaalsele väljale on omane dimensionaalsus, ja pealegi on nii mustal, valgel kui hallil värvusel omaenda suhteliselt sisukas iseloom. Tegelikult peaksime ootama midagi veel palju lihtsamat, millel meie elamuses analoogi polegi. Tõenäoliselt ei suuda me neid elamusi empaatiliselt ette kujutada paremini kui pimedana sündinu nägemist või inimene tunnet olla nahkhiir; kuid me võime vähemasti intellektuaalselt teada midagi nende alusstruktuurist.

Et sellele ideele vähem hullumeelset ilmet anda, mõtiskleme, mis võiks elamusega juhtuda keerukuse skaalal allapoole liikudes. Alustame tuttavatest juhtumitest inimesega, kelle väga keerukast infotöötlusest sünnivad meile tuttavad keerukad elamused. Kui liikuda veidi vähem keerukate süsteemide poole, siis ei ole vist kuigivõrd põhjust kahelda selles, et koertel või isegi hiirtel on teadvus. Mõned autorid on selles kahelnud, kuid ma arvan, et sageli on põhjuseks fenomenaalse teadvuse segiajamine eneseteadvusega. Hiirtel ei pruugi olla peaaegu üldse minatunnet ja ka mitte introspektsioonivõimet, kuid tundub olevat täiesti usutav, et on olemas mingisugunegi tunne olla hiir. Hiired tajuvad oma keskkonda infovoo mustritena, mis ei erine kuigivõrd nendest, mida töötleb ka meie aju, kuigi nendel on mustrid tunduvalt vähem keerukad. Loomulik hüpotees on, et “pertseptuaalsele muutkonnale”, mida me teame hiirtel olevat, vastab “fenomenaalne muutkond”. Hiire pertseptuaalne muutkond on üpris mitmekesine — hiir suudab teha oma tajudes palju eristusi —, ja seega võib mitmekesine olla ka tema fenomenaalne muutkond. Näiteks võib uskuda, et igale eristusele, mida hiire nägemissüsteem suudab teha ja keskkonna tajumisel kasutada, vastab mingi fenomenaalne eristus. Tõestada seda ei saa, kuid tundub, et just niisugune on kõige loomulikum ettekujutus hiire fenomenoloogiast.

Liikudes skaalal allapoole üle sisalike ja kalade nälkjateni näeme, et samasugused mõttekäigud on võimalikud ka siin. Ei ole vist kuigi palju alust arvata, et fenomenoloogia peaks jada ühes punktis korraga kustuma, samas kui tajupsühholoogia on ikka veel üpris keerukas. Kui see peaks siiski nii olema, siis on kusagil kas radikaalne katkestus, kus keerukad elamused asenduvad ühel hoobil elamuste täieliku puudumisega, või siis hakkab kaduma fenomenoloogia sünkroonsus tajudega, nii et mõnda aega püsib suhteliselt rikkalik pertseptuaalne muutkond kõrvuti tunduvalt kesisema fenomenaalse muutkonnaga. Esimene hüpotees on väheusutav, teine aga eeldaks, et vahepealsetel süsteemidel on siseelu kummaliselt irdus kognitiivsetest võimetest. Alternatiiv on kindlasti vähemalt sama tõepärane. Olla kala on arvatavasti palju vähem huvitav kui ​**[lk.295]**​ olla inimene, kuna lihtsamale psühholoogiale vastab lihtsam fenomenoloogia. Kuid on vist küllalt arukas oletada, et mingisugunegi tunne olla kala on ikkagi olemas.

Kui liikuda skaalal kaladest ja nälkjatest edasi, üle lihtsate neuronvõrkude termostaatideni, siis missuguses punktis teadvus kaob? Kalade ja nälkjate fenomenoloogia ei ole tõenäoliselt primitiivne, vaid suhteliselt keerukas, kajastamaks erinevaid eristusi, mida nad teha suudavad. Küllap peaks fenomenoloogia täielikule kadumisele eelnema mingi äärmuslikult lihtne fenomenoloogia. Mulle tundub, et kõige loomulikumaks kohaks, kus see võiks toimuda, on samavõrra lihtsa “tajupsühholoogiaga” süsteemid, nagu seda on näiteks termostaat. Termostaadis tundub sedaliiki infotöötlus, mis on olemas kalal või nälkjal, realiseeruvat oma kõige lihtsamal kujul, nii et võib-olla on tal siis ka vastavat laadi fenomenoloogia oma kõige lahjemal kujul — niisugune, mis saab hakkama ühe või kahe tema toimimise jaoks olulise eristusega; vähemalt mulle ei tundu alusetuna oletus, et vastavad eristused on olemas ka elamuses.

Kogu see jutt ei olnud muidugi mingi tõestus, et ka termostaatidel peavad tingimata elamused olema, sest muutused võivad keerukuse skaalal allapoole liikudes olla ka teistsugused. Aga toodud pilt tundub olevat üks mõistlikke võimalusi, ja kui järele mõelda, siis isegi mitte vähem loomulik kui teised. Võib arvata, et siinne arutluskäik on lihtsalt sama arutluskäigu jätk, millega me omistame elamusi koertele ja hiirtele. Igatahes niipea kui me hakkame mõtlema, mis võiks toimuda hiire elamuses, ja nende protsesside alustest tema tajupsühholoogias, hakkab sama mõttekäigu rakendamine järjest lihtsamatele süsteemidele tunduma palju loomulikumana kui võib-olla esialgu.

Kui keegi peab mõtet termostaadi elamustest hullumeelseks, siis võlgneb ta meile vähemalt seletuse, miks see mõte hullumeelne on. Arvatavasti põhjendaks ta seda sellega, et termostaadil puudub mingi elamusteks ilmselgelt tarvilik omadus; minule aga ei tundu ükski niisugune omadus ilmselgena. Võib-olla puudub termostaadil mingi oluline infotöötluse komponent, mis on hiirel olemas või mis puudub hiirel ja on olemas inimesel, kuid mina säärast elamuseks nii väga ilmselgelt tarvilikku komponenti ei näe, ja pole ka ilmne, et niisugune komponent peab tingimata olemas olema.

Kui me ütleme, et termostaatidel on elamused, siis ei tähenda see muidugi, nagu oleks neil midagi tõelise vaimse elu taolist. Termostaadil puudub eneseteadvus, tal ei ole vähimalgi määral intelligentsust, ja ma ei tahaks väita, et termostaat võib mõtelda.[[357]](#footnote-358) Mingis osas võib vastuseis teadvusega termostaadi ideele tulla elamuse seostamisest viimati mainitud vaimuomadustega, mis kõik nõuavad peaaegu kindlasti palju suuremat keerukust. Kõigil neil omadustel on tugev psühholoogiline komponent, ja vastavate põhjuslike rollide tagamiseks peaks süsteem arvatavasti üpris keerukas olema. Ent niipea kui teha fenomenaalsete omaduste ja psühholoogiliste omaduste vahele selge vahe, pole teadvusega termostaadi idee enam nii hirmus. Me peame ette kujutama vaid ​**[lk.296]**​ midagi elementaarse “elamussähvatuse” taolist, millega ei kaasne mingeid mõisteid, ühtki mõtet ega mingit keerukat infotöötlust.

Teiseks võib mõni lugeja olla teadvusega termostaadi idee vastu sellepärast, et ta ei leia süsteemis teadvusele kohta. Süsteem tundub olevat liiga lihtne ja tundub, et teadvusele ei leiduks temas rolli. Ent see, kes nii mõtleb, ei ole omandanud mittereduktiivse positsiooni õppetundi: hoolikalt uurideski ei leia me ühestki süsteemist mitte kunagi teadvust, ja me võime toimuvaid protsesse alati mõista ilma teadvuse mõiste abita. Kui teadvus ei ole loogiliselt superventne, siis ei olegi oodata, et me leiame süsteemi organisatsioonis teadvusele “koha” — teadvus on hoopis midagi muud kui süsteemi infotöötluslikud omadused.

Mõni autor pole aga valmis tunnistama teadvusega termostaatide võimalikkust lihtsalt sellepärast, et me tunneme termostaate liiga hästi. Me teame neis toimuvatest protsessidest kõike, ja teadvusest kõnelemiseks ei paista mingit põhjust. Kuid tegelikult ei erine termostaat selles suhtes ajust. Ka siis, kui me hakkamegi ajuprotsesse kunagi täielikult mõistma, ei leia me ikkagi mingit põhjust kõnelda teadvusest. Vahe termostaadiga on ainult selles, et see, mis ajus toimub, on meile praegu veel nii salapärane, et võib kergesti tulla kiusatus oletada, et teadvuse “asupaik” on mingil moel ajuprotsessides, mida me veel ei mõista. Kuid nagu ma olen püüdnud näidata, ei too ka nende protsesside mõistmine iseenesest veel pildile teadvust. Seega on aju ka selles suhtes termostaadiga samas olukorras.

Mõnda häirib võib-olla see, et inimene võib termostaadi ise ehitada, panemata temasse mingit teadvust. Kuid sama kehtib ju ka aju kohta, vähemalt põhimõtteliselt. Kui me (näiteks paljunemise ja arengu käigus) loome aju, tuleb teadvus tasuta kaasa. Nii on ka termostaadiga: ei maksa oodata, et teadvust saaks süsteemis lokaliseerida nagu füüsilist komponenti! Mõnda jällegi teeb rahutuks võib-olla see, et termostaat ei ole elus; kuid on raske näha, miks peaks sel erinevusel nii põhimõtteline tähtsus olema. Eelmises peatükis kirjeldatud kehatut räniaju ei saaks vist samuti elusaks pidada, kuid teadvus võib tal, nagu nägime, ikkagi olla. Ja kui eelmise peatüki argumendid on õiged, siis ei ole bioloogiliste komponentide puudumisel termostaadi koosseisus põhimõtteliselt mingit tähtsust.

Intuitiivne vastuseis võib mingis osas tulla sellest, et termostaadis ei tundu olevat kohta sellele, millele või kellele elamused kuuluda saaksid: kus on termostaadis koht, kuhu mahuks subjekt? Kuid me ei peaks otsima füüsilistes süsteemides subjekti koha täiteks mingit homunkulust. Subjektiks on kogu süsteem, või õigemini kuulub subjekt süsteemi juurde samamoodi, nagu subjekt kuulub aju juurde. Sellest kõnelemiseks pole lihtne õigeid sõnu leida. Me ei ütleks rangelt võttes, et elamused on minu ajul, vaid et nad on minul. Ükskõik kuidas me seda seost ka mõtestaksime, ikka saab seda rakendada ka termostaatidele: rangelt võttes on ilmselt parem öelda mitte, et ​**[lk.297]**​ elamused on termostaadil endal (kuigi ma ütlen mitterangelt kõneldes niiviisi ka edaspidi), vaid et elamused kuuluvad termostaadi juurde. Me ei leia termostaadi “seest” subjekti, ent sama vähe leiame me subjekti ka aju seest.

Tulles tagasi positiivsete punktide juurde, mis kõnelevad lihtsate elamustega süsteemide kasuks: need lubavad meil pääseda mööda vajadusest postuleerida teadvuse hüppeline “süttimine” mingil keerukusastmel. Mõttes, et *n*​-elemendilisel süsteemil ei saa teadvust olla, kuid ​*n*​ + 1-elemendilisel süsteemil juba saab, oleks midagi kentsakat. Ja me ei saa otsuse tegemisest mööda hiilida nii, nagu me võime hiilida mööda küsimusest, kui väikesest juuste arvust alates on inimene kiilaspäine: viimasel juhul võib kõnelda mõningasest semantilisest määramatusest, kuid palju vähem veenvana tunduks väide, et teadvuse olemasolu mingil süsteemil võib samuti määramatu olla. (Seda eriti juhul, kui võtta omaks mittereduktiivne vaade, mille järgi elamuse fakte ei saa kirjeldada elementaarsete faktide kaudu, nii nagu me seletame määramatust kiilaspäisuse küsimuses kindlakujuliste faktide kaudu juuste arvu kohta peas.). Kuigi oleks ju mõeldav, et teadvus ilmubki mingis punktis hüppeliselt, näib iga konkreetne punkt selleks meelevaldsena, mistõttu teooria, mis oskaks sellise otsuse tegemist mööda pääseda, võidaks mõneti lihtsuses.

Ja lõpuks kõneleb elamuste kasuks lihtsatel süsteemidel veel üks kaalutlus: kui elamus on tõepoolest fundamentaalne omadus, siis peaks tema laialdane levik olema täiesti loomulik. Igatahes avalduvad kõik teised fundamentaalsed omadused, mida me tunneme, isegi lihtsates süsteemides, ja neid leidub kogu universumis. Oleks kummaline, kui mingi fundamentaalne omadus realiseeruks esmakordselt alles universumi suhteliselt hilises ajaloos ja isegi siis ainult väga vähestes keerukates süsteemides. Mõttes, et mõni fundamentaalne omadus realiseerub vaid sporaadiliselt, pole otsest vastuolu; ent muudel võrdsetel tingimustel tundub vastupidine võimalus usutavamana.

#### Panpsühhism — kuhumaani?

Kui elamus kuulub juba termostaatidegi juurde, siis võib arvata, et teda leidub kõikjal, sest kõikjal, kus on põhjuslik vastasmõju, on ka informatsioon, ja kus on informatsioon, seal on ka elamus. Infoseisundeid võib leida kivis — näiteks kui ta paisub ja kokku tõmbub —, ja isegi elektroni seisundites. Nii et kui topeltaspektiprintsiip kehtib kitsendusteta, siis kuuluvad elamused ka kivide ja elektronide juurde.

(Ma ei ütleks lausa, et kivil on elamused või et tal on teadvus, analoogiliselt sellega, kuidas ma mitterangelt kõneldes ütleksin, et termostaadil on elamused või et tal on teadvus. Erinevalt termostaadist ei vaatle me kivi infotöötlussüsteemina. Me vaatleme teda lihtsalt kui objekti, mistõttu seos elamusega on vähem otsene. Võib-olla oleks parem öelda, et kivi sisaldab teadvusega süsteeme: arvatavasti on tal palju alamsüsteeme, ​**[lk.298]**​ millest ühegi elamusi ei saa kanooniliselt pidada kivi elamusteks (sama vähe, nagu minu elamusi saab pidada minu töökabineti elamusteks). Seevastu termostaadi juurde kuulub teatav kanooniline inforuum, mistõttu kanoonilistest elamustest rääkimine tundub termostaadi puhul mõistlikumana. Muidugi ei ole seegi kõnepruuk, nagu [ülal](#id_157loose) märgitud, päris range.)

Vaade, mille järgi elamust leidub kõikjal, kus on põhjuslik vastastoime, on intuitsioonivastane. Ent järele mõeldes hakkab see idee pikkamööda järjest ahvatlevamana tunduma, kuna ta integreerib teadvuse paremini looduse korrasse. Kui see vaade on õige, siis ei leidu teadvust maailmas sporaadiliste pudemetena — külluslike teadvuselamuste kujul, mida meelevaldselt produtseerivad üksteisest lahus seisvad keerukad süsteemid —, vaid tema jaotus universumis on ühetaolisem: väga lihtsatel süsteemidel on ka fenomenoloogia väga lihtne, keerukatel süsteemidel keerukas. Nii hakkame nägema teadvust mingis mõttes vähem “erilisena” ja seega mõistuspärasemalt.

Huvi pakub küsimus, kas elamus eeldab ka aktiivset põhjuslikkust. Kas termostaadil võib olla elamusi ka siis, kui tema seisund ei muutu (mis on mingis mõttes samuti mingi välise olukorra “põhjus”, ilma et termostaat aga reaalselt midagi teeks)? Või on tal elamus ainult seisundite vahetumise ajal? Aju puhul näib suurem osa elamust kandvatest põhjuslikest protsessidest olevat aktiivne, kuna asjakohase informatsiooni töötlus toimub pidevalt, neuronid saadavad kogu aeg impulsse jne. Teisest küljest võib olla, et aktiivse ja passiivse põhjuslikkuse vahel ei olegi fundamentaalsel tasemel erinevust, ja sellisel juhul võiks neid kaht käsitleda võrdväärsetena. Ma ei oska sellele küsimusele vastata, kuid intuitiivselt tundub, et elamus peaks eeldama siiski mingisugust aktiivsust.

Ei saa ka välistada veel üht siin seni veel kaalumata võimalust: et lihtsatel süsteemidel ei ole fenomenaalseid omadusi, kuid on *protofenomenaalsed* omadused. Ma mainisin [4. peatükis](#PARA.82-10) võimalust, et fenomenaalseid omadusi konstitueerivad mingid veel fundamentaalsemad omadused. Kui sellised omadused on tõepoolest olemas, siis tundub mõte nende realiseerumisest lihtsates süsteemides loomulikuna. Sellisel juhul ei ole termostaatidel võib-olla elamusi selle sõna tavalises mõttes, vaid nad realiseerivad mingeid hoopis teisi, elamustega sarnanevaid omadusi, mida me täielikult ei mõista (vahest mingeid protoelamusi). See lubaks jääda ülalkirjeldatud ühtse vaade juurde looduse korrale ja aitaks võib-olla kaasa ka “hüppelise kadumise” probleemi lahendamisele. (Kui protofenomenaalsed omadused ongi fundamentaalsed, siis elamused, mida nad konstitueerivad, võivad ikkagi “lülituda sisse” vähehaaval.) Kuna see seisukoht ei ütle, et termostaatidel on täisväärtuslikud elamused, võib ta tunduda ka veidi vähem “pöörasena” kui alternatiiv. Muidugi on selle hinnaks meile tundmatute ja arusaamatute omaduste klassi postuleerimine. Kuid välistada seda võimalust ei saa.

Igal juhul on säärasel vaatel palju ühist nn. *panpsühhismiga* — seisukohaga, mille järgi psüühika on kõigel. Ma ise seda terminit üldiselt ei kasuta, ja seda mitmel põhjusel: (1) ma arvan, et ​**[lk.299]**​ elamustest üksi ei piisa veel selleks, mida me tavaliselt mõtleme psüühika all, kuigi neid võib pidada psüühikaks oma kõige lihtsamal kujul; (2) võib selguda, et protofenomenaalsed omadused jäävad tavapärasest psüühika mõistest veelgi kaugemale; (3) minu arvates ei ole [ülalöeldud](#PARA.158-2) põhjustel rangelt võttes õige öelda, et näiteks kividel on elamused, kuigi mingid elamused võivad nende juurde kuuluda. Termini “panpsühhism” eksitavuse peamine põhjus on aga võib-olla hoopis selles, et ta eeldab seisukohta, mille järgi lihtsate süsteemide, näiteks aatomite elamused on fundamentaalsed ning keerulised elamused on mingil moel selliste lihtsamate elamuste summeerumise tulemus. Kuigi see on üks võimalus, ei ole mingit põhjust arvata, et asi peaks olema just nii: keerukad elamused võivad olla autonoomsemad, kui säärane seisukoht eeldaks. Muu seas eeldab informatsiooniline arusaam teadvusest ettekujutust, mille järgi keerukate elamuste moodustumine toimub holistlikumalt.

Neid mööndusi arvesse võttes pole vist vale öelda, et siinkirjeldatud vaade on üks panpsühhismi teisendeid. Pean siiski märkima, et minu kontseptsiooni metafüüsiliseks aluseks ei ole panpsühhism, vaid naturalistlik dualism koos psühhofüüsikaseadustega. Panpsühhism on lihtsalt üks viis, kuidas elamuse looduslik superventsus füüsika suhtes võiks välja näha. Looduslik superventsus annab mingis mõttes üldise skeleti; pansühhism on vaid üks võimalus üksikasjade väljatöötamiseks.

Mina isiklikult pooldan palju ennemini naturalistlikku dualismi kui panpsühhismi. Viimases tundub olevat väga palju lahtisi otsi. Kuid ma loodan, et sellest, mida ma öelnud olen, piisab näitamaks, et me peaksime mingikujulise pansühhismi võimalikkust tõsiselt võtma: selle vaate vastu ei paista olevat otsustavaid argumente, ja samas leidub mitmesuguseid positiivseid põhjendusi tema kasuks.

#### Topeltaspektiprintsiibi kitsendamine

Isegi kui olla valmis tunnistama elamusi ka väga lihtsatel süsteemidel, võib mõte, et informatsiooniga käib alati kaasas elamus, siiski nõutuks teha. Näiteks tundub teadvuselamuses sisalduv informatsioon olevat vastavuses ainult väikese osaga kogu informatsioonist, mida inimene kognitiivselt töötleb. Eks ole ju ilmselge fakt, et suurem osa meie infotöötlusest toimub teadvuse osavõtuta?

Kui topeltaspektiprintsiip kehtib kitsendusteta, siis on vastus oletatavasti selles, et kogu see “mitteteadlik” informatsioon realiseerub elamuses küll, kuid mitte tingimata just minu elamuses. Näiteks kui ühel minu neuronil on mingi elamus (nagu termostaadilgi on mingi elamus) siis me ei eeldaks, et see elamus moodustaks osa minu enda elamusest, samuti nagu me ei eeldaks ka, et minu elamust muudaks radikaalselt selle neuroni väljavahetamine väikese homunkulusega, kellel oleks teadvs. Samamoodi võivad elamused olla aju mitmesugustel ​**[lk.300]**​ infotöötlusega tegelevatel “teadvuseta” alamsüsteemidel — vastavad elamused kuuluvad lihtsalt teisele subjektile. Ajus on palju mitmesuguseid infotöötlussüsteeme ja süsteem, mis vastab minule — võib-olla süsteem, mis teeb mingi informatsiooni kättesaadavaks mingit kindlat liiki globaalseks juhtimiseks ja raporteerimiseks — on vaid üks neist. Ma ei oota, et mul oleks otsene ligipääs nende teiste süsteemide elamustele, nagu ma ei oota ka, et mul oleks otsene ligipääs teiste inimeste elamustele.

Rahutuks võivad teha ka suhteliselt keerukad infotöötlussüsteemid, mida võib leida maailmas kõikjal, alates minu CD-mängijast ja lõpetades minu maoga. Kas neid kõiki võib pidada keerukate elamustega, teadvusega indiviidideks? Vastuseks tasub märkida, et neis süsteemides ei ole midagi, mis vähegi sarnaneks sidusa kognitiivse struktuuriga meie enda süsteemis, ning seega ei ole kõik nende juurde kuuluvad elamused tõenäoliselt sugugi sarnased meie omadega. Kui näiteks CD-mängija juurde kuuluvad mingid elamused, siis taanduvad nad tõenäoliselt vaid “lamedale” bittstruktuurile; ja kui mingid elamused seostuvad informatsiooniga minu maos, siis pole alust arvata, et nad vastaksid millelegi, mida me mõtleme psüühika all. Meie elamustega samasse liiki kuuluvad elamused tekivad ainult infotöötlussüsteemides, millel on evolutsiooni käigus kujunenud keerukad, koherentsed kognitiivsed struktuurid, mis külluslikult kajastavad meie ettekujutust välisest maailmast. On tõenäoline, et elamussubjektide klass, millede psühholoogiline struktuur lubaks pidada neid tõepoolest agentideks või isikuteks, on väga kitsas.

Ometi võib see elamuste lokkamine, eriti kui see toimub ühe aju piires, jätta ebamugava mulje, mida süvendab veelgi tähelepanek, et üldjuhul on kaunis lihtne leida palju inforuume, mis on mingist ühest konkreetsest inforuumist vaid pisut erinevad — selleks tarvitseb vaid mõnd relevantset põhjuslikku trajektoori teisiti individualiseerida või korraldada vastavad tagajärjed (informatsiooni erinevusest sündivad “eristused”) veidi teisiti. Kas me peaksime uskuma, et kõigi nende inforuumide ümber hõljub palju mitmesuguseid elamussüsteeme? Kui nii, siis võib mul olla suur hulk omavahel väga sarnaseid, kuid pisut siiski ka erinevaid fenomenaalseid sugulasi, kes kasvavad välja protsessidest minu enda peas!

Teine võimalus on kitsendada topeltaspektiprintsiipi, et kahandada füüsiliselt realiseeruvate, fenomenaalsete korrelaatidega inforuumide klassi. Kõige loomulikum strateegia oleks ehk kitsenduste seadmine informatsiooni töötlemise viisile. Ma olen ju nagunii juba öelnud, et minu süsteemis vastab minu elamusele kõige otsesemalt see informatsioon, mis on otseselt kättesaadav globaalseks juhtimiseks. Ent pole kuigi tõenäoline, et see “kriteerium” hakkaks niisugusel kujul fundamentaalseaduses mingit rolli mängima — selleks on ta liiga ebamäärane ja kõrgtasemeline; tõepoolest, me saame seda põhimõtet kasutada ju alles siis, kui me oleme juba individualiseerinud mingi kõrgtasemelise süsteemi, näiteks inimese või aju. Kuid võib-olla leidub täpsem ja lihtsam kriteerium, mis annab sama tulemuse.

​**[lk.301]**​ Üks võimalus on, et otsustav tähtsus on informatsiooni võimendamisel. See tähendaks, et füüsiliselt realiseeruv informatsioon realiseerub ühtlasi elamuses ainult juhul, kui süsteem seda informatsiooni mingil kindlal viisil võimendab, nii et see võib põhjustada mõnel põhjuslikul trajektooril suuri erinevusi. Võiks ehk isegi öelda, et elamuse intensiivsus vastab võimenduse määrale, või midagi sarnases vaimus. See oleks vahest kenasti ühitatav globaalse kättesaadavuse kriteeriumiga, kuigi nüüd võib ilmneda teisi raskusi: näiteks on olemas rohkesti võimendatud informatsiooni, mis intuitiivselt ei tundu teadvusse jõudvat; samuti ei ole ilmne, kuidas tuleks täpsustada võimendamise mõistet ennast.

Teine võimalus oleks seada kitsendusi süsteemis realiseeruvale põhjuslikkuse tüübile. Me [nägime](#PARA.156-1), et kus on põhjuslikkus, seal on alati ka informatsioon; kuid võib-olla sobib elamust kandvate inforuumide individualiseerimiseks ainult mingit kindlat liiki põhjuslikkus. Näiteks sobivad selleks võib-olla ainult mingis mõttes “aktiivsed” põhjuslikud suhted, või siis peavad põhjuslikud suhted olema mingis mõttes “​loomulikud”. Intuitsioon ütleb, et paljud inforuumid, mis meie seni pakutud kriteeriume rahuldavad, on mingis mõttes ebaloomulikud; võib-olla saaks vastavat kitsendust kuidagi selgemini formuleerida. Ka nii jääks meile tõenäoliselt ikkagi väga avar infoseisundite klass, kuid astronoomilisele vohamisele tuleks ehk lõpp.

Ma ei oska öelda, missugune peaks olema vajalik kitsendav kriteerium. Kuid see ei tähenda, et seda ei olegi. Võib isegi leiduda kitsendav kriteerium, mis ahendab vastavaid inforuume nii, et termostaatides ja teistes sellistes lihtsates süsteemides liiklev informatsioon jääb mängust hoopis välja. Minu enda intuitsioon ütleb, et vajalikku topeltaspektiprintsiibi kitsendust on arvatavasti võimalik leida, kuid see ei tarvitse olla niisugune, mis välistaks nii lihtsad süsteemid, nagu seda on termostaadid. Minule tundub paljude sarnaste elamuste vohamine ühes ajus intuitsioonivastasemana kui elamuse olemasolu lihtsates süsteemides, kuigi kummagagi ei ole asi täiesti selge. Igatahes võib prototeooria viimistlemine teooriaks hakata kulgema väga mitmel viisil.

### 5. Informatsiooni metafüüsika

Kuid endiselt on vastamata küsimus: kuidas mõista informatsiooni topeltaspektilisuse kontseptsiooni ontoloogiat? Kui tõsiselt võtta neid arutlusi inforuumidest ja -seisunditest: kas nad on lihtsalt kasulikud konstruktid, või on nad mingil moel ontoloogiliselt fundamentaalsed? Kas informatsioon on primaarne, või on primaarsed siiski füüsiline ja fenomenaalne sfäär, informatsioon aga on vaid kasulik ühenduslüli?

Seda kõike võib mõista mitmeti. Kõige otsesem ja kõige vähem avantüristlik võimalus oleks tunnistada informatsiooni füüsikaline realiseering ja fenomenaalne realiseering teineteisest täiesti lahus seisvateks nähtusteks, millede vahel pole ainsatki ontoloogilist ühenduslüli peale seaduspärase seose ja teatava struktuurse isomorfismi. ​**[lk.302]**​ Niiviisi vaadates jääb ontoloogia sirgjooneliseks omadusdualistlikuks ontoloogiaks, mille järgi fenomenaalsed omadused seisavad füüsikalistest omadustest täiesti lahus, kuid nende vahel on seaduspärane seos. Jutte topeltaspektsust tuleb siin mõista deflatiivselt, lihtsalt värvika viisina kõnelda kahesugustest omavahel korreleeruvatest ja sarnase struktuuriga omadustest. Informatsioon on aga lihtsalt kasulik vahend selle ühise struktuuri iseloomustamiseks; mingit ontoloogiliselt “sügavat” vastet tal ei ole.

See võib olla täiesti õige vaateviis, kuid on ka huvitavamaid võimalusi. Enamik neist eeldab tõsisemat suhtumist informatsiooni rollisse. Järgnevalt asun ma vaatlema üht säärast võimalust. Ma hoiatan lugejat, et arutlus sukeldub kaunis sügavale spekulatiivsesse metafüüsikasse, kuid vaevalt olekski võimalik vältida teadvuse ontoloogiat käsitledes spekulatiivset metafüüsikat.

####  Alguses oli bitt

Aeg-ajalt avaldavad füüsikud mõtet, et informatsioonil on universumi füüsikas keskne koht, ja ütlevad mõnikord isegi, et füüsikalised omadused ja seadused ongi võib-olla tuletised informatsioonilistest omadustest ja seadustest. Seda arusaama — “alguses oli bitt” [​*it from bit*​] — on väljendanud Wheeler (1989, 1990) ja Fredkin (1990), ja seda on oma artiklites vaadelnud ka Zurek (1990) ja Matzke (1992, 1994). Kui see on õige, siis on ehk võimalik leida informatsioonile meie ontoloogias tõsisem roll. Et seda paremini ette kujutada, vaatlen üht peamist viisi, kuidas informatsiooni saaks pidada füüsikas fundamentaalseks. See ei ole ainus viis, kuidas mõtet “alguses oli bitt” arendatud on (näiteks on see mõnevõrra erinev Wheeleri positsioonist[[358]](#footnote-359)), kuid just see viis tundub mulle kõige loomulikumana kõnealuse arusaama mõtestamiseks. See tõlgendus põimub, nagu [allpool](#ORG.8.5.2) näeme, tihedalt [4. peatükis (lk. 153-55)[[359]](#footnote-360)](#ORG.4.4.1.4) käsitletud “russelliaanlike ideedega”.

See vaateviis toetub tähelepanekule, et fundamentaalsed füüsikalised seisundid, millest kõnelevad füüsikateooriad, individualiseeruvad faktiliselt infoseisunditena. Vaadeldes mõnd parameetrit, näiteks massi või laengut, leiame eest vaid elementaarsete eristavate erinevuste ruumi. Füüsika ei ütle, mis on mass või mis on laeng: ta annab meile vaid väärtuste hulga, mida need suurused omandada võivad, ning ütleb meile, kuidas nad võivad mõjutada teisi parameetreid. Füüsikateooriate poolest võiksid konkreetsed massi- või laenguseisundid väga hästi olla puhtad infoseisundid: oluline on vaid nende asukoht inforuumis.

See paistab välja asjaolust, et füüsika ei võta endale mingeid kohustusi nende seisundite realiseerumise viisi suhtes. Füüsikateooria eesmärke teenib ühtviisi hästi nende infoseisundite ükskõik missugune realiseering, senikaua kui see ei riku seisundite vaheliste põhjusliku või dünaamiliste suhete struktuuri. Senikaua kui nende suhete kuju ei muutu, paistab füüsiline maailm meie tajumissüsteemidele ikka ühesugusena: meil ei ole välises maailmas ligipääsu mingitele teistele realiseeringu omadustele ​**[lk.303]**​ peale põhjuslike suhete süsteemi struktuuri kuju (vähemalt kuivõrd meie fenomenaalsed omadused ei ole otseses seoses realiseerivate omadustega.)

On väljendatud isegi mõtet, et universum ongi hiiglaslik arvuti. Fredkin (1990) on oletanud, et universum on tohutu rakkautomaat, mille kõige madalatasemelisemaks realiseeringuks on üüratu bittstruktuur.[[360]](#footnote-361) Leckey (1993) on aga öelnud, et võib-olla ongi kogu aegruumi aluseks informaatiline protsess omaette registriga iga maailmas realiseeruva fundamentaalsele parameetri kohta. Senikaua kui nende registrite vahel valitsevad õiged põhjuslikud suhted, ei märka ükski olend midagi. See informaatiline mõttekäik illustreerib suurt hulka võimalikke viise, kuidas füüsikalised entiteedid, mida me “tunneme”, realiseeruda võivad: selleks on tarvis vaid, et mingid entiteedid mängiksid sobivaid põhjuslikke rolle. Kõnealuse oletuse võib liigitada n.-ö. füüsika metafüüsika valda kuuluvaks spekulatsiooniks aegruumi enda põhjuslikku struktuuri kandva ontoloogia üle.

Füüsika ise säärase metafüüsikaga muidugi ei tegele. Küsimustes, kuidas parameetrid realiseeruvad ja kas nad realiseeruvad just nii- või naasugusel viisil, võib füüsika jääda täiesti neutraalseks. Füüsika poolest võiks maailma olemus ammenduda maailma informatsioonilise kirjeldusega. Kui ongi olemas mingeid veelgi fundamentaalsemaid “realiseerivaid” omadusi, siis ei mängi need füüsikalistes teooriates otseselt mingit rolli, ja nii võib tulla ahvatlus neist üldse loobuda.

Nii saaksime pildi maailmast kui puhta informatsiooni maailmast. Igale maailma fundamentaalsele parameetrile vastab üks inforuum ning kõikjal, kus füüsika näeb nende parameetrite realiseeringut, realiseerub mingi infoseisund vastavast ruumist. Senikaua kui nende infoseisundite omavahelised suhted on õiged, on kõik nii, nagu olema peab. Selle maailmapildi kohaselt pole siin rohkem midagi öelda. Ei ole olemas midagi peale informatsiooni.

Nii saangi ma aru kontseptsioonist, mille järgi maailma “alguses oli bitt”. Selles kontseptsioonis — maailm kui puhtas infovoog, ühegi muu substantsita informatsiooni kõrval — on kummalist ilu. (Mõni selle vaate versioon lubab vaadelda aegruumi kui elementaarset karkassi, millele need inforuumid kinnituvad; teiste versioonide järgi konstitueerivad inforuumide vahelised suhted ka aegruumi ennast.) Maailm koosnebki elementaarsetest erinevustest ühes põhjuslike ja dünaamiliste suhetega erinevuste vahel. Selle kontseptsiooni järgi oleks eksitus püüda öelda maailma kohta veel midagi enamat.

####   ​Informatsiooni fenomenoloogilised juured

Sellise maailmapildiga on nähtavasti kaks peamist raskust. Esimese seab meie ette teadvus ise. Teadvuses paistab olevat veel midagi ​**[lk.304]**​, mis ei mahu puhta inforuumi piiridesse. Fenomenaalsetel omadustel on seesmine olemus, mis ei ammendu asukohaga inforuumis. Tundub, et puhtalt informatsioonilises maailmapildis ei leidu nendele seesmistele omadustele kohta.

Teine raskus seisneb selles, et puhta infovoo mõiste koherentsus pole ilmne. Võib tunduda, et kõnealuse vaate kohaselt on maailm liiga sisutühi, et olla maailm. Kas võib olla elementaarseid erinevusi, mis ei tulene erinevustest mingis nende aluseks olevas kvaliteedis​[[361]](#footnote-362)? Võib pidada tõepäraseks, et igal konkreetsel erinevusel maailmas on oma kandja, s.t. peab leiduma miski, mis erineb.

See raskus on väga sarnane [4. peatükis](#PARA.96-2) (lk. ​153​[[362]](#footnote-363)) käsitletud “puhta põhjuslikkusvoo” kontseptsiooni raskustega, mille üks teisendeid see vaade on. Too kontseptsioon kaotas maailmast kõik olemuslikud omadused, jättes alles vaid põhjuslike suhete maailma, kus ei paista põhjusliku töö tegijat. Siinne kontseptsioon näeb ehk veidi parem välja, kuna selle järgi on olemas infoseisundid, mida põhjuslikud suhted omavahel seostavad. Kuid need seisundid on silmatorkavalt sisutühjad: nad on üksteisest lihtsalt erinevad, kuid neil puudub igasugune omaenda olemus. Lugeja leiab võib-olla, et säärane pilt seesmise olemuseta maailmast ei kujutagi maailma.

Tõepoolest, mulle võidakse vastu öelda, et inforuumidel peab olema ka mingi omaenda olemus. Kahel fundamentaalsel omadusel võib olla ühesugune informatsiooniline struktuur — näiteks võivad mõlemad omandada reaalarvulisi väärtusi mingil kontiinumil. Kui füüsika on puhas informatsioon, siis pole mingit võimalust nende kahe inforuumi realiseeringuid teineteisest eristada. Kuid mingisugune erinevus peab nende vahel olema, sest need kaks omadust figureerivad üsna erinevates seadustes ja avaldavad maailma teistele omadustele erisugust mõju. Seega peab olema veel midagi, mis neid realiseeringuid eristab; midagi, mis ulatub puhta informatsiooni piiridest väljapoole. Selleks eristuseks on nähtavasti tarvis mingit seesmist omadust.

Nende raskustega võib püüda toime tulla mitmel viisil. Võib otsustada, et teine raskus ei ole siiski fataalne, ja jääda rahule puhta informatsiooni füüsikaga; edasi võiks püüda kaasata sellesse füüsikasse fenomenaalsed omadused, otsides neile mingeid seaduspäraseid seoseid informatsiooniga. Teise raskusega võiks püüda toime tulla ka mingite olemuslike omaduste postuleerimisega, mis oleksid füüsikaliste inforuumide aluseks, esimese raskuse võitmiseks aga tuua fenomenaalsed omadused mängu mingitel sõltumatutel kaalutlustel.

Kõige põnevam strateegia on siiski püüda saada mõlemast raskusest üle üheskoos. Esimene raskus viib mõttele, et fenomenaalsed omadused annavad meile otsese teadmise maailma seesmisest olemusest, mis ei mahu puhta informatsiooni piiridesse, teine aga mõttele, et maailmal peab olema seesmine olemus, mis neid infoseisundeid kannab. Sellisel juhul võib selguda, et infoseisundite kandjana tarvilik seesmine olemus kuulub tihedalt kokku meile fenomenoloogias avaneva seesmise olemusega. Võib isegi olla, et üks neist konstitueerib mingil määral ​**[lk.305]**​ teist. Nii saame säästliku ja elegantse ontoloogia ning ületame kaks raskust ühe hoobiga.

See mõte on jällegi lähedases suguluses Russelli ideega, millest oli juttu  [4. peatükis](#PARA.96-3) — et maailma tundmatud olemuslikud omadused on fenomenaalsed (või protofenomenaalsed). Russell vajas neid omadusi, et rajada neile põhjuslikud suhted, millest füüsika kõneleb, meie aga vajame neid selleks, et rajada neile infoseisundid (eristavad erinevused), mida füüsika postuleerib. Probleem on sisuliselt sama. Mõlemal juhul tundub, et me saame kaks lahendust ühe lahenduse hinnaga. Füüsikalise maailma mõtestamiseks on tarvis mingeid seesmisi omadusi, teiselt poolt aga tuleb leida koht seesmistele omadustele, mida ilmutab meile fenomenoloogia. Tundub, et need kaks probleemi hakkavad hästi kokku.

Seega on idee selles, et inforuumid, mida nõuab füüsika, juurduvad omakorda fenomenaalsetes või protofenomenaalsetes omadustes. Iga sellise inforuumi realiseering on tegelikult fenomenaalne (või protofenomenaalne) realiseering. Iga kord, kui realiseerub mingi omadus, näiteks mass või laeng, on selle omaduse taga mingi seesmine omadus — fenomenaalne või protofenomenaalne, lühidalt öeldes *mikrofenomenaalne* omadus. Meil on hulk mikrofenomenaalseid alusruume, üks iga fundamentaalse füüsikalise omaduse kohta, ja need ruumid ongi nende inforuumide kandjad, mida füüsika nõuab. Viimseteks erinevusteks ongi need mikrofenomenaalsed erinevused.

Loomulikult eeldab see positsioon taas mingis versioonis “padupansühhismi”, kuid seesugune pansühhism ei ole, nagu ma olen [juba](#PARA.158-8) püüdnud näidata, nii ebamõistlik, kui tavaliselt arvatakse. Võttes omaks oletuse, mille ma olen [juba](#PARA.158-1) välja pakkunud — et fenomenaalseid omadusi võib leiduda kõikjal, kus on informatsiooni —, võiksime leida neile omadustele ka kasuliku rolli.

Ontoloogiat, milleni me sel viisil jõuame, võiks õigusega nimetada *topeltaspektiontoloogiaks*​. Füüsika vajab infoseisundeid, kuid teda huvitavad nende seisundite juures ainult nende suhted, mitte nende seesmine olemus; fenomenoloogia vajab infoseisundeid, kuid hoolib ainult nende seesmisest olemusest. Kõnealuse kontseptsiooni järgi on olemas infoseisundite üksainus, neid kaht aspekti teineteisega ühendav aluskomplekt. Võiksime öelda, et nende seisundite seesmised aspektid on fenomenaalsed, välised aga füüsikalised. Või aforistlikult: elamus on informatsioon seestpoolt; füüsika on informatsioon väljastpoolt.

####  Aga makroskoopiline fenomenoloogia?

Kõik see toimib ontoloogiana väga kenasti, metsikupoolsest väljanägemisest hoolimata. Aga enne kui liiga hoogu sattuda, mõelgem ühele tohutu suurele siiani vastamata küsimusele: Kuidas ühildada seda ontoloogiat psühhofüüsikateooria üksikasjadega? Kuidas ühildada teda näiteks psühhofüüsikaliste reeglipäradega makroskoopilisel tasemel? Raskus on selles, et siin kehtib topeltaspektiprintsiip ​**[lk.306]**​ fundamentaalfüüsika tasemel, kus füüsikaliselt realiseeruval mikroskoopilisel informatsioonil on ka fenomenaalne realiseering. Kuid teadvuseteooria jaoks on tarvis, et fenomenaalne realiseering oleks ka füüsiliselt realiseeruval makroskoopilisel informatsioonil. Ja pole sugugi ilmne, et sellist “makroskoopilist fenomenoloogiat” saab tuletada mikroskoopilisest fenomenoloogiast.

Pealtnäha ei tundu sugugi, nagu oleks meie teadvuselamus mingi (näiteks) meie aju füüsikalistele põhiomadustele vastavate mikrofenomenaalsete omaduste summa. Meie elamus tundub palju terviklikuma ja homogeensemana, kui seda oleks ükskõik missugune lihtne summa. See on üks versioon “teralisusprobleemist”, mille Sellars (1965) esitas materialismile: kuidas saaks elamus olla identne tohutu hulga füsioloogiliste sündmuste kogumiga, arvestades ühtede homogeensust ja teiste peeneteralisust? Analoogiline probleem on eriti aktuaalne sedasorti russelliaanlike vaadete jaoks, mille üle ma praegu arutlen[[363]](#footnote-364). Kui kõik fenomenoloogia juured on mikrofenomenoloogias, siis on raske näha, kust tuleb sile, struktuurne makroskoopiline fenomenoloogia: selle asemel ootaks pigem mingit “sakilist”, struktuuritut fenomenide kuhjumit.

Selle küsimuse lahendamiseks on mitmesuguseid võimalusi. Esiteks võib püüda kujutada asja nii, nagu kehtiks topeltaspekti­ontoloogia kõigil tasemetel, mitte ainult mikroskoopilisel. See tähendab, et isegi makroskoopilisel tasemel on füüsikaliste inforuumide aluseks fenomenaalne realiseering. On võimalik väita, et mikroskoopiline tase ei ole mingis eristaatuses: siin on kõik lihtsam, kuid midagi ontoloogiliselt erilist ei tarvitse selles tasemes olla. Argumendid, mida me oleme esitanud informatsioonilise vaateviisi kohta füüsikalisele maailmale, kehtivad ka makroskoopilisel tasemel. Võiks öelda, et isegi sellel tasemel leidub ruumi makroskoopilistele eristavatele erinevustele, milledest igaüks võib realiseeruda vastavas fenomenoloogias.

Häda on selles, et kõigile neile omaette fenomenaalsetele realiseeringutele ei pruugi ruumi jätkuda. Tundub, et kui fundamentaalsed füüsikalised tunnused on fenomenaalsetes inforuumides juba realiseerunud, siis on makroskoopilisel informatsioonil juba alus olemas: eristavate erinevuste aluseks on mikroskoopiliste füüsikaliste parameetrite konfiguratsioonid, millel on omakorda juured mikrofenomenoloogias. Sellegipoolest võiks püüda kehtestada mingi fenomenaalne alus, kuid see oleks nähtavasti liiane ja teoreetiliselt vähem elegantne kui vastav samm mikroskoopilisel juhul. Liiasusest lahtisaamiseks võiks lugeda makroskoopilise aluse primaarseks, kuid siis oleks raske selgitada juhte kinniste mikroskoopiliste süsteemidega jms. Seega ei ole selge, kas “aluse otsimise” vaade topeltaspektiontoloogiale saab makroskoopilisel tasemel otseselt toimida.

​**[lk.307]**​ Teiseks võiks otsida võimalust, kuidas need mikrofenomenaalsed omadused võiksid makroskoopilist fenomenoloogiat konstitueerida. Esimesel pilgul ei tundu, et makroskoopiline fenomenoloogia võiks olla nende omaduste mingi lihtne summa või kogum; vastasel korral kerkiks meie ette kohe “sakilisuse” probleem. Ent võib-olla on viga lihtsalt selles, et me ei mõista — nagu Nagel (1986) on seda väljendanud — osa ja terviku suhet vaimu sfääris. Teisiti öeldes, meil puudub täpne ettekujutus sellest, kuidas madaltaseme mikrofenomenaalsd omadused annaksid “summeerudes” kokku kõrgtaseme fenomenoloogia. Me kipume mõtlema neist asjadest füüsikaliste analoogiate abil ja kujutame seda summeerumist samamoodi, nagu mikrofüüsika summeerub makrofüüsikaks. Kuid see ei tarvitse olla õige mõtteviis. Võib-olla konstitueerub fenomenoloogia mingil hoopis teisel viisil.

Näiteks võib asi olla nii, et mikrofenomenaalsed omadused summeeruvad makrofenomenoloogiaks viisil, mis ei kajasta mitte nende aegruumilist, vaid informatsioonilist ühendstruktuuri. Kui mingi hulk neid omadusi realiseerib omavaheliste põhjuslike suhete varal ühiselt mingi keeruka infoseisundi, siis võiks ehk oodata, et suvalisel niiviisi sündival makrofenomenoloogial on selle infoseisundi vorm. On ju mikrofenomenaalsete omaduste keskseks rolliks infoseisundite realiseerimine, mistõttu poleks väga üllatav, kui infostruktuur mängiks mingit rolli omaduste omavahelistes konstitutiivsetes suhetes. Kui see oleks nii, siis oleks kõigil säärastel makrofenomenaalstel seisunditel “sile” informatsiooniline struktuur, nagu algne topeltaspekti printsiip ennustaski. Seda ei ole kerge mõista, kuid me ei võigi oodata, et meie tavapärane arusaam füüsikalisest maailmast kehtiks ka fenomenaalses maailmas. Seega oleks parem arusaam fenomenoloogia enda olemusest võib-olla kooskõlas just niisuguste konstitutiivsete suhetega.

Kui peaks selguma, et mitte mingisugune konstitutiivne seos niiviisi toimida ei saa, võiksime proovida kolmandat võimalust, milleks on makrofenomenoloogia ühendamine mikrofenomenoloogiaga seaduste vahendusel.[[364]](#footnote-365) Näiteks võib-olla kehtib lihtsalt seadus, et kui mikrofenomenaalsed seisundid realiseerivad omavaheliste põhjuslike suhete varal (“eristava erinevuse” põhimõttel) mingit kindlat liiki infoseisundit, siis tekib seeläbi ka sama seisundi otsene fenomenaalne realiseering. See lahendaks teoreetilised probleemid, kuid lahenduse hinnaks oleks ontoloogia keerustumine. Meie ontoloogia ei oleks enam lihtne ontoloogia, kus fenomenoloogia on füüsikaliselt realiseeruva informatsiooni seesmine aspekt: mingi osa fenomenoloogiast jääks seaduste vahendusel selle informatsiooni “ripatsiks”, nagu me näeme ka tavapärasema omadusdualismi puhul. Nii jääks russelllik vaade küll veel täiesti koherentseks, ent kaotaks osa oma atraktiivsusest.

Nii või teisiti jätan ma selle küsimuse siin lahtiseks. See on kindlasti kõige suurem takistus russelliaanliku vaate teel, kuid ei ole ilmne, et sellest pole võimalik üle saada. Kui selle raskusega võib toime tulla, siis hakkab teine strateegia tunduma eriti paljutõotavana; kuid võib-olla on takistusest ülesaamiseks ​**[lk.308]**​ tarvis mingit täiesti uut ideed. Optimistlikult vaadates võime näha selles probleemis — kuidas ühildada psühhofüüsikateooriat nii meie fenomenoloogia ja tema füüsikalise kandja makrotasandi faktidega kui ka Russelli kontseptsiooni mikrotasandilise ontoloogiaga — üht neist olulisest kitsendustest, mis võivad meid pikapeale viia üksikasjaliku teadvuseteooriani. Üks teooria konstrueerimise raskusi ongi ju kitsenduste puudus. Võib-olla juhataks just see probleem meid õigesse suunda.

Kui peaks selguma, et meid ei rahulda ükski neist strateegiatest, siis tuleb russelliaanlikust vaatest taganeda ja võtta omaks mingi teistsugune arusaam metafüüsikast. Võib teha katset näiteks puhta informatsiooni metafüüsikaga kui ühe võimalusega füüsikalise maailma mõistmiseks, seejärel aga ühendada sellega kuidagi fenomenoloogia, võib-olla näiteks seaduspärase seose kaudu puhta informatsiooniga. Võib kaaluda ka lihtsalt taganemist “taltsa” ontoloogia juurde, mille järgi on eraldi olemas nii füüsikaline maailm kui ka fenomenaalne maailm, kumbki omaenda seesmise olemusega, ning kus neid ühendavad teineteisega seaduspärased seosed, mis toetuvad mingil moel informatsiooniprintsiibile. See tähendaks, et jutte topeltaspektsusest tuleks võtta vähem tõsiselt ja ontoloogia tuleks mõnevõrra vähem elegantne, kuid ikkagi on võimalik, et ka see tee viib täiesti rahuldava teooriani.

### 6. Lahtised küsimused

Psühhofüüsikaseaduste informatsiooniline mudel, mille ma visandasin, jätab tohutu hulga küsimusi lahtiseks. Selleks, et jõuda siinsest pildist lõpliku teooriani, tuleks leida vastused kõigile nendele küsimustele. [Eelmises jaos](#ORG.8.5) mainisin ma selle vaate ontoloogia mõningaid raskusi. Kuid arvukalt küsimusi on ka seaduste kuju kohta ja selle kohta, kuidas meie fenomenoloogiat õigupoolest seletada tuleks. Mõned neist küsimustest on järgmised:

1. Miks toimub inforuumi fenomenoloogiline realiseerumine just nii, aga mitte teisiti? Näiteks kui meie fenomenaalne värvusruum võib inverteeruda, siis tundub mõnevõrra meelevaldne, et ta on niisugune, nagu ta on. Kas selle küsimuse lahendamiseks tuleb juurde tuua uusi seadusi või postuleerida kontingentseid konstante?
2. Kas fenomenaalse inforuumi iseloomu määrab selle ruumi struktuur (​võib-olla välistamata inversioonide võimalikkust​[[365]](#footnote-366))? Näiteks on nii värvus- kui ka maitseruum pealtnäha lihtsad kolme- või neljamõõtmelised ruumid, kuid vaatamata sarnasele kujule on neil väga erinev iseloom. Saaks püüda näidata, et nende struktuuride sarnasus on illusioon ja et kui me põimime mõlemad suuremasse struktuuri — nähes näiteks värvuselamustes üht osa ​**[lk.309]**​ visuaalsete elamuste sügavast kogustruktuurist —, siis sarnasus kaob. Kuid endiselt jääb lahtiseks küsimus: kas visuaalse värvusinfo projitseerumine fenomenoloogiasse saab toimuda ainult enam-vähem samasugusel viisil, nagu töötavad meie värvusaistingud, või leidub ka hoopis teistsuguseid võimalusi? Mulle tundub, et õige on pigem esimene kui teine vastus, kuid ei ole kaugeltki selge, kuidas seda põhjendada tuleks.
3. Seni olen ma oma mudeli abil arutlenud peamiselt lihtsate tajuelamuste, näiteks värvuselamuste üle. Ei ole selge, kuidas mudelit laiendada, nii et ta lubaks käsitleda ka peenemaid elamusi, näiteks keerukaid emotsionaalseid elamusi või jooksvalt pähe tulevaid mõtteid. Kas selline laiendamine on võimalik?
4. Missugune formaalne struktuur vastab kõige paremini fenomenaalse informatsiooni struktuurile? Missugused topoloogilised ruumid sobiksid elamuse relatsioonilise struktuuri tabamiseks? Kas tuleks võtta tarvitusele mingi spetsiifilisem struktuur, näiteks meetriline ruum või diferentseeruv muutkond? Elamuse kombinatoorne struktuur on veelgi huvitavam: näiteks visuaalse välja struktuuri kujutamine tavalise mitmemõõtmelise kontiinumina oleks tõenäoliselt liiga suur lihtsustus. Kuidas oleks kõige parem edasi anda kogu struktuuri? Kas sel eesmärgil tuleb muuta inforuumi definitsiooni?
5. Kuidas seletada minu mudeli abil teadvuse ühtsust? Teisiti öeldes, mis teeb minu nägemiselamused, kuulmiselamused jne. kõik ühe ja sama subjekti elamusteks? Ma kahtlustan, et vastus oleneb asjaomase informatsiooni töötelmise viisist, nii et teadvuse ühtsus kajastab relevantse informatsiooni kättesaadavust mingil kindlal viisil integreerimiseks. Ent pole selge, kuidas seda täpselt formuleerida.
6. Missuguste on täpsed kriteeriumid, mille järgi saaks teha kindlaks, milline informatsioon ajus vastab minu enda elamustele? Kas leidub mingi sellekohane konkreetne põhjuslik trajektoor või mingit kindlat liiki põhjuslikkusvoog? Arvatavasti on siin keskset rolli mängimas otsene kättesaadavus globaalseks juhtimiseks või miski muu, mis individualiseerib informatsiooni ja asjakohased trajektoorid.​

See, et meil pole ühelegi neist küsimustest vastust, näitab, kui kaugel on need skemaatilised ideed tõelisest teooriast. Et seda mõista, tarvitseb vaid panna tähele, kui kaugel on need ideed teooriast, mis lubaks täpselt ennustada füüsilise süsteemi fenomenaalseid omadusi tema füüsikaliste omaduste põhjal. Praegusel kujul jääb minu ideel vajaka seletus- ja ennustusvõimest; et panna ta reaalset kasu tooma, tuleb teda oluliselt edasi arendada.

​**[lk.310]**​ Et jõuda sellest ideest rahuldava teooriani, on tarvis tervet hulka uusi ideid. Võib-olla viib läbimurdeni [eelmises jaos](#ORG.8.5) vaadeldud probleemi uurimine: kuidas viia fenomenaalne makrotaseme informatsioon vastavusse vaatega informatsioonile kui “seesmisele omadusele” mikrotasemel? Teise läbimurdeni võiks ehk viia katse leida kitsendus, mille rakendamise eraldaks füüsiliselt realiseeruvate inforuumide klassist inforuumid, mis realiseeruvad elamuses. Läbimurdeid võib tulla ka suundadest, millega ma pole osanud arvestadagi.

See idee võib osutuda ka täiesti ekslikuks. See ei üllataks mind; tegelikult ongi minu arvates tõenäolisem, et fundamentaalse teooria võti peitub hoopis mujal. Kuid ma esitasin need ideed siiski, sest me peame hakkama nende küsimuste üle mõtlema, ja isegi halb sellesuunaline näide võib olla õpetlik. Samuti loodan ma, et mõned meie arutluste käigus vaadeldud ideed — ideed fenomenaalsete otsustuste seletamise, elamuse kõikjalolu ning elamuse, informatsiooni ja seesmiste füüsikaliste omaduste vahelise seose kohta — võivad osutuda kasulikuks isegi siis, kui kanda nad üle teistsugusesse mudelisse. Võib-olla jääb adekvaatsemasse teadvuseteooriasse alles mingi maik ka siin esitatud ideedest, isegi kui üksikasjades osutub olevat suuri erinevusi.

Sageli heidetakse seesugustele teadvuseteooriatele ette, et need on liiga spekulatiivsed ja mittekontrollitavad. Kuid ma arvan, et minu siinse “teooria” tegelik häda on mujal: ta ei anna küllalt täpseid ennustusi. Kui meil oleks umbkaudu sama lihtne teooria, mis suudaks meie infotöötlussüsteemi füüsikalisi fakte teades ennustada kõiki meie elamuste spetsiifilisi fakte — kasvõi ainult neidki, mida me teame esimese isiku seisukohalt —, oleks see juba tähelepanuväärne saavutus ja annaks meile väga tõsist alust tunnistada see teooria õigeks. Praegu meil sellist teooriat ei ole, kuid pole põhjust arvata, et selleni ei olegi võimalik jõuda.

# ​**[lk.311]**​ IV OSA: Rakendused

## ​**[lk.313]**​ 9. peatükk. Tugev tehisintellekt

### 1. Masinteadvus

Kas masinal võib olla teadvus? Kas arvutit saab programmeerida nii, et tal oleks tõeline vaim? Nende küsimuste üle on viimastel aastakümnetel tohutult palju vaieldud. Suur osa *tehisintellekti*​ (ehk AI [​*artificial intellect*​]) valdkonnas tehtud uuringutest on võtnud endale eesmärgiks jäljendada arvutiga vaimset elu. Seni pole edusammud olnud kuigi suured, kuid selle suuna pooldajad väidavad, et meil on põhjust uskuda, et kunagi on arvutitel tõepoolest mõistus. Vastased aga kinnitavad, et arvuti võimetel on piirid, mida inimesel ei ole, mistõttu teadvusega vaimu sünd ainuüksi arvutuslike protsesside varal ei tule kõne allagi.

Tüüpilisi vastuväiteid tehisintellektile on kahesuguseid. Esiteks *välised* vastuväited, millega püütakse tõestada, et arvutisüsteem ei saa iial käitudagi nagu kognitiivsed süsteemid: inimesel on funktsionaalseid võimeid, mida ei saaks olla ühelgi arvutil. Mõnikord väidetakse näiteks (vt. Dreyfus 1972), et kuna need süsteemid järgivad reegleid, ei saa nad ilmutada inimesele omaselt loovat või paindlikku käitumist. Teised on väitnud, et arvutid ei suuda iial dubleerida inimese matemaatilist taipu, kuna erinevalt inimesest kitsendab arvutisüsteemide võimalusi Gödeli teoreem (Lucas 1961; Penrose 1989).

Välistele vastuväidetele on olnud raske veenvust anda — seda takistab füüsikaliste protsesside arvutusliku simulatsiooni kõigitine edukus. Eelkõige on meil nähtavasti põhjust uskuda, et füüsikaseadused on arvutatavad, ja seega peaks olema võimalik inimese käitumist vähemasti arvutuslikult simuleerida. Mõnikord vaieldakse sellele vastu, et füüsikaseadustes on midagi, ​**[lk.314]**​ mis ei ole arvutatav (nii Penrose), või et on olemas ka mittefüüsikaline põhjuslikkus (Lucas), kuid on selge, et nende vastuväidete esitajatel tuleb pidada ebavõrdset võitlust.

Rohkem on levinud vastuväited, mida ma nimetan *seesmisteks*​. Nende esitajad on nõus vähemalt kaaluma oletust, et arvutid võivad simuleerida inimese käitumist, kuid kinnitavad, et mõistust neil sellegipoolest olla ei saa. Muu seas väidetakse, et arvutil ei oleks siseelu: ei teadvuselamusi ega tõelist arusaamist. Parimal juhul võiks arvuti vaimset elu simuleerida, kuid mitte seda luua. Sedaliiki vastuväidetest kõige tuntum on John Searle’i “hiina toa” argument (Searle 1980). Nende vastuväidete järgi võib arvutisüsteemil olla parimal juhul mõistuse tühi kest: nad oleksid ränist sombid.

Autoreid, kes pooldavad mittereduktiivset vaadet teadvuselamusele, on sageli köitnud seesmised vastuväited tehisintellektile, kusjuures paljud väidavad, et pelgal arvutil teadvust olla ei saa. Tõepoolest, need, kellele teadvuse probleem on muljet avaldanud, on selle iseloomustamiseks osutanud teadvusele kui omadusele, mis on meil, kuid mida ei oleks ühelgi arvutil! Paljudel on raske uskuda, et tehislikus, mittebioloogilises süsteemis võiks sündida teadvuselamus.

Mittereduktiivne arusaam teadvusest ei too aga automaatselt kaasa pessimistlikku vaadet tehisintellektile. Neil kahel küsimusel ei ole midagi ühist. Esimene puudutab füüsikaliste süsteemide ja teadvuse vahelise seose tugevust: kas füüsikalised protsessid konstitueerivadki teadvuse, või teadvus vaid tekib füüsikaliste protsesside pinnal? Teine puudutab seose kuju: missugune nimelt peab olema füüsiline süsteem, et temas võiks sündida teadvus? Muidugi pole ilmne, et õigete arvutuste sooritamisest võiks sündida teadvus; ent sama vähe on ilmne ka, et teadvus võiks sündida aju närviprotsessidest. Pealtnäha pole ju mingit põhjust arvata, et arvuti peaks selles suhtes ajule alla jääma. Kui me oleme juba kord nõustunud üllatava faktiga, et ajus võib sündida teadvus, siis ei läheks üllatus suuremaks, kui selguks, et teadvust saab luua ka arvutuslikul teel. Seega tuleb küsimust pidada lahtiseks, niipea kui me võtame omaks mittereduktiivse vaate teadvusele.

Käesolevas peatükis liigun siit edasi ja püüan näidata, et tehisintellekti ambitsioonid on mõistlikud. (Vt. joonist 9.1.) Ma pooldan seisukohta, mida Searle nimetab *tugevaks tehisintellektiks*​: et on olemas mittetühi arvutuste klass, millesse kuuluva arvutuse implementatsioonist piisab mõistuseks, sealhulgas ka teadvuselamuseks. See piisavus on muidugi ainult looduslikult paratamatu: loogiliselt on subaline arvutamine võimalik ilma teadvuseta. Kuid nii on see ju, nagu oleme näinud, ​**[lk.315]**​ ka ajuga. Masinteadvuse perspektiivide hindamisel aktuaalses maailmas huvitavad meid ainult looduslik võimalikkus ja looduslik paratamatus.

Joonis 9.1. Bloomi maakond tugevast AI-st. (​© 1985, Washington Post Writers Group. Reprinted with permission​)

* Ma mõtlen. Järelikult …
* ma olen. Ma olen!

- Ma mõtlen! Järelikult ma olen elus. Mul on elu, mõtlemine ja psüühika! Magus teadvus! Ja surematu hing …  — Sutsti!

(Et lugeja ei peaks seda järeldust [eelmises peatükis](#ORG.8.4.2) esitatud panpsühhistlikke ideede valguses triviaalseks, tähendan, et käesolevas peatükis pole midagi, mis seal avaldatud mõtetest sõltuks. Tõepoolest, ma väidan, et teadvuseks piisab õige arvutuse implementatsioonist, ja veel enamgi: et õige arvutuse implementatsioon on piisav sama külluslikuks teadvuselamuseks, nagu seda on meie oma.)

Suurema osa tugeva tehisintellekti kaitsmiseks vajalikust tööst tegin ma juba [7. peatükis](#ORG.7), organisatsioonilise invariantsuse printsiipi põhjendades. Kui sealsed mõttekäigud on õiged, siis on õige ka see, et suvalisel süsteemil, millel on õige funktsionaalne organisatsioon, on teadvus, olenemata tema koostisest. Seega me juba teame, et näiteks see, et miski on valmistatud ränist, ei takista seda, et tal võiks olla teadvus. Jääb veel vaid selgitada välja seos arvutuste ja funktsionaalse organisatsiooni vahel, tuvastamaks, et sobiva arvutuse implementatsioonist piisab asjakohase funktsionaalse organisatsiooni tagamiseks. Niipea kui see on tehtud, järeldub edasi kohe tugeva tehisintellekti võimalikkus. Ma vastan siin ka mitmele tugeva tehisintellekti projekti vastasele argumendile.

### 2. Arvutuste implementeerimisest[[366]](#footnote-367)

Oma standardsel kujul käsitleb teoreetiline informaatika ainuüksi abstraktseid objekte: Turingi masinaid, Pascali programme, lõplikke seisundiautomaate jne. Need on matemaatilised entiteedid, mis asustavad matemaatilist ruumi. Seevastu reaalses maailmas on kognitiivsed süsteemid konkreetsed objektid, millel on füüsiline kuju ja mis vastasmõjustuvad põhjuslikult teiste füüsilise maailma objektidega. Kuid sageli tahame me teha teoreetilise informaatika abil järeldusi konkreetsete objektide kohta reaalses maailmas. Selleks on tarvis silda abstraktse teooria ja konkreetsete objektide vahel.[[367]](#footnote-368).

​**[lk.316]**​ Selleks sillaks on *implementatsiooni* [​*implementation*​] mõiste. Implementatsioon on suhe abstraktsete arvutuslike objektide (lühidalt “arvutuste”) ja füüsikaliste süsteemide vahel, kui füüsikaline süsteem “realiseerib” mingit arvutust või kui arvutus “kirjeldab” mõnd füüsikalist süsteemi. Arvutusi implementeeritakse sageli sünteetilistes, ränipõhistes arvutites, kuid neid võib implementeerida ka muul viisil. Näiteks öeldakse looduslike süsteemide, näiteks inimese aju kohta sageli, et nad implementeerivad arvutusi. Informaatilisi kirjeldusi rakendatakse füüsikalistele süsteemidele kõikvõimalikes valdkondades. Igal sellisel korral on vaikivalt või otsesõnu mängus implementatsiooni mõiste.

Implementatsiooni mõistet analüüsitakse harva üksikasjalikult, tavaliselt aktsepteeritakse seda ilma pikemata. Kuid tugeva tehisintellekti põhjendamiseks on tarvis implementatsiooni mõiste üksikasjalikku kirjeldust. Tugeva tehisintellekti tees on sõnastatud informaatika mõistetes ja ta ütleb, et teadvuseks piisab kohase arvutuse implementatsioonist. Selle väite hindamiseks on tarvis teada, missugustel juhtudel nimelt implementeerib füüsiline süsteem üht või teist arvutust. See välja selgitatud, võime kombineerida saadud teadmist psühhofüüsika seaduste analüüsiga, mille me ülal esitasime, et teha kindlaks, kas neist saaks tuletada soovitud järelduse.

Mõned autorid on väitnud, et implementatsiooni ei olegi võimalik sisukalt kirjeldada. Näiteks Searle (1990b) on väitnud, et see, kas miski on implementatsioon, ei ole objektiivne fakt, vaid sõltub vaatlejast: iga süsteemi võib sobivalt tõlgendades pidada suvalise arvutuse implementatsiooniks. Muu hulgas kinnitab Searle, et tema toaseina võib vaadelda kui tekstiredaktori WordStar implementatsiooni. Kui see oleks nii, siis oleks raske näha, kuidas võiksid informaatika mõisted mängida mingisugustki fundamentaalset rolli teoorias, mis viimselt käsitleb konkreetseid süsteeme. Mis puutub tugevasse tehisintellekti, siis oleks see kas sisutu või eeldaks mingil tugeval kujul pansühhismi. Kuid ma arvan, et seesugune pessimism on alusetu: implementatsiooni objektiivseks kirjeldamiseks pole ühtki takistust. Käesolevas peatükis esitangi ma skemaatilisel kujul ühe säärase kirjelduse. (See on veidi tehniline, kuid peatüki ülejäänud osa peaks olema arusaadav, isegi kui tehnilised üksikasjad vahele jätta).

Igasugune seletus selle kohta, mida tähendab arvutuse implementatsioon, sõltub arvutuste klassist, millest parajasti jutt käib. On olemas terve hulk mitmesuguseid informaatilisi formalisme ja neile vastavaid arvutuste klasse: Turingi masinad, lõplikud seisundiautomaadid, Pascali programmid, konnektsionistlikud võrgud, rakkautomaadid jne. Põhimõtteliselt tuleb iga säärase formalismi puhul seletada, mida tähendaks tema implementatsioon. Ma selgitan siin implementatsiooni tähendust ainult ühe formalismi, nimelt kombinatoorsete seisundiautomaatide näitel. See arvutuste klass on piisavalt üldine, et laiendada seda, mida me tema implementatsiooni kohta ütleme, teistele klassidele.

Kombinatoorne seisundiautomaat on lõplike seisundiautomaatide nutikam sugulane. Lõpliku seisundiautomaadi (​*finite-state automaton*​, FSA) definitsiooni koosseisu kuuluvad tema võimalikud sisendandmed, seesmised seisundid ja väljundandmed (mida kõiki on lõplik hulk) ning ​**[lk.317]**​ seisundisiirete tabel. FSA seesmine seisund on lihtne, sisestruktuurita element ​*Sᵢ*​; samasugused on ka sisend- ja väljundandmed. Seisundisiirdete tabel ütleb, missugused uued seisundid ja väljundandmed vastavad igale võimalikule seisundi ja sisendandmete kombinatsioonile. Siirdetabeli põhjal saab tuletada, kuidas mingi algseisundiga FSA aja jooksul muutub ja missuguseid andmeid ta väljundisse annab, kui sisendandmed on nii- või naasugused. FSA arvutusliku struktuuri moodustabki see suhteliselt lihtne kogum siirdesuhteid struktuurita seisundite vahel.

Lõplikud seisundiautomaadid ei ole piisavalt võimsad representeerimaks enamiku praktiliste arvutuste struktuuri, kuna nende arvutuste seisunditel ja seisundisiiretel on tavaliselt keerukas seesmine struktuur. Mitte ükski FSA kirjeldus ei suuda anda edasi mõne Pascali programmi, Turingi masina või rakkautomaadi kogu struktuuri. Seepärast on kasulikum keskenduda mõnele kindla­struktuuriliste seesmiste seisunditega automaatide klassile.

Kombinatoorsed seisundiautomaadid (​*combinatorial-state automata*​, CSA) erinevad FSA-dest ainult selle poolest, et nende seisunditel on kindel struktuur. CSA seisundit kujutab vektor $\left[S^{1},S^{2},…,S^{n}\right]$, mis võib olla lõplik või lõpmatu, kuid mina võtan siin vaatluse alla ainult lõpliku juhu. Vektori elemente võib vaadelda kui seesmise seisundi komponente; nad vastavad rakkautomaadi rakkudele või Turingi masina lindipesadele ja lugemispea seisundile. Vektori iga element $S^{i}$ võib omandada väärtusi lõplikust hulgast $S\_{j}^{i}$, kus $i$ on elemendi number ja $j$ on elemendi väärtuse indeks tema võimalike väärtuste vektoris. Neid väärtusi võib vaadelda kui masina koguseisundi “osaseisundeid”. Sarnase keeruka struktuuriga on ka sisend- ja väljundandmed: sisendandmed moodustavad vektori $\left[I^{1},…,I^{k}\right]$, väljundandmed vektori $\left[O^{1},…,O^{m}\right]$.

CSA definitsioon koosneb seisundi-, sisend- ja väljundvektorite definitsioonist ja siirdetabelist — reeglitest, mis määravad, kuidas CSA seisund aja jooksul muutub. Siirdereeglid ütlevad, kuidas seisundivektori iga elemendi uus seisund oleneb sisendandmetest ja senisest seisundist. Väljundvektori igale elemendile määrab siirdereegel senise seisundi põhjal uue väljundväärtuse. Iga lõplikku CSAd saab kirjeldada sama võimsusega FSAna, kuid FSA kirjelduses läheb kaduma suurem osa struktuurist, mis on CSA jaoks oluline. Kuid vaimu aluseks oleva organisatsiooni kujutamisel CSA-de abil on sel struktuuril keskne tähtsus.

Nüüd võime hakata juba ka selgitama, mis on implementatsioon. Arvutussüsteemid, näiteks CSAd, on abstraktsed objektid, mille formaalse struktuuri defineerivad nende seisundid ja siirdesuhted. Füüsikalised süsteemid on konkreetsed objektid, mille põhjusliku struktuuri määravad kindlaks nende seesmised seisundid ja seisundite vahelised põhjuslikud suhted. Mitteformaalselt ütleme, et füüsikaline süsteem *implementeerib* arvutust, kui süsteemi põhjuslik struktuur kajastab ​**[lk.318]**​ arvutuse formaalset struktuuri. See tähendab, et süsteem implementeerib arvutust, kui on võimalik korraldada vastavus tema seisundite ja arvutuse seisundite vahel nii, et füüsikalised seisundid, millede vahel on põhjuslik seos, vastavad formaalsetele seisunditele, millede vahel on vastav formaalne seos.

CSA-de implementatsioonist ettekujutuse saamiseks tarvitseb vaid rakendada seda intuitiivset ideed. Füüsiline süsteem on mingi CSA implementatsioon, kui tema seisundid saab lahutada osaseisunditeks, sisend- ja väljundandmed lahutada nende andmete osaseisunditeks ning korraldada vastavuse tema osaseisundite ja CSA osaseisundite vahel nii, et füüsiliste seisundite, sisendite ja väljundite põhjuslikud siirdesuhted peegeldaksid arvutussüsteemi vastavate formaalsete seisundite, sisendite ja väljundite siirdesuhteid.

CSA implementatsiooni formaalne kriteerium on järgmine:[[368]](#footnote-369)

Füüsiline süsteem $P$ on CSA $M$ *implementatsioon*​, kui tema seesmised seisundid saab lahutada komponentideks $\left[s^{1},…,s^{n}\right]$ ja seada osaseisundid $s^{i}$ vastavusse automaadi $M$ osaseisunditega $\left[S^{1},…,S^{n}\right]$ ning korraldada analoogilised vastavused ka mõlema süsteemi sisend- ja väljundandmete vahel, nii et seisund, millesse süsteem $P$ satub suvalise seisundi $\left[s^{1},…,s^{n}\right]$ ja suvaliste tema sisendisse saabuvate andmete $\left[i^{1},…,i^{k}\right]$ korral, vastab äsjakirjeldatud vastavuse mõttes alati seisundile $\left[S’^{1},…,S’^{n}\right]$, millesse siirdub CSA $M$, kui ta on vastavas seisundis $\left[S^{1},…,S^{n}\right]$ ja tema sisendisse saabuvad vastavad andmed $\left[I^{1},…,I^{k}\right]$ ning kui andmed $\left[o^{1},…,o^{l}\right]$, mida $P$ seejuures väljastab, vastavad sama vastavuse mõttes andmetele $\left[O^{1},…,O^{l}\right]$, mida väljastab $M$.

Võime leppida kokku, et füüsilise süsteemi seisundi lahutamisel osaseisundite vektoriks peab vektori iga elemendi väärtus olema superventne süsteemi eri osas, tagamaks, et põhjuslik organisatsioon seostaks omavahel süsteemi erinevaid komponente.[[369]](#footnote-370) Vastasel korral ei ole ilmne, et kirjeldatud põhjuslik struktuur leidub füüsilises süsteemis tõepoolest kõigis oma üksikasjades. Toodud definitsioonis on selle ja teiste üksikasjade muutmiseks veel kõpitsemisruumi. Implementatsiooni mõiste ei ole kivisse raiutud ja seda võib eesmärgist olenevalt avardada või kitsendada. Kuid kirjeldatud üldine kuju on ühine kõigile arusaamadele implementatsioonist.

Võib tunduda, et CSA-d ei ole FSA-dega võrreldes kuigi pikk samm edasi. Saab ju iga lõpliku CSA puhul leida vastava FSA, mis käitub sisendis ja väljundis täpselt samamoodi. Kuid nende vahel on siiski olulisi erinevusi. Esiteks ja eelkõige on CSA-del palju kitsamad implementatsionoonitingimused kui vastavatel FSA-del. CSA implementatsioon eeldab paljude eraldi osade keerukat põhjuslikkku vastastoimet; seega suudab süsteemi kirjeldus CSA-na tabada tema põhjuslikku organisatsiooni palju üksikasjalikumalt. Teiseks saab CSA-de abil ühetaoliselt kirjeldada nii lõplike kui ka lõpmatute masinate implementatsioonitingimusi. Ja kolmandaks võimaldavad CSA-d otseselt kajastada paljudele informaatilistele objektidele (näiteks Turingi masinatele ja rakkautomaatidele) omast keerukat formaalset organisatsiooni. Samaväärses FSA-s läheb suur osa sellest struktuurist kaduma.

Tõepoolest, ülaltoodud definitsiooni abil on lihtne formuleerida implementatsioonikriteeriume ka muudliiki arvutustele. Näiteks Turingi masina implementatsiooni defineerimiseks tarvitseb asjaomane masin vaid CSA kujule ümber kirjeldada ning seejärel rakendada ülaltoodud definitsiooni. Selleks tuleb kirjeldada Turingi masina seisundit hiiglasuure vektorina, mille üks element tähistab masina lugemispea asukohta ning igale lindipesale vastab üks ülejäänud elementidest, mis nimetab seda pesa täitva sümboli ja ühtlasi ütleb, kas masina lugemispea on selle pesa kohal või mitte. Vektoritevahelised siirdereeglid tulenevad loomulikul teel lugemispea ja lindi käitumise kirjeldusest. Turingi masina puhul on vektorid muidugi lõpmatud, kuid lõpliku juhu implementatsioonitingimusi on hõlpus laiendada ka lõpmatule juhule. Kuna Turingi masin on säärasel viisil CSA kujule ümberkirjeldatav, siis võib öelda, et füüsiline süsteem on mingi Turingi masina implementatsioon parajasti siis, kui ta on vastava CSA implementatsioon. Kuna sarnane ümberkirjeldamine on võimalik ka teiste formaalsete süsteemide, näiteks rakkautomaatide ja Pascali programmide puhul, siis on võimalik leida implementatsioonitingimused ka kõigile arvutustele, mis kuuluvad mõnesse nendest klassidest.

See annab meile arvutussüsteemi implementatsiooni täiesti objektiivse kriteeriumi. Arvutussüsteemi implementatsioon ei kuku tühjusse kokku​, nagu oletab Searle. On tõsi, et leidub arvutusi, mille implementatsiooniks on suvaline füüsiline süsteem. Näiteks on suvaline füüsiline süsteem ühe elemendi ja ühe seisundiga CSA implementatsioon, ja peaaegu sama palju on ka kahe seisundiga CSA-de implementatsioone. Samuti on tõsi, et enamik füüsilisi süsteeme implementeerib rohkem kui üht arvutust olenevalt sellest, kuidas süsteemi olek seisunditeks lahutada. Selles ei ole midagi üllatavat: on täiesti ootuspärane, et minu tööarvuti, nagu ka minu aju, implementeerib tervet hulka arvutusi.

Oluline on aga see, et pole mingit põhjust uskuda, nagu oleks iga süsteem iga CSA implementatsioon. Kui CSA on keerukas, siis on tema implementatsiooniks nõutava põhjusliku organisatsiooniga füüsilisi süsteeme väga vähe. Kui võtta CSA, mille seisundivektorid koosnevad tuhandest elemendist, millest igal on kümme võimalikku väärtust, siis saame tõenäosuseks, et juhuslikult valitud füüsiliste seisundite kogumis on olemas vajalikud põhjuslikud suhted, [7. peatükis](#PARA.135-6) esitatud mõttekäikude abil vähem kui ${1}/{\left(10^{1000}\right)^{\left(10^{10000}\right)}}$. (Tegelikult veel palju vähem, sest siirded peavad töötama alati ühtviisi).[[370]](#footnote-371) [[371]](#footnote-372)

Mis saab siis Searle'i väitest, et informaatiline kirjeldus sõltub vaatlejast? Mingil vähesel määral on see tõepoolest nii: iga füüsiline süsteem implementeerib tervet hulka arvutusi ja vaatleja eesmärkidest oleneb, millisele neist ta tähelepanu koondab. Kuid tehisintellekti ja kognitiivset informaatikat ​**[lk.320]**​ ei ohusta see kuidagi. Tõsiasjaks jääb, et iga konkreetne füüsiline süsteem kas on faktiliselt mingi konkreetse arvutussüsteemi implementatsioon või ei ole seda, ja mingi kindla arvutussüsteemi implementatsioonideks kvalifitseerub ainult mingi kitsas füüsiliste süsteemide klass. Ja sellest piisab informaatiliste kontseptsioonide metafüüsiliseks ja seletuslikuks mõttekuseks.

Väide, et üks või teine füüsiline süsteem on mingi kindla keeruka arvutuse implementatsioon, ütleb midagi väga olulist süsteemi põhjusliku struktuuri kohta — midagi niisugust, millest võib olla palju kasu kognitiivsete seletuste otsimisel ning mis võib aidata ka mõista teadvuse alust. Ainult väga spetsiifilise põhjusliku organisatsiooniga süsteemidel on lootust neid tugevaid implementatsioonilisi kitsendusi rahuldada. Seega ei ole ohtu, et meie mõttekäigud minetaksid sisukuse, ning arvutuse mõistest võib loota kindlat alust kognitiivsete süsteemide analüüsile.

### 3. Tugeva tehisintellekti kaitseks

CSA implementatsioon on hämmastavalt sarnane funktsionaalse organisatsiooni realiseeringuga. Tuletame meelde, et funktsionaalse organisatsiooni määratlevad tema abstraktsete komponentide arv, iga komponendi seisundite arv ja sõltuvussuhete süsteem, mis kirjeldab, kuidas iga komponendi seisundid sõltuvad varasematest seisunditest ja sisenditest ning kuidas väljundid sõltuvad varasematest seisunditest. CSA mõiste ongi tegelikult selle mõiste otsene formaliseering.

Tõepoolest, suvalise sedaliiki funktsionaalse organisatsiooni, nagu kirjeldatud [7. peatükis](#ORG.7.1), saab hõlpsasti abstraheerida CSA-ks. Tarvitseb vaid sätestada, et CSA seisundivektorites on organisatsiooni iga komponendi kohta üks element ning et CSA seisundite vahelised formaalsed seisundisiirded vastavad põhjuslikele sõltuvussuhetele komponentide vahel. Funktsionaalse organisatsiooni realiseering ei erine peaaegu üldse vastava CSA implementatsioonist. On mõned väikesed erinevused, näiteks sisendite ja väljundite erinev käitlemine, kuid need ei ole olulised.

Implementatsiooni käsitlus, mille ma siin välja pakkusin, toob seega selgesti esile seose põhjusliku ja informaatilise organisatsiooni vahel. Tänu sellele me näeme, et informaatiliste kirjelduste rakendamine füüsikalistele süsteemidele annab meile põhimõtteliselt kätte nende süsteemide põhjusliku organisatsiooni formaalse kirjelduse. Informaatika mõisted annavad meile ideaalse keele seesuguse abstraktse põhjusliku organisatsiooni täpseks kirjeldamiseks. Võib öelda, et see ongi põhjus, miks informaatika mõisted kognitiivteaduses kõikjal nii laialdaselt tarvitusel on. Keeruka kognitiivsüsteemi käitumise seletamisel on süsteemi abstraktne põhjuslik organisatsioon peaasi, ja informaatilised formalismid annavad ideaalse ​**[lk.321]**​ mudeli selle organisatsiooni kirjeldamiseks ja analüüsimiseks.[[372]](#footnote-373)

See seos teeb tugeva tehisintellekti kaitsmise lihtsaks. Ma olen juba esitanud argumente organisatsioonilise invariantsuse printsiibi kasuks, mis ütleb, et suvalise teadvuselamustega süsteemi puhul on teise samasuguse peeneteralise funktsionaalse organisatsiooniga süsteemi teadvuselamused kvalitatiivselt täpselt samad. Kuid me teame, et suvalise funktsionaalse organisatsiooni saab abstraheerida CSA-ks, mis implementeerub iga kord, kui realiseerub see organisatsioon. Sellest järeldub, et iga teadvusega süsteemi ​*M* funktsionaalse organisatsiooni saab peentes detailides abstraheerida CSA-ks ​*M*​, nii et iga süsteem, mis on CSA ​*M* implementatsioon, on sama funktsionaalse organisatsiooni realiseering, ja seega on tal teadvuselamused, mis on kvalitatiivselt eristamatud lähtesüsteemi omadest. Selles aga seisnebki tugeva tehisintellekti tees.

Näiteks aju neurontasemel kirjelduse saab abstraheerida  CSA-ks, kus iga neuroni kohta on üks seisundivektori element ja iga elemendi kohta osaseisund, mis kajastab neuroni võimalike seisundite diapasooni. CSA seisundisiirdereeglid kujutavad iga neuroni seisundi sõltuvust teiste neuronite seisunditest ja tema seisundite seoseid sisendite ja väljundiga. Kui ajus on ka teisi relevantseid komponente komponendid peale neuronite, võib ka need kirjeldusse sisse võtta. Iga seda CSA-d implementeeriva füüsikalise süsteemi funktsionaalne organisatsioon dubleerib neurontasemel kõigis üksikasjades aju funktsionaalset organisatsiooni. Invariantsusprintsiibi järgi on selle süsteemi elamused eristamatud ajuga seonduvatest elamustest.

Ahvatlev oleks vaadelda arvutit lihtsalt kui sisend-väljundseadet, mille sisendi ja väljundi vahel on ruumi vaid mõnedele formaal-matemaatilistele manipulatsioonidele. Selline vaateviis jätab aga tähelepanuta esmatähtsa tõsiasja, et arvutis on, nagu ajuski, talitlemas sisukas põhjuslik dünaamika. Tõepoolest, tavalises arvutis, mis implementeerib minu aju simulatsiooni üksikute neuronite haaval, valitsevad vooluahelate vahel reaalsed põhjuslikud seosed, mis peegeldavad täpselt neuronite vaheliste põhjuslike seoste mustreid. Igale neuronile vastab teda esindav mälupiirkond, ja iga selline piirkond realiseerub füüsiliselt elektrilise potentsiaalina mingis füüsilises ruumipiirkonnas. Ja nende ahelate vahelistest põhjuslikkusmustritest, nagu ka neuronite vahelistest põhjuslikkusmustritest ajus, sünnivadki kõik teadvuselamused.

Tugeva tehisintellekti teesi saab kaitsta ka otseselt, hääbuvate kvaalide argumendi ja hüplevate kvaalide argumendi abil. Iga kahe suvalise CSA implementatsiooni vahele jääb terve spekter juhtumeid, kus implementatsioonide füüsilised komponendid vahetuvad ükshaaval välja nii, et nende ja ülejäänud süsteemi põhjusliku vastastoime muster jääb kogu aeg samaks. Kui ühel süsteemidest on teadvus ja kui CSA peeneteraline funktsionaalne organisatsioon ühtib abstraktsel tasemel selle ühe süsteemi omaga, ​**[lk.322]**​ siis tuleneb mainitud argumentidest, et ka teisel süsteemil on teadvus ning et tema teadvuselamused on eristamatud esimese süsteemi omadest. Kui teisel süsteemil ei oleks teadvust, siis peaks vahepeale jääma mõni hääbuvate kvaalidega süsteem. Kui teisel süsteemil ei oleks teadvust või tema teadvuselamused erineksid esimese süsteemi omadest, siis saaks konstrueerida vahepealse hüplevate kvaalidega süsteemi. Need järeldused on ebaveenvad [7. peatükis](#ORG.7) öeldud põhjustel. Ent kui tunnistada, et kvaalid ei saa niiviisi hääbuda ega hüpelda, siis järeldub sellest, et mõlema algse süsteemi elamused on teineteisest eristamatud ning tugeva tehisintellekti tees peab paika.

Siin on üks väike peensus. Argument lähtub eeldusest, et aju organisatsiooni saab abstraheerida CSA kirjelduseks. Selleks pole tarvis muud, kui et selle organisatsiooni kirjeldus piirduks lõpliku arvu osadega, millest igal on lõplik arv relevantseid seisundeid. Kuid sellele võidakse vastu vaielda: näiteks, et võib-olla on iga neuroni kirjeldamiseks, temas kulgevate protsesside olulise pidevuse määratlemiseks tarvis lõpmatut hulka seisundeid; ja mõni teine väidab võib-olla, et üleminekud nende lõpmata paljude seisundite vahel ei ole arvutatavad. Võtan seesugused vastuväiteid arutuse alla [hiljem](#ORG.9.5.3); esialgu aga olen meeleldi valmis nõustuma järeldusega, et kui kognitiivne dünaamika on arvutatav, siis sünnib õiget liiki informaatilises organisatsioonis teadvus. Teisiti öeldes, mind huvitavad siin rohkem seesmised kui välised vastuväited. [Käesolevas peatükis allpool](#ORG.9.5) vaatlen ma aga ka mõningaid väliseid vastuväiteid.

### 4. Hiina tuba ja teisi vastuväiteid

Loomulikult on tugeva tehisintellekti vastased aeg-ajalt esitanud selle seisukoha vastu konkreetseid argumente. Kõige tuntuma neist võlgneme John Searle'ile, kes avaldas selle 1980. aastal oma artiklis “Vaimud, ajud ja programmid” [“Minds, Brains, and Programs”] ja mujal. Ma vastan nendele argumentidele ülalkirjeldatud mudeli abil.

#### Hiina tuba

Oma kuulsas tugeva tehisintellekti vastases argumendis väidab Searle (1980), et suvaline programm on implementeeritav, ilma et sellest sünniks vaim. Selleks esitab ta mõttekäigu, mida ta peab vastunäiteks tugeva tehisintellekti teesile: “hiina toa”, kus sümbolitega manipuleeriv inimene simuleerib hiina keele oskajat. Hiina toa mõte on tuua näide süsteemist, mis implementeerib suvalist programmi, kuid millel vastav teadvuselamus puudub.

Oma argumendi esialgse versiooniga ründab Searle pigem masinintentsionaalsust kui masinteadvust, väites, et hiina toal puudub “arusaamine”. Sellegipoolest on üsna selge, et põhimõtteliselt käib jutt teadvusest. Otseselt kinnitab argumendi tuum — kui argument on korrektne —, seda, et hiina toa süsteemil puuduvad ​**[lk.323]**​ teadvusseisundid, näiteks hiina keele mõistmise teadvuselamus. Searle'i arvates nõuab intentsionaalsus teadvust, ja seega näitab argument, et hiina toal puudub ka intentsionaalsus. Kuid mitte kõik ei ole sellega nõus. Igal juhul võime me teadvuse ja intentsionaalsuse vahelise seose küsimusest siinkohal mööda vaadata ning püstitada küsimuse ainult teadvuse kohta. Nii saavad vaidlusalused küsimused ehk mõnevõrra selgemaks.

(Teisiti öeldes, Searle'i järeldused saab jagada kaheks osaks: (1) teadvuseks ei piisa ühestki programmist; ja (2) intentsionaalsuseks ei piisa ühestki programmist. Searle usub, et esimesest järeldub teine, teised aga eitavad seda. Küsimus tõuseb kõige selgemini esile, kui koondada tähelepanu tugeva tehisintellekti argumendi esimesele punktile: kõik osapooled nõustuvad, et kui (1) on tõene, siis on tugeva tehisintellekti kõige huvitavama vormiga lõpp, ja isegi Searle tunnistaks, et punkti (1) kummutamisest järelduks, et hiina toa argument ei pea paika. Nüüd võib teadvuse ja intentsionaalsuse vahelise seost vaadelda kui omaette küsimust, mis pole tehisintellekti vastase argumendi jaoks kriitiline.

Nii väldime olukorda, kus vastased vaidlevad vastu punktile (2), mitte hoolides punkti (1) kummutamisest. Näiteks võivad vastused, mis keskenduvad hiina toa ja keskkonna vahelisele seosele [Fodor 1980; Rey 1986], ja vastused, mis annavad protseduurilisi või funktsionaalseid selgitusi intentsionaalsuse kohta [Boden 1988; Thagard 1986], heita või mitte heita valgust intentsionaalsuse küsimusele, kuid ei tee kuidagi tõepärasemaks väidet, et hiina toal on teadvus. Seepärast jääb mulje, et probleemi, mille kõnealune stsenaarium tehisintellekti ette seab, pole vaadeldudki. Parimal juhul on kahtluse alla seatud vaid abistav eeldus, et intentsionaalsus eeldab teadvust.[[373]](#footnote-374))

Hiina toa argument seisneb järgmises mõttekäigus. Võtame suvalise programmi, mis suudab väidetavalt matkida mõnd teadvuse aspekti, näiteks hiina keele mõistmist või punase värvuse aistimist. Siis võib selle programmi implementeerida ükskeelne inglise keele kõneleja — nimetame teda edaspidi *deemoniks* — must-valges toas. Deemon järgib käsitsi kõiki reegleid, mida järgib programm, pidades paberilipikutel arvestust kõigi asjakohaste seesmiste seisundite ja muutujate üle, kustutades ja uuendades neid vastavalt vajadusele. Võib kujutleda, et deemoniga on ühenduses robotkeha, mis võtab tajuanduritelt vastu digitaalseid sisendandmeid, manipuleerib nendega vastavalt programmi spetsifikatsioonile ning saadab digitaalsed väljundandmed täituritele. Nii saame programmi täiusliku implementatsiooni. Sellest hoolimata tundub olevat ilmne, et deemon ei mõista teadvuslikult hiina keelt ja et tal ei ole punast elamust. Seepärast ei piisa programmi implementatsioonist nendeks teadvuselamusteks. Teadvuseks on tarvis midagi enamat kui sobiva programmi implementatsioon.

Tugeva tehisintellekti toetajad on tavaliselt öelnud vastuseks, et hiina keelt ei mõista deemon ise, vaid mõistmine ja teadvus tuleb omistada hoopis süsteemile, mille koosseisu kuuluvad nii deemon kui ka paberilipikud. Searle peab seda vastust ilmselgelt ​**[lk.324]**​ ebaveenvaks. On ju tõepoolest midagi intuitsioonivastast väites, et agendist ja paberitükkidest koosneval süsteemil on kollektiivne teadvus. Siinkohal satub poleemika ummikusse. Tehisintellekti pooldajad väidavad, et süsteemil on teadvus, vastased peavad seda järeldust naeruväärseks, ja tundub, et siit on raske edasi liikuda. Mina arvan siiski, et argumendid, mida ma olen juba esitanud, annavad võimaluse ummikust pääsemiseks ja tugeva tehisintellekti positsioonile jõudmiseks.

Võtame näidisprogrammiks kombinatoorse seisundiautomaadi, mis simuleerib parajasti punast ja mahlakat õuna vaatleva hiina keele kõneleja organisatsiooni neurontasemel. Toas viibiv deemon on selle automaadi implementatsioon; tal on iga lipiku kohta üks seisundivektori element ja ta uuendab lipikuid igal programmi​sammul vastavalt seisundisiirdereeglitele. Nüüd võime tuua mängu hääbuvate ja hüplevate kvaalide argumendid, konstrueerides juhtude spektri, mille ühes otspunktis on esialgne hiina keele kõneleja ja teises hiina tuba.[[374]](#footnote-375) Seda ei ole raske teha. Alustuseks võime asendada hiina keele kõneleja aju neuronid ükshaaval tillukeste deemonitega, millest igaüks dubleerib vastava neuroni sisend-väljundfunktsiooni.[[375]](#footnote-376) Naaberneuronitelt stiimuleid saades sooritab deemon vajalikud arvutused ning stimuleerib omakorda naaberneuroneid. Sedamööda kuidas üha uued ja uued neuronid asenduvad deemonitega, võtavad need juhtimise üle; lõpuks jõuame süsteemini, kus kolju on puupüsti täis miljardeid deemoneid, mis reageerivad üksteise signaalidele ja sensoorsetele sisendandmetele, sooritavad arvutusi ning saadavad omakorda signaale teistele deemonitele ja täituritele. (Kui keegi vaidleb vastu, et kolju ei suudaks iial kõiki neid deemoneid mahutada, siis võime muuta stsenaariumi ja paigutada raadiosideseadmed väljapoole koljut.)

Edasi hakkame järk-järgult vähendama deemonite arvu ja kahekordistama nende tööd. Kõigepealt asendame kaks naaberdeemonit ühe deemoniga, kes teeb ära mõlema deemoni töö. Uus deemon hakkab pidama arvestust mõlema neuroni siseseisundi kohta, mida ta peab simuleerima: ta võib pidada märkmeid kummagi kohta omaette paberilipikul ning uuendada mõlema lipiku sisu vastavalt naaberdeemonitelt saadud signaalidele ja teise lipiku sisule. Nii ühendame deemoneid aina edasi, nii et lõpuks jääb meile üksainus deemon ja mitu miljardit tibatillukest lipikut. Iga lipikut voib kujutleda paiknevana vastava neuroni algses asukohas, nii et deemon sibab ajus ringi, uuendades iga lipikut vastavalt naaberlipikute sisule ja vajadusel ka sensoorsetele sisendandmetele.

Kõigist neist muudatustest hoolimata ühtib sel viisil saadud lõppsüsteemi funktsionaalne organisatsioon esialgse aju omaga. Neuronite vahelisi põhjuslikke suhteid, nagu need kehtisid lähtejuhul, jäljendavad vahejuhtudel põhjuslikud suhted deemonite vahel ning lõppjuhul põhjuslikud suhted lipikute vahel. Sel viimasel juhul kujutavad põhjuslikke suhteid deemoni toimingud — lipik mõjutab deemoni seisundit ja see omakorda naaberlipikut —, kuid suhted on sellegipoolest põhjuslikud. ​**[lk.325]**​ Vaadeldes süsteemi talitlemist kiirendusega, näeme põhjuslike vastasmõjude keerist, mis vastab täpselt neuronkeerisele.

Seega võime nüüd lasta käiku hääbuvate kvaalide ja hüplevate kvaalide argumendid. Kui lõppsüsteemil teadvuselamused puuduvad, siis peab kusagil vahepeal leiduma hääbunud teadvuselamustega süsteem. See ei ole usutav põhjustel, mida me juba teame. Võime mõtteliselt vaadelda ka olukorda, kus neuronkontuuri dubleerib teine kontuur, mida implementeerivad deemonid (või üksainus deemon) ja paberilipikud ja kus saab sisse lülitada kord ühe, kord teise kontuuri. Nagu ennegi, viiks see kvaalide hüplemiseni ilma, et funktsionaalne organisatsioon muutuks, mistõttu süsteemil ei oleks mitte kuidagi võimalik erinevust märgata. Jällegi näib palju tõenäolisemana oletus, et kvaalid jäävad kogu aeg ühesugusteks.

Seega on mõistlik järeldada, et lõppsüsteemil on täpselt samasugused teadvuselamused nagu esialgsel süsteemil. Kui neuronsüsteemis sündisid erepunased elamused, siis sünnivad need ka deemonite süsteemis ja deemonist vahetalitajaga paberilipikute võrgus. Kuid see lõppjuht on ju vaid hiina toa süsteemi koopia. Seega oleme leidnud positiivse aluse uskumaks, et süsteemil on tõesti teadvuselamused, näiteks hiina keele mõistmise või punase värvuse aistimise elamus.

Selline vaatlusviis teeb selgeks kaks asja, mis hiina toa kirjelduses, nagu Searle seda esitab, märkamatuks jäävad. Esiteks ei ole “paberilipikud” hiina toas lihtsalt hunnik formaalseid sümboleid. Nad moodustavad konkreetse dünaamilise süsteemi, mille põhjuslik organisatsioon on otseses vastavuses esialgse aju organisatsiooniga. See asjaolu võib jääda märkamata sümbolitega manipuleerimisega meie ettekujutuses seostuva aeglase tempo ning neid manipulatsioone sooritava deemoni kohalolu tõttu, kuid sellegipoolest loob teadvuselamuse just lipikute vaheline konkreetne dünaamika. Teiseks on deemoni roll siin täiesti teisejärguline. Meid huvitav põhjuslik dünaamika leiab aset lähtejuhu neuronitele vastavate lipikute vahel, deemoni ülesanne piirdub põhjuslike seoste vahendamisega. Kujutluspilt kolju sees ringi vudivast deemonist teeb silmanähtavaks asjaolu, et süsteemi elamuste omistamine deemonile oleks tõsine tasemete segiajamine. Asjaolu, et deemon on teadvusega agent, võib ahvatleda arvama, et kui süsteemil üldse kusagil elamusi on, siis on need deemonis; tegelikult aga on deemoni teadvus süsteemi toimimise seisukohalt täiesti ebaoluline. Deemoni töö teeks ära lihtne otsingutabel. Süsteemi otsustavaks aspektiks on sümbolite vaheline dünaamika.

Searle'i väide avaldab meie intuitsioonile muljet programmi kummalise implementeerimisviisi tõttu, mis ähmastab relevantset põhjuslikku dünaamikat. Vaadates aga kaugemale kujutlustest, mida toob mõttesse irrelevantse deemoni olemasolu ja sümbolite ümberpaiknemise aeglane kiirus, näeme, et hiina toa põhjuslik dünaamika peegeldab täpselt põhjuslikku dünaamikat kolju sees. Ja oletus, et süsteemis sünnib elamus, ei tundu enam nii ebatõenäolisena.

​**[lk.326]**​ Ühes Searle’i argumendi versioonis jätab deemon arvutusreeglid meelde ja implementeerib programmi iseenda sees. Loomulikult ei suuda inimesed tegelikkuses meelde jätta sadatki reeglit ja sümbolit, rääkimata paljudest miljarditest, kuid supermälumooduliga deemonit, kes suudaks meelde jätta kõik reeglid ja kõigi sümbolite seisundid, võib ette kujutada küll. Sellisel juhul võime jällegi eeldada, et süsteemis sünnivad teadvuselamused, mis ei ole deemoni elamused. Searle väidab, et kui siin üldse kusagil elamusi on, siis peavad need olema deemonil, sest kogu töötlus toimub deemoni sees. Kuid õigem on vaadelda seda olukorda kui näidet kahe mentaalse süsteemi realiseeringut ühesainsas füüsilises ruumis. Organisatsioon, millest sünnivad hiina keele oskaja elamused, on hoopis midagi muud kui organisatsioon, millest sünnivad deemoni elamused. Hiina keele oskaja organisatsiooni moodustavad miljardite aadresside vahelised põhjuslikud suhted supermälumoodulis; ka siin tegutseb deemon jällegi vaid nii-öelda põhjuslikkuse abistajana. See saab ilmseks, kui vaadelda juhtude spektrit, kus koljus ringi sõeluv deemon jätab vähehaaval meelde reegleid ja sümboleid, nii et lõpuks teab ta neid kõiki. Vastav struktuur nihkub järk-järgult koljust deemoni supermällu, kuid elamus jääb kogu aeg ühesuguseks ja täiesti irdu deemoni elamustest.

Võidakse arvata, et kuna minu argument eeldab aju organisatsiooni dubleerimist neurontasemel, siis kinnitab ta vaid tugeva tehisintellekti nõrka erikuju, mis seostab intellekti kindlalt bioloogiaga. (Arutledes vastuväite üle, mida ta nimetab “ajusimulaatori” argumendiks, väljendab Searle üllatust, et vastuses, mille on andnud tehisintellekti pooldaja, figureerib inimese bioloogia üksikasjalik simulatsioon.) Niiviisi arvaja ei ole aga mõistnud selle argumendi jõudu. Ajusimulatsiooniprogramm on vaid jäämäe tipp. Niipea kui me teame, et ühesainsaski programmis võib sündida vaim isegi siis, kui implementeerida seda hiina toa kombel, kaotab Searle’i universaalsusele pretendeeriv argument kogu oma jõu: me teame, et hiina toa deemon ja lipikud võivad tõepoolest kanda sõltumatut vaimu. Ja siis avanevad lüüsid tervele programmide spektrile, mis võiksid kandideerida teadvuselamuse kandja kohale. Spektri piirid jäävad lahtiseks, kuid hiina tuba meile sel teel jalgu ei jää.

#### Süntaks ja semantika

Teine argument, mille on samuti esitanud Searle (1984), kõlab nii:

1. Arvutiprogramm on süntaktiline.
2. Süntaksist ei piisa semantikaks.
3. Vaim on semantiline.
4. Seepärast ei piisa ühegi programmi implementatsioonist vaimuks.

​**[lk.327]**​ Ka seda mõttekäiku esitatakse kui argumenti intentsionaalsuse kohta, kuid teda võib võtta ka kui argumenti teadvuse kohta. Searle'i jaoks on keskseks intentsionaalsuse liigiks fenomenoloogiline intentsionaalsus, nagu see on omane teadvusele.

Seda argumenti võib tõlgendada ja kritiseerida mitmel viisil, kuid tema peamine viga on, et ta ei hinda vääriliselt implementatsiooni otsustavat rolli. Programmid on abstraktsed informaatilised objektid ja nad on puhtalt süntaktilised. Kindlasti ei või ükski programm iseenesest pretendeerida sellele, et tal oleks vaim. Seevastu programmide implementatsioonid on konkreetsed, põhjusliku dünaamikaga süsteemid ja nad ei ole puhtalt süntaktilised objektid. Implementatsioonil on reaalses maailmas põhjuslik kaal, ja just tänu sellele kaalule sünnivad teadvus ja intentsionaalsus. Süntaktiline on programm; semantiline sisu on implementatsioonil.

Searle oleks võinud öelda, et mingis mõttes on isegi implementatsioonid süntaktilised, võib-olla näiteks sellepärast, et implementatsioonide dünaamika määravad nende formaalsed omadused. Ent niipea kui me tarvitame terminit “süntaks” niisuguses tähenduses, et implementatsioonid osutuvad süntaktilisteks, ei pea enam samas tähenduses paika, et süntaksist ei piisa semantikaks. Annaks ehk pidada tõenäoliseks, et abstraktsete sümbolite staatilistel kogumitel ei ole mingeid seesmisi semantilisi omadusi, ent palju vähem ilmne on, et formaalselt kirjeldatud põhjuslikud protsessid ei saa olla vaimu kandjad.

Seda Searle’i argumenti saaks parodeerida nii:

1. Retseptid on süntaktilised.
2. Süntaksist ei piisa mureduseks.
3. Küpsised on muredad.
4. Seepärast ei piisa küpsisteks retsepti implementatsioonist.

Asja niiviisi vaadates on viga kohe ilmne. Argument ei tee vahet retseptide vahel, mis on süntaktilised objektid, ja retseptide implementatsioonide vahel, mis on füüsilised süsteemid reaalses maailmas. Jällegi teeb kogu töö ära implementatsioonisuhe, mis seob abstraktset valda konkreetsega. Retsept määratleb kaudselt niisuguste füüsiliste süsteemide klassi, mida võib pidada tema implementatsioonideks, ja muredus jms. omadused kuuluvad nendele süsteemidele. Analoogiliselt vastab igale programmile füüsikaliste süsteemide klass, mida võib pidada selle programmi implementatsioonideks, ja nendes süsteemides võivadki sündida niisugused omadused, nagu seda on vaim.

#### Simulatsioon on vaid simulatsioon

Üks populaarseid vastulauseid tehisintellektile (nt. Searle 1980, Hamad 1989) ütleb, et fenomeni simulatsioon ei ole sama mis tema replikatsioon. Näiteks kui me simuleerime arvutuslikult seedimist, siis tegelikkuses mitte mingi toidu seedimist ei toimu. Simuleeritud orkaan ei ole tegelik orkaan: orkaani ​**[lk.328]**​ arvutisimulatsioon ei tee kedagi märjaks. Soojuse simuleerimisel ei teki reaalset soojust. Miks peaksime seega ootama, et vaimu simulatsiooni tagajärjel tekib tõeline vaim? Miks peaksime ootama, et just sel ühel juhul, erinevalt teistest, ei ole arvutuslik protsess mitte pelk simulatsioon, vaid päris asi?

Kahtlemata ei ole simulatsioon paljude omaduste puhul replikatsioon. Simuleeritud soojus ei ole tõeline soojus. On aga ka omadusi, mille puhul simulatsioon on tõepoolest replikatsioon. Näiteks on põhjusliku silmusega süsteemi simulatsioon põhjusliku silmusega süsteem. Seega on küsimus tegelikult järgmine: kuidas eristada tüüpe ​*X*​, millede simulatsioon on samuti ​*X*​, tüüpidest, millede puhul see nii ei ole?

Mina oletan, et vastus on järgmine: omaduse ​*X* simulatsioon on ise ​*X* parajasti siis, kui omadus olla ​*X* on *organisatsiooniline invariant*​. Organisatsioonilise invariandi definitsioon on samasugune nagu ennegi: üks või teine omadus on organisatsiooniline invariant, kui ta sõltub ainult teda kandva süsteemi funktsionaalsest organisatsioonist ja mitte ühestki muust asjaolust. Füüsikalise süsteemi informaatiline simulatsioon võib suuta reprodutseerida süsteemi abstraktset põhjuslikku organisatsiooni ja tagada, et see põhjuslik organisatsioon replitseerub igas implementatsioonis, olenemata implementatsiooni kandja koosseisust. Selline implementatsioon replitseerib kõik lähtesüsteemi organisatsioonilised invariandid, kuid süsteemi muud omadused lähevad kaduma.

Omadus olla orkaan ei ole organisatsiooniline invariant, kuna ta sõltub ka mitteorganisatsioonilisetest omadustest, näiteks kõnealust omadust kandva süsteemi kiirusest, kujust ja füüsilisest koosseisust. (Süsteem, kus samu põhjuslikke vastasmõjusid implementeeritaks väga aeglaselt suure hulga piljardipallidega, ei oleks orkaan.) Analoogiliselt sõltuvad seedimine ja soojus neid kandva süsteemi füüsilise koosseisu aspektidest, mis ei ole läbinisti organisatsioonilised. Seedimissüsteemis võib bioloogilised komponendid järk-järgult välja vahetada nii, et hapete ja aluste vaheliste reaktsioonide asemele astuvad põhjuslikult isomorfsed vastasmõjud metallitükkide vahel, ja tulemust ei saa enam pidada seedimiseks. Seega ei ole niisuguste omadustega süsteemi puhul mõtet oodata, et samad omadused oleksid ka selle süsteemi simulatsioonil.

Kuid fenomenaalsete omadustega on teine lugu. Fenomenaalsed omadused on, nagu ma [7. peatükis](#PARA.133-7) näidata püüdsin, organisatsioonilised invariandid. Ja kui see on nii, siis järeldub siit, et fenomenaalsete omadustega süsteemi õigel simulatsioonil on ka endal need fenomenaalsed omadused, kuna ta jäljendab kõigis üksikasjades lähtesüsteemi funktsionaalset organisatsiooni. Organisatsiooniline invariantsus teeb teadvuse põhimõtteliselt erinevaks teistest mainitud omadustest ja avab tee tugevale tehisintellektile.

### 5. Välised vastuväited

Tugeva tehisintellekti vastastest argumentidest on mind kõige rohkem huvitanud seesmised vastuväited, kuna need on käesoleva raamatu teema seisukohalt kõige olulisemad. Kuid ma vähemalt mainiksin ära ​**[lk.329]**​ ka mõned välised vastuväited. Nagu ma [ülal](#PARA.167-3) juba ütlesin, on tehisintellekti vastu esitatud väliste argumentide positsioon *prima facie* tugev[[376]](#footnote-377): on kõigiti põhjust uskuda, et füüsikaseadused on vähemalt meie praeguse arusaamise järgi arvutatavad ning et inimese käitumine tuleneb füüsikaseadustest. Kui see on nii, siis järeldub sellest, et arvutussüsteem võib simuleerida inimese käitumist. Siiski vaieldakse sellele aeg-ajalt vastu, mistõttu ma siiski annan neist vastuväidetest väikese ülevaate.

####  Reeglite järgimist vaatlevad vastuväited

Võib-olla kõige vanem kõigist välistest tehisintellekti vastastest argumentidest on see, mis ütleb, et kuna arvutussüsteemid järgivad alati ainult reegleid, siis ei saa nad iial omandada mõningaid inimlikke võimeid, näiteks loovust või paindlikkust. See on kõigist välistest vastuväidetest mitmes suhtes kõige nõrgem, muu hulgas oma udususe ja ebamäärasuse tõttu. Tõepoolest, sellele võib kergesti vastata, et neurontasemel võib inimese aju olla ju vägagi mehhaaniline ja reflektoorne, kuid see ei välista sugugi loovust ja paindlikkust makroskoopilisel tasemel. Loomulikult ei tarvitse oponent nõustuda pidama neurontaset mehhaaniliseks, kuid nii või teisiti ei paista leiduvat ühtki head argumenti teesi kasuks, et informaatiline dünaamika, mis toimib põhjuslikul alustasemel, on ühitamatu loovuse ja paindlikkusega makroskoopilisel tasemel.

Seesuguseid vastuväiteid soodustab mingil määral võib-olla arvutussüsteemide vaikiv samastamine *sümboolsete arvutussüsteemidega* — süsteemidega, mis sooritavad sümboolseid manipulatsioone kontseptuaalsete representatsioonidega kõrgtasemel, äärmuslikul juhul koguni süsteemidega, mis tuletavad eeldustest järeldusi jäigalt esimest järku loogika järgi​. Nendel juhtudel on sel vastuväitel võib-olla mingil määral alustki, ehkki isegi selle üle võib vaielda. Kuid nii või teisiti on arvutussüsteemide klass palju mahukam. Näiteks ajusimulatsioon madaltasemel on samuti arvutus, kuid mitte seesugune sümboolne arvutus. Ka mingil vahepealsel tasemel on kognitiivteaduste konnektsionistlikud mudelid püüdnud rakendada arvutusi, mille sisuks ei ole sümbolitega manipuleerimine. Sellistel juhtudel võib süsteem mingil tasemel reegleid järgidagi, ilma et see kajastuks otseselt käitumise tasemel; ja konnektsionistid väidavadki sageli, et nende meetod võimaldab näidata, kuidas madaltaseme mehhaanilisusest kasvab välja kõrgtaseme paindlikkus. Või nagu Hofstadter (1979) on öelnud: tase, millel ma mõtlen, ei pruugi olla tase, millel ma *sum*​[[377]](#footnote-378).[[378]](#footnote-379).

#### Gödeli teoreemile apelleerivad vastuväited

Mõnikord kinnitatakse, et Gödeli teoreem näitab, et arvutussüsteemide võimalustel on piirid, mida inimestel ei ole. Gödeli teoreem ütleb meile, et igas mittevastuolulises formaalses süsteemis, mis on piisavalt võimas, et hõlmata teatavaid aritmeetikatehteid, leidub tõene lause — süsteemi *Gödeli-lause* — mida see süsteem ei suuda tõestada. Ja argument ütleb, et kuna meie mõistame, et Gödeli-lause on tõene, ​**[lk.330]**​ siis on meil mingi võime, mida formaalsel süsteemil ei ole. Sellest järeldub, et mitte ükski süsteem ei suuda täpselt formaliseerida inimlikke võimeid. (Seesuguseid argumente on esitanud muuhulgas Lucas [1961] ja Penrose [1989, 1994]).

Lühike vastus neile argumentidele on, et pole mingit põhjust uskuda, et inimesedki suudaksid asjaomaste Gödeli-lausete tõesust mõista. Parimal juhul näeme, et kui süsteem on vastuolutu, siis on tema Gödeli-lause tõene; kuid pole põhjust uskuda, et me suudame tõestada suvalise formaalse süsteemi vastuolutust.[[379]](#footnote-380) [[380]](#footnote-381) Eriti kehtib see keeruliste formaalsete süsteemide, näiteks inimese aju väljundit simuleeriva süsteemi puhul: võib väga hästi selguda, et sellise süsteemi vastuolutuse kindlakstegemine ületab meie võimete piirid. Seega võib olla, et igaüht meist saab simuleerida keeruka formaalse süsteemiga, mille vastuolulisust või vastuolutust me ei suuda näidata[[381]](#footnote-382). Kui see on nii, siis ei ole meiegi võimelised omaendi Gödeli-lausete tõeväärtust hindama.

Sel göödellikul argumendil on palju variante, millele on omakorda esitatud vastuargumente ja möödapääsuteid. Ma ei hakka neid siinkohal arutama. (Käsitlesid pikemalt artiklis Chalmers 1995c). Need küsimused toovad palju huvitavaid ja virgutavaid mõtteid, kuid minu arvates võib õigusega öelda, et väidet, nagu göödellikud kitsendused inimeste puhul ei kehtiks, pole mitte keegi suutnud veenvalt põhjendada.

####  Mittearvutatavusele ja mittevastuolulisusele apelleerivad vastuväited

Ülaltoodud vastuväited olid “kõrgtasemel” argumendid kognitiivse talitlemise mittearvutatavuse kasuks. Kuid tehisintellekti positsiooni võidakse rünnata ka madaltasemel, väites, et füüsiline talitlemine ei ole arvutatav. Penrose (1994) väidab, et mittearvutatav element võib sisalduda näiteks korrektses kvantgravitatsiooniteoorias. Ainsa kinnituse sellele järeldusele leiab ta aga mainitud göödellikus argumendis. Teoreetiline füüsika ise ei paku sellisele järeldusele mingit tuge ja see tähendab, et kui göödellik argument kummutada, siis kaob igasugune põhjus mittearvutatavatesse füüsikaseadustesse uskumiseks. Tõepoolest, võidakse väita, et kui igal aju elemendil, näiteks neuronil, on ainult lõplik hulk relevantseid seisundeid ja kui ka relevantseid elemente on ainult lõplik hulk, siis peab aju relevantne põhjuslik struktuur olema väljendatav informaatilise kirjeldusega.

See toob meid viimase vastuväite juurde, mille järgi ajuprotsessid on võib-olla mingis olemuslikus mõttes pidevad, samas kui arvutuslikud protsessid on diskreetsed, ning et see pidevus võib olla meie kognitiivsele võimekusele oluliselt omane, missugusel juhul seda võimekust ei suuda matkida ükski diskreetne simulatsioon. Võib-olla läheb neuroni lähendamisel süsteemiga, millel on ainult lõplik hulk seisundeid, kaduma midagi tema talitluseks eluliselt tähtsat. Oponent võib viidata näiteks teatavate mittelineaarsete süsteemide “tundlikule sõltuvusele algtingimustest”, millest tuleneb, et isegi väike ümardusviga ühes ​**[lk.331]**​ töötluse etapis võib viia ulatuslike makroskoopiliste erinevusteni mõnes hilisemas etapis. Kui ajuprotsessid on tõepoolest niisugused, siis annab suvaline diskreetne ajusimulatsioon teistsuguseid tulemusi kui tegelikud, pidevad protsessid.

Siiski on põhjust arvata, et pidevus​[[382]](#footnote-383) ei saa meie kognitiivse pädevuse seisukohalt nii oluline olla. Taustmüra olemasolu bioloogilistes süsteemides tähendab, et mitte ükski protsess ei saa nõuda täpsusmäära, mis ületab mingi piiri. Mingist piirist (milleks võib olla täpsus 10-10 mingil sobival skaalal) edasi minnes teevad juhusliku taustmüra vätimatud fluktuatsioonid igasuguse suurema täpsuse ülearuseks. See tähendab, et kui me lähendame süsteemi seisundit sellele täpsustasemele (või kindluse mõttes veidi kaugemale, näiteks 10-20-ni), siis annab meie simulatsioon niisama häid tulemusi, nagu suudab usaldusväärselt anda süsteem ise. On tõsi, et mittelineaarsete efektide tõttu võib see lähendus viia teistsuguse käitumiseni kui see, mida produtseerib vastaval juhul algne süsteem ise; ent kui bioloogiline müra oleks olnud veidi teistsugune, siis oleks ka süsteem ise võinud produtseerida ennustatud käitumise. Soovi korral võime modelleerida isegi müraprotsessi ennast.[[383]](#footnote-384) Tulemuseks saame simulatsiooni, mis ühtib käitumuslikelt võimetelt lähtesüsteemiga, kuigi tema tegelik käitumine võib mõnel spetsiifilisel juhul olla erinev. Moraal on selles, et kui jutt on meie kognitiivsete võimete dubleerimisest, siis pole hea lähendus halvem kui päris asi.

On tõsi, et piiramatu täpsusega süsteemil võib olla kognitiivseid võimeid, mida ei saaks iial olla ühelgi diskreetsel süsteemil. Näiteks võiks täiuslik pidev süsteem lahendada peatumisprobleemi — ülesande, mida ei suuda lahendada ükski diskreetne süsteem — reaalarvu abil, mille *n*​. kahendnumber on 1 parajasti siis, kui Turingi masin number ​*n* suvaliste sisendandmete korral peatub.[[384]](#footnote-385) Kuid müra olemasolust tuleneb, et niisugust süsteemi ei suudaks usaldusväärselt implementeerida ükski bioloogiline protsess. Täpsusel, millele võivad loota bioloogilised süsteemid, on lõplikud piirid, ja see tähendab, et diskreetsetel süsteemidel ei ole võimaluste piirid ahtamad kui inimeste ja loomade ajudel.

### 6. Kokkuvõte

Kõigest sellest võib järeldada, et tehisintellekti ambitsioonide teel ei näi olevat põhimõttelisi takistusi. Välistel vastuväidetel ei tundu olevat kuigi palju jõudu. Seesmised vastuväited teevad võib-olla rohkem muret, kuid ükski neid vastuväiteid toetavatest argumentidest ei tundu lähemal analüüsimisel veenvana; ja tõepoolest, kui minu eelmistes peatükkides esitatud argumendid peavad paika, siis on meil tõsist positiivset põhjust uskuda, et sobiva arvutuse implementatsiooniga tuleb kaasa teadvuselamus. Seega on masinteadvuse väljavaated head — vähemasti põhimõtteliselt, kuigi mitte veel praktikas.

​**[lk.332]**​ Ma pole öelnud kuigi palju selle kohta, missugune on see arvutus, millest võiks piisata teadvuselamuseks. Enamikus argumentides olen ma näitena kasutanud aju neuronhaavalist simulatsiooni, kuid tõenäoliselt leidub ka palju teisi teadvuseks piisavaid arvutuste liike. Näiteks võib olla, et arvutussüsteem, mis peegeldab aju põhjuslikku organisatsiooni palju jämedamal tasemel, suudab siiski simuleerida momente, mis on teadvuselamuse sündimiseks olulised. Ja on tõenäoline, et implementeerimisel võivad külluslikke teadvuselamusi anda ka mingid hoopis teised arvutused, mis vastavad hoopis teistsugusele põhjuslikule organisatsioonile.

See pilt on ühevõrra kooskõlas nii sümboolse kui ka konnektsionistliku vaateviisiga kognitsioonile ning teistegi informaatiliste vaateviisidega. Tõepoolest võikski öelda, et arvutuste keskne roll tunnetuse uurimisel tuleb sellest, et informaatiliste konstruktsioonidega saab simuleerida peaaegu ükskõik missugust põhjuslikku organisatsiooni. Me võime näha informaatilistes formalismides ideaalset vahendit põhjusliku organisatsiooni mustrite väljendamiseks ja, veel enamgi, ühenduses implementatsioonimeetoditega ka ideaalset vahendit nende replitseerimiseks. Ükskõik, missugune põhjuslik organisatsioon tunnetuse ja teadvuse mõttes ka määravaks osutub, me võime eeldada, et informaatiline käsitlus suudab seda simuleerida. Oleks võimalik isegi väita, et arvutussüsteemide sageli mainitud universaalsuse taga peitubki just seesugune paindlikkus. Tehisintellekti pooldajad ei ole sidunud endid ühegi arvutuste liigiga kui vaimuks piisavaga; tehisintellekti tees mõjub nii veenvalt just tänu arvutussüsteemide klassi mahukusele.[[385]](#footnote-386) [[386]](#footnote-387).

Küsimus, missugune arvutuste klass on inimese vaimu replitseerimiseks piisav, jääb seega lahtiseks, kuid meil on tõsist põhjust arvata, et see klass ei ole tühi.

## ​**[lk.333]**​ 10. peatükk. Kvantmehhaanika interpreteerimine

### 1. Kaks müsteeriumi

Kvantmehhaanika probleem on peaaegu sama raske nagu teadvuse probleem. Kvantmehhaanika annab meile hämmastavalt täpse meetodi empiiriliste vaatluste tulemuste ennustamiseks, kuid pildist, mille ta meile maailmast annab, on erakordselt raske aru saada. Kuidas saab meie maailm olla niisugune, nagu ta peab olema, et kvantmehhaanika ennustused osutuksid õigeteks? Selles küsimuses ei ole üksmeelele ligilähedalegi jõutud. Sageli näib, et kvantmehhaanika probleemile, nagu ka teadvuse probleemile, ei olegi võimalik rahuldavat lahendust leida.

Paljud autorid on arvanud, et nende kahe kõige mõistatuslikuma probleemi vahel võib olla sügav seos (nt. Bohm 1980; Hodg­son 1988; Lock­wood 1989; Pen­rose 1989; Squires 1990; Stapp 1993; Wig­ner 1961). Kui müsteeriume on kaks, siis on ahvatlev oletada, et neil on ühine allikas. Ahvatlust suurendab veelgi asjaolu, et kvantmehhaanika probleemides on tähtsal kohal vaatluse mõiste, mille juurde kuulub kindlalt suhe mingi subjekti elamuse ja ülejäänud maailma vahel.

Kõige sagedamini on oletatud, et kvantmehhaanikas peitubki teadvuse füüsikalise seletuse võti. Kuid nagu me oleme näinud​, ei ole sel projektil iial lootust eesmärgile jõuda. Viimselt tumestab teadvuse “kvantteooriaid” samasugune seletuslünk nagu klassikalisi teooriaidki. Nii või teisiti tuleb elamust vaadelda kui midagi, mis väljub maailma füüsikaliste omaduste piiridest. Võib-olla on kvantmehhaanikal mingi roll keha–vaimu seose kirjeldamisel, kuid kvantteooria üksi ei saa meile öelda, miks teadvus eksisteerib.

​**[lk.334]**​ Kuid nende kahe probleemi vaheline seos võib olla ka keerulisem. Kuigi kvantmehhaanika teadvust ei seleta, pole võimatu, et teadvuseteooria suudab heita valgust kvantmehhaanika probleemidele. On ju väga paljud autorid nõus, et mõlemal probleemidel on mingi kokkupuude vaatlemise ja elamusega. Loomulik on oletada, et elamuse teooria võib aidata meil neis probleemides selgusele jõuda. On kõneldud teadvuse aktiivsest rollist kvantteoorias — näiteks on oletatud, et teadvus põhjustabki lainefunktsiooni kollapsi —, kuid mina kavatsen näidata, et teadvuse roll neis asjus ei ole nii otsene. Eelkõige kavatsen ma näidata, et kvantteooria probleemid saab ümber mõtestada maailma füüsikalise struktuuri ja meie maailmaelamuse vahelise seose probleemideks ning et seega võib sobiv teadvuseteooria pakkuda tuge ühele kvantmehhaanika mitteortodokssele interpretatsioonile.

### 2. Kvantmehhaanika tugisambad

Kvantmehhaanika peamiseks tugisambaks on matemaatiline meetod eksperimentaalsete mõõtmistulemuste ennustamiseks. Kirjeldan siin üht selle meetodi varianti, jättes kõrvale mitmed tehnilised üksikasjad, et hõlmata olulisi tunnusjooni võimalikult lihtsalt. Käesolevas jaos kirjeldan ma seda mudelit ainult kui meetodit empiiriliste prognooside leidmiseks, jättes lahtiseks küsimuse, kas see kirjeldab otseselt füüsikalist reaalsust. Sügavamad interpretatsioonilised probleemid tulevad vaatluse alla [järgmises jaos](#ORG.10.3).

Klassikalises mudelis käib füüsikalise süsteemi seisundi kirjeldamine väga lihtsalt. Näiteks osakese seisundi kirjeldamiseks tuleb täpselt öelda, missugused väärtused on igal omadusel mingist omaduste hulgast, näiteks asukohal ja impulsil. Sellist lihtsat väärtust võib nimetada *baasväärtuseks*​[[387]](#footnote-388). Kvantmudelis ei ole asi nii lihtne. Siin tuleb süsteemi seisundit üldjuhul avaldada *lainefunktsiooni*​ ehk *seisundivektori* kujul. Vastavaid omadusi ei saa enam väljendada lihtsate väärtustena, vaid baasväärtuste teatavat liiki kombinatsioonina. Kvantseisundit võib vaadelda kui lihtsamate seisundite superpositsiooni.

Sellise omaduse kõige lihtsama näitena võib mainida spinni — suurust, millel on ainult kaks baasväärtust[[388]](#footnote-389). Olgu nende nimed “üles” ja “alla”. Kvantmehhaanikas ei ole osakese spinn aga mitte alati lihtsalt “üles” või “alla”, vaid üldjuhul tuleb teda väljendada väärtuste “üles” ja “alla” ühendina, kus kummalgi on erinev kompleksarvuline kaal. Seega on kõige parem vaadelda osakese spinni vektorina kahemõõtmelises vektorruumis. Kõige loomulikum on kujutleda seda suurust “üles”- ja “alla”-spinni superpositsioonina, kus kummalgi on oma kaal.

​**[lk.335]**​ Sama lugu on asukoha ja impulsiga, ainult et nendel suurustel on baasväärtusi mõlemal lõpmata palju. Klassikalisel osakesel võivad nii asukoht kui ka impulss omandada suvalisi väärtusi lõpmatust kontiinumist. Kvantosakese asukohta väljendab seega lõpmatumõõtmeline vektor, kus igal neist asukohtadest on erinev kaal. Kõige parem on vaadelda seda vektorit lainena, millel on igas ruumipunktis kindel amplituud; funktsioon ruumipunktide hulgast amplituudide hulka ongi *lainefunktsioon*​. Ka kvantosakese impulssi võib vaadelda kui lainet, kus igale impulsi baasväärtusele vastab mingi amplituud. Asukohta või impulssi võib jällegi kujutada vastavate baasväärtuste superpositsioonina.

Kuna need seisundid on lihtsalt vektorid, siis saab neid väga mitmel viisil komponentideks lahutada. Kuigi kahemõõtmelist spinnivektorit on sageli kasulik vaadelda kui “üles”- ja “alla”-komponendi summat, võib teda lahutada ka väga mitmel muul viisil, olenevalt vektorruumi baasi valikust. Kõik need baasid on võrdselt “looduslikud”; loodus ei eelista ühtki neist. Täpsemini öeldes kujutab nii osakese asukohta kui ka tema impulssi üksainus vektor. Lahutades seda vektorit ühe baasi järgi, saame asukoha amplituudid, teise järgi impulsi amplituudid. Üldjuhul sõltub ühele või teisele juhule sobiv lahutus suurusest, mis meid huvitab, ja eriti, nagu peagi jutuks tuleb, suurusest, mida me soovime mõõta.

Rohkem kui ühest osakestest koosnevate süsteemide seisunditega on asi mõnevõrra keerulisem, kuid põhimõte jääb samaks. Võtame süsteemi, mis koosneb kahest osakesest ​*A* ja ​*B*​. Üldiselt ei saa niisuguse süsteemi seisundit väljendada osakestele ​*A* ja ​*B* vastavate lainefunktsioonide lihtsa kombinatsioonina; kahe osakese seisundid on sageli mittelahutuvad[[389]](#footnote-390). Seepärast tuleb süsteemi seisundit väljendada lainefunktsioonina keerukamas ruumis. Ent seda funktsiooni võib siiski vaadelda kahe osakesega süsteemi lihtsamate seisundite superpositsioonina, nii et üldine pilt jääb ka siin kehtima. Nii on see ka veelgi keerukamate süsteemide puhul: ka siin on seisundeid kõige parem kujutada seisundite superpositsioonidele vastavate lainefunktsioonidena.

Kõik see on intuitsioonivastane, kuid mitte veel paradoksaalne. Võttes seda formalismi täht-tähelt kui reaalsuse tõsikindlat kirjeldust, ei ole tema mõtestamise raskused liiga suured. Mõni autor leiab, et kuna see formalism eeldab, et maailma objektidel kindlakujuline objektiivne seisund puudub, siis on ta vastuolus “objektiivse” vaatega maailmale. Kuid see ei ole õige järeldus. Kõnealuse pildi järgi sobib lainefunktsioon objekti seisundi väljendamiseks paremini kui diskreetsed suurused, kuid ta on täiesti selgepiiriline. See pilt ütleb meile lihtsalt seda, et reaalsuse alustase on lainesarnane. See nõuab uutviisi mõtlemist, kuid sellega võib harjuda. Jääb ju mikroskoopilise reaalsuse alustase​**[lk.336]**​ väga kaugele makroskoopilisest tasemest, millega me tavaliselt kokku puutume, ja seepärast polegi nii üllatav, et tal on harjumatuid omadusi. Raskusi valmistavad kvantmehhaanika teised iseärasused.

Kvantmehhaanika tuuma moodustavad kaks printsiipi, mis määravad lainefunktsiooni dünaamika: *Schrödingeri võrrand* ja *mõõtmispostulaat*​. Üheskoos determineerivad need kaks väga erinevat printsiipi süsteemi lainefunktsiooni muutumise ajas.

Suurema osa kvantmehhaanika sisust võtab kokku Schrödingeri võrrand — diferentsiaalvõrrand, mis annab süsteemi lainefunktsiooni muutumise kirjelduse (peaaegu) igas olukorras. Võrrandi täpne kuju ei ole meie jaoks oluline. Meie jaoks on kõige tähtsam, et see diferentsiaalvõrrand on *lineaarne*​: kui meil on kaks seisundit *A* ja *B*​, mis muutuvad nii, et ​*A*​-st saab ​*A′* ja ​*B*​-st ​*B′*​, siis saab seisundist, mis on seisundite ​*A* ja ​*B* superpositsioon, seisundite ​*A′* ja ​*B′* superpositsioon. Samuti tasub märkida, et Schrödingeri võrrandi dünaamika iseloomu tõttu muutuvad suhteliselt diskreetsed seisundid aja jooksul enamasti laialivalguvamaks. Seisund, mis on algul mingisse kitsasse vahemikku kuuluvate väärtuste superpositsioon, muutub üldjuhul pikapeale superpositsiooniks, mida moodustavate väärtuste vahemik on palju suurem. Ja lisaks kõigele muule on Schrödingeri võrrand täiesti deterministlik.

Schrödingeri võrrand on suhteliselt lihtne ja kergesti mõistetav. Just siin peitub kvantteooria tegelik sisu​[[390]](#footnote-391). Kvantteooria rakendamisel nii praktilistele kui ka eksperimentaalsele probleemidele moodustab suurema osa tööst mitmesuguste seisundite ajalise muutumise Schrödingeri dünaamika arvutamine.

Schrödingeri võrrand ei saa aga siiski olla viimane sõna. Sellest võrrandist tuleneb, et valdav enamik füüsikalistest seisunditest muutub aja jooksul peagi superpositsiooniks, mille osaseisundid paiknevad väga laias vahemikus. Kuid selline pilt ei lähe kokku meie tähelepanekutega maailma kohta. Osakese asendi mõõtmine annab meile ühe kindla väärtuse, mitte väärtuste superpositsiooni, nagu ennustaks Schrödingeri võrrand. Kui Schrödingeri võrrand ammendakski kogu kvantdünaamika, siis jõuaks maailm isegi makroskoopilisel tasemel hullupööra superpositsioonilisse seisundisse. Ent kogemus ütleb meile, et see ei ole nii. Osutitel on kindlad asukohad, liikuvatel objektidel on kindel mõõdetav impulss jne. Seega peab siin olema veel midagi, mis viiks meid selle võrrandi juurest diskreetsete sündmusteni, millest kõneleb meie kogemus.

Kvantfüüsika kanoonilise formalismi teiseks osaks on *mõõtmispostulaat*​ (ehk *kollapsi-*​ või *projektsioonipostulaat*​). See ütleb, et on olukordi, kus Schrödingeri dünaamikat rakendada ei saa. Täpsemalt öeldes: mõõtmise hetkel kollabeerub lainefunktsioon selgekujulisemaks. Kollabeerumise viis oleneb mõõdetavast omadusest. Näiteks kui me mõõdame osakese spinni ​**[lk.337]**​, siis kollabeerub see, isegi kui osakese seisund oli enne seda superpositsiooniline, seisundiks, kus spinn on kas “üles” või “alla”. Osakese asukoha mõõtmisel kollabeerub lainefunktsioon seisundiks, kus osakesel on täiesti kindel asukoht.[[391]](#footnote-392) Lõppseisund vastab endiselt mingile lainefunktsioonile, ent niisugusele, kus kogu amplituud koondub ühte kindlasse punkti; kõikjal mujal on amplituud null. Igale suurusele, mida me mõõta võiksime, vastab mingi operaator, ja mõõtmise hetkel kollabeerub seisund selle operaatori omaseisundiks. Operaatori omaseisund on alati seisund, kus vastaval mõõdetaval suurusel on kindel väärtus. Sellest järeldub, et mingit suurust mõõtes on tulemuseks alati ühene väärtus, täpselt nagu kogemus meile ütlebki.

Kollapsi dünaamika ei ole deterministlik, vaid tõenäosuslik. Kui osakese seisund on mingite asukohtade superpositsioon, siis me teame, et kui tema asukohta mõõta, kollabeerub ta mingisse seisundisse, kus tema asukoht on ühene, kuid seda asukohta ennast me ette ei tea. Mõõtmispostulaat annab iga võimaliku seisundi kohta vaid tõenäosuse, et süsteem kollabeerub sellesse seisundisse. Iga seisundi puhul väljendab seda tõenäosust[[392]](#footnote-393) lainefunktsiooni amplituudi ruut sellele seisundile vastavas täpses ruumipunktis. Näiteks kui osakese spinni moodustab superpositsioon, mille komponentideks on “üles”-spinn amplituudiga ${1}/{2}$ ja “alla”-spinn amplituudiga ${\sqrt{3}}/{2}$, siis kollabeerub spinn mõõtmisel tõenäosusega ${1}/{4}$ “üles”-spinniks ja tõenäosusega ${3}/{4}$ “alla”-spinniks. Lainefunktsiooni amplituudid on alati niisugused, et vastavate tõenäosuste summa on $1$.

### 3. Kvantmehhaanika interpreteerimine

Ühtekokku annavad need kaks printsiipi erakordselt võimsa matemaatilise meetodi eksperimentaalsete mõõtmistulemuste ennustamiseks. Selleks koostame süsteemi seisundit väljendava lainefunktsiooni ja arvutame Schrödingeri võrrandi abil välja tema ajalise muutumise enne mõõtmishetke. Mõõtmishetkel leiame lainefunktsiooni selleks hetkeks arvutatud amplituudide põhjal ühte või teise seisundisse kollabeerumise tõenäosuse ning tõenäosuse, et mõõtmine annab nii- või teistsuguse tulemuse. Teooria ennustused on leidnud nii vääramatult katselist kinnitust, et teaduslikke teooriaid, mis ennustamisega sama edukalt toime tulevad, on väga vähe. Matemaatilise meetodina on see teooria peaaegu kuulikindel.

Keeruliseks läheb asi siis, kui hakata küsima, kuidas see meetod töötada saab. Missugused reaalse maailma omadused annavad ​**[lk.338]**​ talle nii suure ennustustäpsuse​?[[393]](#footnote-394) See on kvantmehhaanika interpretatsiooni probleem. Selle käsitlemiseks on palju võimalusi, millest ükski ei ole päris veatu.

#### Esimene variant: mõista arvutusmeetodit täht-tähelt

Esimene loomulik reaktsioon oleks võtta kvantmehhaanika formalismi puhta kullana, nagu seda teaduses enamiku teooriatega tehaksegi. Meetodi aluseks on Schrödingeri võrrandi abil[[394]](#footnote-395) leitav lainefunktsioon ning mõõtmispostulaat, ja kuna meetod töötab, siis peaksime eeldama, et ta annab meile otseselt pildi sündmustest maailmas. See tähendab, et süsteemi reaalseks seisundiks ongi laineseisund, mida väljendab lainefunktsioon ja mis muutub ajas vastavalt dünaamikale, mida väljendavad kvantmehhaanika kaks alusprintsiipi. Enamasti muutub seisund ajas vastavalt Schrödingeri võrrandile, mõõtmishetkel aga vastavalt mõõtmispostulaadile. Selle pildi järgi koosneb maailm lainetest, mis muutuvad ajas enamasti lineaarselt ja superpositsioonis, ent aeg-ajalt, siis, kui toimub mõõtmine, kollabeeruvad kindlapiirilisemateks seisunditeks.

Kuid seda pilti ei ole kerge mõista. Kõigi raskuste allikaks on mõõtmispostulaat, mille järgi kollaps toimub mõõtmise hetkel. Kuid mida pidada mõõtmiseks? Kuidas loodus teab, millal on toimunud mõõtmine? “Mõõtmine” ei ole kindlasti mingi loodusseaduste alustermin; et teha mõõtmispostulaadist midagi, mida annaks kuidagiviisigi fundamentaalseaduseks pidada, tuleb mõõtmise mõiste asemele võtta midagi selgemat ja sügavamat. Kui lainefunktsiooni kollaps on maailmas objektiivselt eksisteeriv protsess, siis on meil tarvis selgeid, objektiivseid kriteeriume, mis lubaksid selle protsessi toimumist tuvastada.

Üks ilmselt ekslik lahendus oleks öelda, et kollaps toimub alati, kui kvantsüsteem vastasmõjustub mõõteriistadega. Probleem on selles, et mõõteriista mõiste mõjuks fundamentaalseaduse formuleeringus sama väheveenvalt kui mõõtmise mõiste. Enne oli meil tarvis kriteeriume, mis ütleksid, mida lugeda mõõtmiseks; nüüd on tarvis kriteeriume, mis ütleksid, mida lugeda mõõteriistaks.

Kvantmehhaanika algaegadel kõneldi sageli, et mõõteseade on klassikaline süsteem ning et mõõtmine toimub iga kord, kui kvantsüsteem vastasmõjustub klassikalise süsteemiga. Kuid see on selgelt ebarahuldav vastus. Kvantteooria tahab olla universaalne teooria ja peaks olema mõõteriistas toimuvatele protsessidele samavõrd kohaldatav nagu kõigile teistelegi protsessidele. Kui me ei soovi tunnistada, et maailmas on kaht põhimõtteliselt erinevat liiki füüsikalisi objekte — aga see eeldaks täiesti uue teooria loomist —, siis ei sobi termin “klassikaline süsteem” fundamentaalsetesse loodusseadustesse paremini kui “mõõtmine”.

​**[lk.339]**​ Ühe analoogilise idee järgi toimub mõõtmine iga kord, kui kvantsüsteem vastasmõjustub makroskoopilise süsteemiga. Ent ka “makroskoopiline” ei ole ilmselgelt mõiste, mis sobiks figureerima fundamentaalseadustes. Ta tuleks asendada millegi täpsemaga, nagu näiteks “süsteem, mille mass on üks gramm või rohkem”. Ent fundamentaalseadus, mis sisaldaks midagi sellesarnast, jätaks erakordselt meelevaldse mulje.

Kollapsi määratlemiseks ei tundu leiduvat ainsatki vähegi vastuvõetavat füüsikalist kriteeriumi. Mikrotaseme kriteeriumid — näiteks et kollaps toimub siis, kui süsteem vastasmõjustub prootoniga — võib katseandmete põhjal välistada. Alternatiiv oleks, et kriteerium viitaks mingile füüsikalisele omadusele kõrgtasemel, nii et kollaps toimuks siis, kui süsteemil on mingi kindel kõrgtasemeline konfiguratsioon. Kuid iga säärane kõrgtaseme omadus oleks meelevaldne, ja seniajani pole ühtki head kandidaati ka välja pakutud. Ning oletus, et mikroskoopiliste süsteemide Schrödingeri dünaamika kaotab korraga kehtivuse, niipea kui need süsteemid satuvad mingitesse erilistesse konfiguratsioonidesse, oleks samuti kuidagi väga kummaline.

Siiani on välja pakutud ainult üks kriteerium, millest vähegi midagi loota on: et mõõtmine toimub siis, kui kvantsüsteem mõjutab mõne olendi teadvust. See ei ole vähemalt nii ähmane ja meelevaldne kriteerium kui eelmised.[[395]](#footnote-396) Matemaatilise formalismi interpretatsioon, mis sellele kriteeriumile vastab, on vormilt kaunis elegantne ja lihtne, ja ta on täht-tähelistest interpretatsioonidest ainus, mis on vähegi laiemalt levinud. Selle interpretatsiooniga tulid esimesena välja London ja Bauer (1939), kuid kõige sagedamini seostatakse teda Wigneriga (1961).

Paneme tähele, et see interpretatsioon eeldab keha–vaimu dualismi. Kui teadvus oleks lihtsalt üks füüsikaline omadus teiste seas, siis seisaksid meie ees jälle kõik ülalmainitud raskused: meie positsioon osutuks jällegi üheks teisendiks “kõrgtaseme omaduse” positsioonist, mille järgi füüsikalise süsteemi lainefunktsiooni kollaps toimub mingite keerukate füüsiliste konfiguratsioonide kontekstis. Ent kui dualism on õige, siis võib sellest saada kollapsile tõeliselt fundamentaalne kriteerium. Ka lubaks asjaolu, et kollapsi põhjus oleks füüsikaliste protsesside suhtes väline, teooriat oluliselt lihtsustada. Kõiki puhtalt füüsikalisi süsteeme reguleeriks Schrödingeri dünaamika üksi; mõõtmiste dünaamikal, mis on hoopis teistsugune, oleks sõltumatu allikas[[396]](#footnote-397).

See interpretatsioon viib aga intuitsioonivastastele järeldustele. Võtame mingi mõõteseadme, näiteks elektroni seisundit mõõtva seadme osuti. Olgu selle seisund algselt superpositsiooniline. Kui läheduses ei ole teadvust, siis määrab kogu süsteemi käitumise lineaarne Schrödingeri dünaamika: kuna elektroni erinevad diskreetsed seisundid viiksid osuti erinevatesse diskreetsetesse seisunditesse, siis järelikult viib elektroni superpositsiooniline seisund ka osuti superpositsioonilisse seisundisse. Teisiti öeldes, teooria ennustab, et osuti näitab paljudesse suundadesse korraga! ​**[lk.340]**​ Ühte kindlasse suunda hakkab ta näitama alles siis, kui ma teda vaatan.

Veelgi kummalisematele järeldustele viib Schrödingeri kassi stsenaarium. Kinnisesse kasti pannakse kass, ja elektroni spinni mõõdetakse mingi mõõteseadmega, mis on seadistatud nii, et ta surmab kassi parajasti siis, kui elektroni spinn on “üles”. (Olgu kass narkoosi all, nii et tema teadvus ei mängi rolli.) Kui elektron on algselt superpositsioonilises seisundis, siis on satub kass mõõtmishetkel seisundisse, mis on elususe ja surnususe superpositsioon! Alles siis, kui kasti vaatab teadvusega olend, selgub üheselt, kas kass on surnud või elus.

Selle pildi järgi on kõik makroskoopilised süsteemid kogu aeg suuremastaabilises superpositsioonis, kui läheduses ei ole teadvust. Seega pidi kogu universum olema algul hiiglaslikus superpositsioonis, enne kui esimene teadvusetiba tema seisundis ühel hoobil kollapsi esile kutsus. See kõlab võib-olla pööraselt, kuid kõige otsemat teed just niisuguse järelduseni viib kvantmehhaanika printsiipide ainus vastuvõetav sõnasõnaline interpretatsioon. Ma loodan, et siit saab selgeks, kui kummaline teooria on kvantmehhaanika ja kui tõsised on takistused tema interpreteerimise teel.

Selle intuitsioonivastase tulemusega võiks ju ehk leppidagi, kuid mina sellist interpretatsiooni siiski ei poolda. Esiteks on see vastuolus seisukohaga, mida ma olen püüdnud põhjendada — et teadvust võib kohata kõikjal. Kui teadvus on isegi väga lihtsatel süsteemidel, siis toimub kollapseid selle interpretatsiooni järgi väga madalal tasemel ja väga sageli. See ei sobi kokku füüsika andmetega, mis näitavad, et madaltasemelised superpositsioonid püsivad sageli üpris pika aja vältel kollabeerumatuna. Teine raskus seisneb selles, et puudub ligilähedaltki korralik teooria selle kohta, mis laadi on see mõju, mida teadvus peab avaldama[[397]](#footnote-398), et kollaps toimuda saaks, või missuguse kuju see kollaps omandab. Neile küsimustele annaks pakkuda palju erinevaid vastuseid, kuid üksikasjade väljatöötamiseks ei tundu leiduvat ainsatki veenvat viisi.

Raskusi valmistab juba kollapsi mõiste ise. Esiteks peab kollaps olema mittelokaalne: kui kahe osakese seisundid on omavahel põimingus, põhjustab ühe osakese mõõtmine otsekohe ka teise osakese seisundi kollabeerumise. Selline pilt aga pingestab suhteid relatiivsusteooriaga. Näiteks nõuab mittelokaalne kollaps nähtavasti viidet mingile eelistaustsüsteemile. Ilma sellise taustsüsteemita jääb teise osakese kollapsi toimumise aeg ebamääraseks, sest puudub selge definitsioon, mida tähendab samaaegsus eri kohtades.

Üldse ei sobi kogu kollabeerumisprotsess kuigi ladusalt kokku ülejäänud füüsikaga. Täht-tähelt võttes peaks kollaps olema silmapilkne, mittepidev, ajas asümmeetriline mittelokaalne protsess, radikaalselt erinev kõigist teistest protsessidest, millesse uskumiseks teoreetiline füüsika meile põhjust annab. On kummaline, et sirgjoonelise, pideva, ajas sümmeetrilise ja lokaalse Schrödingeri võrrandi kõrval eksisteerib säärane isemoodi protsess. Schrödingeri võrrandi, kvantteooria ​**[lk.341]**​ keskse võrrandi elegantsuse ja võimsuse kõrval näib kollaps tõepoolest peaaegu meelevaldselt külgetraageldatud elemendina. On väga kentsakas mõelda, et maailmas on kõige fundamentaalsemal tasemel toimimas kaks niivõrd erinevat dünaamikat.

Need argumendid ei ole muidugi veel kaugeltki otsustavad, ja interpretatsioon, mille järgi lainefunktsiooni kollapsit põhjustab teadvus, väärib väga tõsist tähelepanu. Kuid siiski arvan ma, et tuleks otsida mõnd teist interpretatsiooni, mis annaks meile looduse alusprotsessidest lihtsama ja sirgjoonelisema pildi.

#### Teine variant: püüda saada mõõtmispostulaat tasuta

Kõigi täht-tähelise interpretatsiooni raskuste allikaks on mõõtmispostulaadi tunnistamine fundamentaalseaduseks. On ahvatlev oletada, et see postulaat ei olegi alusprintsiip, vaid osutub järelduseks mingitest veelgi fundamentaalsematetest printsiipidest. Selleks on kaks võimalust. Me võime püüda võtta tarvitusele mõne uue alusprintsiibi, mis annaks samad tulemused, kuid oleks vähem kahtlane kui mõõtmispostulaat. See on [neljanda võimaluse](#ORG.10.3.4) strateegia. Teine võimalus oleks proovida tuletada needsamad tulemused teadaolevatest alusprintsiipidest, näiteks Schrödingeni võrrandist. Teisiti öeldes, me võiksime proovida saada mõõtmispostulaadi tasuta.

Selle strateegia intuitiivseid motiive pole raske mõista. Intuitsioon ütleb, et superpositsiooniefektid avalduvad esmajoones mikroskoopilisel tasemel, makroskoopilisel aga võib-olla “koonduvad” mingil moel olematuks. Võib olla, et kui mikroskoopilisi superpositsioone on palju, siis vastasmõjustuvad nad omavahel nii, et tulemuseks on suhteliselt kindlapiiriline makroskoopiline seisund. Keerukate konfiguratsioonide mõnede matemaatiliste omaduste tõttu on meil ehk võimalik näha efektiivses​[[398]](#footnote-399) kollapsis mikromääramatuse tagajärge. Fundamentaalse tõenäosusliku kollapsi asemele astuks siis emergentne statistiline protsess keerukas süsteemis.

Sellele oletusele on korduvalt proovitud matemaatilist kuju anda, kusjuures selleks on sageli otsitud abi termodünaamika statistilistelt printsiipidelt (nt. Daneri, Loinger ja Prosperi 1962). Paraku on kõik need katsed ebaõnnestunud, ja nüüdseks on valdavaks saanud arvamus, et nii see pidigi minema. Kuna Schrödingeri dünaamika on lineaarne, siis on alati võimalik konstrueerida olukordi, kus mikroskoopilised superpositsioonid viivad välja makroskoopiliste superpositsioonideni. Kui elektroni “üles”-spinn viib ühe makroskoopilise seisundini ja “alla”-spinn teiseni, siis peab superpositsioonilise elektroni tagajärjeks olema makrotasemel superpositsiooniline seisund. (Albert 1992, lk. 75, põhjendab seda punkti väga lihtsa mõttekäiguga.) See tähendab, et senikaua kui me ei too mängu uusi alusprintsiipe, tuleb makroskoopilisel tasemel oodata superpositsiooni.

Nendel strateegiatel võib olla tõepoolest midagi pakkuda. Statistika seesugune rakendus, samuti Gell-Manni ja Hartle'i (1990) jt. uuemad tööd “dekoherentsuse” alal näitavad, et sageli ei ole superpositsiooniline lainefunktsioon mitte kaootiline segadik, vaid ​**[lk.342]**​ lahutub selgesti eristuvate makroskoopiliste seisundite suhteliselt kindlapiiriliseks superpositsiooniks. Need makroskoopiliste seisundite üksteisest lahku “dekohereerumisel” on nendevahelised interferentsiefektid minimaalsed. Vähemalt aitab see mõte leida superpositsioonilises lainefunktsioonis tuttava klassikalise maailma elemente. Kuid lainefunktsioon jääb endistviisi superpositsiooniliseks ning kõigis neis uurimistöödes ei öelda midagi selle kohta, miks peaks makroskoopilise superpositsiooni elementidest aktualiseeruma vaid üks. Seega on alusprobleemi kallal tarvis veelgi vaeva näha. Sellest tööst võib olla kõige rohkem kasu ühenduses mõne teise võimalusega kvantmehhaanika interpreteerimiseks, näiteks [viiendaga](#ORG.10.5) (vt. allpool).

#### Kolmas võimalus: millest ei saa rääkida …

Tegevfüüsikute seas paistab valitsevat seisukoht, et polegi tarvis küsida, mis toimub reaalses maailmas väljaspool kvantmehhaanika matemaatilist aparaati. Matemaatika töötab, ja kogu lugu. Sellel vaatel on kaks versiooni. Esimese järgi ongi maailmas võib-olla midagi faktiliselt​[[399]](#footnote-400) toimumas, kuid meil ei ole võimalik selle kohta iial midagi täpsemat teada saada. Matemaatika annab meile kogu empiirilise informatsiooni, mida on võimalik saada, ja kõik muu on seega puhas spekulatsioon. Parem on muretsemine jätta ja edasi arvutada. Praktikas on see positsioon mõistlik, kuid ei rahulda neid, kes tahavad, et füüsika ütleks meile midagi reaalsuse kohta alustasemel. Matemaatika töötab küll, aga tahaks saada vähemalt aimugi, kuidas see võimalik on. Võib-olla ei saagi me midagi kindlalt teada, aga küsida maksab ikka.

Teine versioon on karmim ja ütleb, et selle kohta, mis maailmas toimub, ei olegi fakte — faktid lõpevadki sellega, et matemaatika töötab. Enamasti ei öelda seda nii selgelt välja — võib-olla sellepärast, et otsekoheses sõnastuses oleks seda peaaegu võimatu uskuda. Meile pakutakse pilti tegelikkusest, milles pole kohta maailmale! Nii jõuame välja idealismini, mille järgi on olemas vaid meie kujutlused, või millegi väga sarnaseni. Enne kasti avamist ei ole Schrödingeri kass elus ega surnud, tema seisund ei ole ka superpositsiooniline — ta ei ole üldse mitte mingis seisundis. Loobudes uurimast, mis mõõtmiste taga faktiliselt toimub, eitab see vaade sõltumatult eksisteerivat reaalsust.

Sageli peetakse selle seisukoha versiooniks ka Bohri ja tema kolleegide nn. Kopenhaageni interpretatsiooni, kuigi Bohri kirjatööd on mõnevõrra mitmetimõistetavad ja neist pole kerge aru saada. Neis töödes võib kohati aimata ka [esimese võimaluse](#ORG.10.3.1) (mõista arvutusi täht-tähelt)[[400]](#footnote-401) elemente, samuti kolmanda, s.t. käesolevas jaos kõne all oleva võimaluse epistemoloogilist versiooni. Bohr pani suurt rõhku mõõteseadme “klassikalisusele”; tema seisukohtadest võib välja lugeda, et objektiivsed seisundid on ainult klassikalistel (s.t. makroskoopilistel) objektidel. Küsimused objekti tegeliku seisundi kohta, mida kirjeldab superpositsioon, on lihtsalt keelu all. Kuid selline seisukoht ​**[lk.343]**​ eeldab süsteemide lahutust klassikalisteks süsteemideks ja kvantsüsteemideks, ja raske on leida selliseks lahutuseks objektiivset alust; raske on ka ette kujutada, et makroskoopiliselt tasemelt mikroskoopilisele laskudes reaalsus lihtsalt “hajub”. Paljud on leidnud, et kui Bohri seisukohta tõsiselt võtta, viib see eelmises lõigus​[[401]](#footnote-402) käsitletud tugeva operatsionalismini. Ka too vaade annab reaalsusest alustasemel pildi, mis ei olegi pilt.

#### Neljas võimalus: postuleerida täiendavaid füüsikalisi alusprintsiipe

Kui mõõtmispostulaadi sõnasõnaline interpretatsioon on vastuvõetamatu ja teda ei saa tuletada olemasolevatest füüsikalistest printsiipidest, siis on loomulik oletada, et mängus on veel midagi. Postuleerides mõne uue füüsikalise alusprintsiibi senistele lisaks, suudaksime ehk seletada kvantmehhaanika matemaatilise aparaadi tõhusust vähem kahtlasel moel.

Esimene võimalus on jätta kollapsi idee kehtima, kuid seletada seda teisiti. See strateegia ei loobu eeldusest, et füüsikalised alusseisundid on lainefunktsioonid, mida reguleerib Schrödingeri võrrand, ent võtab tarvitusele uued printsiibid, et selgitada, kuidas mikroskoopilistest superpositsioonidest sünnib makroskoopiline diskreetsus.

Selle strateegia kõige tuntum näide on Ghirardi, Rimini ja Weberi (1986; vt. ka Bell 1987a) nn. GRW interpretatsioon[[402]](#footnote-403).[[403]](#footnote-404) See interpretatsioon postuleerib fundamentaalseaduse, mis ütleb, et iga elementaarosakese asukohavektoriga võib igal hetkel väga väikese tõenäosusega toimuda mikroskoopiline “kollaps”. (Tõenäosus, et osakese kollaps toimub lähema sekundi jooksul, on umbes 10-¹⁵). Sellise kollapsiga kaasneb mittelahutuvusefektide tõttu​[[404]](#footnote-405) üldjuhul kollaps ka makroskoopilises süsteemis, mille koosseisu ta kuulub. Igas makroskoopilises süsteemis on selliseid osakesi palju, ja järelikult on iga makroskoopiline süsteem peaaegu igal ajahetkel suhteliselt diskreetses seisundis. On võimalik näidata, et tulemus on väga lähedal sellele, mida ennustab mõõtmispostulaat.

Teine võimalus on kaotada vajadus kollapsi järele, eitades superpositsioonilist lainefunktsiooni kui reaalsuse alustaseme kujutust. Kui asukohal jts. omadustel on kindlad väärtused juba baastasemel, siis ei ole kollapsit tarviski. Selline teooria postuleerib alustasemel “peitmuutujad”, mis annavad otsese seletuse tegelikule diskreetsusele makroskoopilisel tasemel. Selle valiku hind on, et nüüd on tarvis uusi printsiipe, mis selgitaksid, miks lainefunktsiooni ajalise muutumise printsiip ja kollapsiprintsiip nii hästi toimivad.

Kõige silmapaistvamaks näiteks on selles seoses teooria, mille töötas välja Bohm (1952). Tema teooria järgi on baasosakestel kogu aeg kindel asukoht. Lainefunktsioonile jääb osakese asukoha muutumist suunava omamoodi “pilootlaine” roll, lainefunktsiooni ennast kirjeldab ​**[lk.344]**​ Schrödingeri võrrand. Mõõtmispostulaadi tõenäosuslikud ennustused interpreteeruvad ümber statistilisteks seadusteks. Selle teooria järgi tuleb välja, et enne mõõtmist ei ole kunagi võimalik teada osakese täpset asukohta, vaid ainult tema lainefunktsiooni. Mõõtmispostulaat ütleb meile, kui suurel osal nii- või teistsuguse lainefunktsiooniga osakestest on üks või teine kindel asukoht. Seepärast annab ta parimad statistilised ennustused, mida me oma teadmatust arvestades oodata võime.

Kõik sedaliiki ideed toovad jällegi kaasa raskusi. Nii GRW kui ka Bohmi interpretatsioon panevad erilist rõhku kindlale asukohale, rikkudes sellega asukoha ja impulsi vahelist sümmeetriat kvantmehhanika matemaatilises aparaadis. See on mõistlik, kui eesmärgiks on ennustamine — kuna leidub argumente selle kasuks, et meie otsustusel makroskoopiliste andmete ühesuse kohta on aluseks alati kindel asukoht (näiteks mõõteseadme osuti asukoht) —, kuid teooria kaotab sellega elegantsuses. Sarnastel põhjustel valmistab tõsiseid raskusi nende ideede ühitamine relatiivsusteooriaga.

GRW teoorial on teisigi puudusi. Kõige tõsisem neist on ehk see, et rangelt võttes ei tulene sellest teooriast üldsegi, et makroskoopiline maailm peaks olema diskreetne. Makroskoopilist seisundit kujutab endistviisi superpositsiooniline lainefunktsioon: kuigi suurem osa selle amplituudist koondub ühte kohta, erineb amplituud nullist kõikjal, kus erineb nullist kollabeerumata lainefunktsiooni amplituud. Nii tulevad jälle tagasi hädad superpositsiooniga. Osuti näitab endiselt mitmesse suunda korraga, isegi pärast mõõtmist. On tõsi, et enamiku puhul neist suundadest on amplituud väga väike, kuid on raske mõista, miks peaks superpositsioon olema väikese amplituudi puhul vastuvõetavam kui suure.

Bohmi teoorial on tehnilisi raskusi vähem kui GRW interpretatsioonil, kuid temast tuleneb kummalisi järeldusi. Kõige silmatorkavamaks neist on erakordselt tugev mittelokaalsus. (Iga peitmuutujate teooria, mis on kooskõlas arvutuslike ennustustega, peab olema mittelokaalne põhjustel, millest on juttu artiklis Bell 1964.)[[405]](#footnote-406) [[406]](#footnote-407) Asi ei ole mitte lihtsalt selles, et ühe osakese omadused võivad silmapilkselt mõjutada mõne eemal paikneva osakese omadusi. Selgub, et üheainsa osakese trajektoori määramiseks tuleb võtta arvesse osakeste lainefunktsioone võib-olla isegi teistes galaktikates! Kõigil osakestel on oma roll globaalse lainefunktsiooni koosseisus, mis kirjeldab ühekorraga kõigi osakeste trajektoore kogu universumis.

Võib-olla kõige põhimõttelisem alus neis interpretatsioonides kahtlemiseks seisneb siiski selles, et nad postuleerivad lihtsuse taha midagi keerukat. Ent hoolimata oma hädadest on kvantmehhaanika matemaatiline aparaat erakordselt lihtne ja elegantne. Need interpretatsioonid aga püüavad muuta ja seletada seda lihtsat mudelit keerukate ja üpris *ad hoc* lisaprintsiipide abil. Veidi väiksemal määral on see nii GRW interpretatsiooniga, mille lisakeerukused piirduvad kahe uue fundamentaalkonstandiga ning asukoha ja impulsi vahelise sümmeetria rikkumisega; kuid sellegipoolest on erakordselt “õnnelik” juhus, et konstantide väärtused ​**[lk.345]**​ on just niisugused, et nende abiga on võimalik saada kätte peaaegu täpselt samad ennustused, mida annab ka standardne mudel. Bohmi interpretatsiooni lisakonstruktsioon on tõsisem: Bohm postuleerib lisaks asukohtade ühesusele veel ühese lainefunktsiooni ja terve uue keeruka fundamentaalprintsiibi, mille järgi osakeste asukohad determineerib lainefunktsioon; samuti rikub ta algse mudeli sümmeetriat.

Võiks öelda, et neist interpretatsioonidest jääb mulje, nagu oleks maailm Descartes'i kurja deemoni konstrueeritud — püüavad nad meid ju veenda, et maailmakord on teistsugune, kui ta tegelikult on. Nagu ütlevad Albert ja Loewer (1989): Bohmi Jumal ei mängi täringuid, kuid tal on pahatahtlik huumorimeel. Stsenaarium, kus Bohmi keerukas interpretatsioon juhtub andma samu ennustusi, mida annab ka lihtne mudel, erineb ainult astmelt olukorrast, kus keegi manipuleerib vaadiaju sisendandmeid nii, et sellele jääks mulje otseselt eksisteerivast välisest maailmast. Veel meenutab Bohmi interpretatsioon evolutsiooniteooria “interpretatsiooni”, mille järgi Jumal lõi paar tuhat aastat tagasi fossiilsed leiud oma praegusel kujul ning korraldas nii, evolutsiooniteooria ennustused osutuksid õigeteks. Seletusmudeli lihtsus on toodud ohvriks keerukale hüpoteesile, mis suudab anda samu tulemusi, mida andis juba seninegi teooria.

Kvantmehhaanika mudel on nii lihtne ja elegantne, et ükski alusteooria, millel pole samasugust lihtsust ja elegantsust, ei saa olla lõpuni rahuldav ja veenev. Kui kvantteoorias leiduks anomaaliaid või katsetulemusi, mida mudel täpselt ei ennusta, siis võiks pigem arvata, et tema lihtsus on keeruka jäämäe tipp. Praegusel kujul on mudel aga nii usaldusväärne, et keerukate mehhanismide postuleerimine tema lihtsate ennustuste seletamiseks jätaks väga eriskummalise mulje.

Arvestades, et kõigil kvantmehhaanika interpretatsioonidel on omad raskused, tuleb neid kõiki tõsiselt võtta. Kuid on loomulik, et me püüame jõuda lihtsama pildini maailmast.

#### Viies võimalus: Schrödingeri võrrandist piisab

Kvantmehhaanika keskseks komponendiks on Schrödingeri võrrand. See võrrand sisaldub mingil kujul igas kvantmehhaanika interpretatsioonis. Kõik interpretatsioonid, mida me oleme käsitlenud, lisavad Schrödingeri võrrandile midagi, seletamaks maailma makroskoopilist diskreetsust. Kuid kõige lihtsam interpretatsioon on see, mis ütleb, et Schrödingeri võrrand kehtib ja midagi rohkemat pole vajagi. See tähendab, et lainefunktsioon kirjeldab maailma füüsikalist seisundit täielikult ja Schrödingeri võrrand kirjeldab lainefunktsiooni ajalist muutumist täielikult. Selle interpretatsiooni on esitanud Everett (1957, 1973).

Üks [ülal](#ORG.10.3.2) esitatud strateegia (2. variant, püüda saada mõõtmispostulaat tasuta) ütleb samuti, et Schrödingeri võrrandist piisab, kuid kinnitab, et see võrrand on ühitatav diskreetsusega ​**[lk.346]**​ makroskoopilisel tasemel. Me nägime, et see ei saa väga lihtsatel matemaatilistel põhjustel paika pidada. Everetti interpretatsioon on palju radikaalsem. Tema leiab, et Schrödingeri võrrandit tuleb võtta puhta kullana ning et lainefunktsioon kirjeldab maailma seisundit igal tasemel. Sellest järeldub, et vastupidiselt sellele, mis meile paistab, on maailm isegi makroskoopilisel tasemel superpositsioonilises seisundis.

### 4. Everetti interpretatsioon

Selle interpretatsiooni tõukejõud on ilmsed. Kvantmehhaanika tuumaks on Schrödingeri võrrand. Mõõtmispostulaat ja kõik teised väljapakutud printsiibid näevad välja nagu selle võrrandi lisandused. Miks siis mitte neist vabaneda? Samavõrd ilmne on selle interpretatsiooni puudus. Kui Schrödingeri võrrandist piisab, siis on maailm igal tasemel superpositsiooniline. Aga pealtnäha see nii ei ole: me ei näe kunagi osuteid kahe seisundi superpositsioonis. Miks?

Igatahes on see interpretatsioon vähemasti väga intuitsioonivastane. Ta ütleb, et elektroni seisundit on kõige parem kirjeldada superpositsiooniga, kuid sama käib ka elektroni mõõtva osuti seisundi kohta! Objektiivselt võttes ei ole päris rangelt õige öelda, et osuti näitab üles või alla. Tegelikult on ta üles- ja alla-seisundi superpositsioonis. Sama käib peaaegu kõigi asjade makroskoopiliste seisundite kohta: seisund, mida nende lainefunktsioonid kirjeldavad, ei vasta peaaegu kunagi üheleainsale “diskreetsele” seisundile. Tuleb välja, et superpositsioon on kõikjal. Aga miks paistab maailm meile siis diskreetsena?

Everetti vastuseks sellele küsimusele on superpositsiooni mõiste laiendamine, nii et superpositsioon hõlmaks isegi vaimu. Kui me võtame Schrödingeri võrrandit tõsiselt, siis on seisundite superpositsioonis nii elektroni mõõtev osuti kui ka osutit tajuva inimese aju. Aju seisund on kirjeldatav kahe seisundi superpositsioonina: ühes seisundis tajub inimene, et osuti näitab üles, teises, et alla. Everetti otsustav samm seisneb oletuses, et kummalegi neist kahest seisundist vastab omaette vaatleja. Pärast mõõtmist on vaatlejaid kaks. Üks neist tajub osutit näitavat üles, teine alla. Sellest järeldub, et mõlemad vaatlejad elamuses on maailma üks diskreetne seisund.

Edasi näitab Everett, et selle skeemi järgi on neil vaatlejatel enamik omadusi, mida me ootame vaatlejatel olevat, ning et nii saab tuletada enamiku kvantmehhaaniliste arvutuste ennustustest. Näiteks ei ole raske näha, et nüüd ei ole kummalgi kahest superponeeruvast seisundist ligipääsu teisele seisundile, nii et ​**[lk.347]**​ vaimu superpositsioon ei reeda end ainsaski seisundis. On isegi võimalik näidata, et kui mõõtmist sooritav vaatleja tajub teist vaatlejat, kes mõõdab sama suurust, siis on tajutavad mõõtmistulemused omavahel kooskõlas, nii et maailm näib täiesti koherentsena. Lühidalt, iga üksiku vaatleja maailmaelamus on paljuski just niisugune, nagu me ootame, olgugi maailm ise superpositsioonilises seisundis.

Seda interpretatsiooni ei tohiks segi ajada *hargmaailmade*​ interpretatsiooniga, mis ütleb, et iga kord, kui toimub mingi mõõtmine, hargneb maailm sõna otseses mõttes paljudeks omaette maailmadeks. Ühes maailmas näitab osuti üles, teises sellest maailmast täiesti sõltumatus maailmas alla. Asja niimoodi kujutades on see positsioon kõike muud kui lihtne. Esiteks nõuab ta mingit uut ja väga fundamentaalset printsiipi hargnemisprotsessi kirjeldamiseks. Pole kaugeltki selge, millal see hargnemine nimelt toimuma peaks (selline on mõõtmisprobleemi nüüdne kuju), ja jääb täiesti ebaselgeks, mida need hargemise tagajärjel ilmunud maailmad endast kujutama peaksid. Selleks, et toimuks sõna otseses mõttes hargnemine, peab lainefunktsioon “lahutuma” paljudeks komponentideks. Kuid lainefunktsiooni lahutamiseks on palju võimalusi ja kvantmehhaanika ei anna mingit alust eelistada üht võimalust teisele. See interpretatsioon tundub olevat veelgi keerulisem ja veelgi enam *ad hoc* kui kõikvõimalikud “kollapsipõhised” interpretatsioonid, ja pole kuigi palju põhjust teda omaks võtta.

Seda hargnemiskontseptsiooni omistatakse sageli Everettile (paljus tänu Everetti tööde tõlgendustele DeWitti artiklites (1970,1971)), kuid Everetti enda avaldatud töödest pole seda leida. Everetti enda seisukoht ei ole päris selge, kuid seda on palju loomulikum tõlgendada minu pakutud viisil; just niisugust tõlgendust soovitavad ka Albert ja Loewer (1988) ning Lockwood (1989). Selle vaate kohaselt ei toimu objektiivselt mingit hargnemist, vaid lainefunktsioon areneb aja möödudes seisundite superpositsiooniks, mille osaseisundeid on kõige parem käsitleda üheainsa maailma komponentidena. Everetti seisukohta nimetatakse mõnikord *hulgimaailmade interpretatsiooniks* (juba nimi viib mõtte hargmaailmade vaatele), kuid seisukohta, mille üle mina siin arutlen, on täpsem nimetada *ühe suure maailma* interpretatsiooniks​. Maailmu on üksainus, kuid see üks maailm sisaldab rohkem, kui võinuks arvata.[[407]](#footnote-408).

Selle vaate järgi toimub hargnemine kui üldse, siis ainult vaatlejate vaimus. Kuna superpositsioonid mõjutavad subjekti ajuseisundit, siis tekib mitu omaette vaimu, millest igaüks vastab superpositsiooni ühele komponendile. Iga niisugunen vaim tajub omaenda diskreetset maailma, mis vastab maailmale, mida tajume meie — nimetagem seda *minimaailmaks*​, et hoida seda lahus superpositsioonilisest *maksimaailmast*​. Reaalne maailm on maksimaailm, minimaailmad eksisteerivad aga vaid subjektide vaimudes. Everett nimetab oma seisukohta *suhteliste seisundite* interpretatsiooniks: minimaailma seisundit, kus osutid näitavad diskreetseid positsioone, võib pidada maailma seisundiks ainult vaatlejapoolse kirjelduse suhtes. Objektiivselt on maailma seisund superpositsiooniline.

​**[lk.348]**​ Selles interpretatsioonis on aga üks otsustava tähtsusega element analüüsimata jäänud. Miks on legitiimne identifitseerida iga niisuguse ajuseisundi komponenti omaette vaatlejaga? Miks mitte kõnelda ühestainsast vaatlejast, kelle vaimuseisundiks on superpositsiooniline segapundar? Ja miks peaks säärasest mittekoherentsest ajuseisundist üldse sündima vaim? Everetti käsitlus vaatab neist esmatähtsatest küsimustest mööda. Tõepoolest, võib tunduda, et seostades ajuseisundi lainefunktsiooni mitme vaimuga, millest igaüks tajub üht diskreetset seisundit, apelleerib Everett ebakorrektselt eelisbaasile, samamoodi nagu seda teeb hargmaailmade interpretatsioon. Lainefunktsioonil ei ole üht objektiivset lahutust — teda saab lahutada komponentideks väga mitmel viisil, olenevalt vektorruumi baasi valikust. Sageli on otstarbekas ja loomulik lahutada lainefunktsioon mingil ühel kindlal viisil, vastavalt ühele konkreetsele baasile. Kuid selline lahutus ei kajasta lainefunktsiooni objektiivseid omadusi. Aju seisundi saab lahutada “ülespoole näitavana tajutud osuti” ja “allapoole näitavana tajutud osuti” seisundiks, ent sama hästi võib teda lahutada kaheks seisundiks nii, et tajud on mõlemas mitmesed. Tundub, et objektiivset lahutust postuleerides väljub Everett Schrödingeri võrrandi võimaluste piiridest.

Everett on jätnud oma käsitlusest välja ühe olulise elemendi: ta ei analüüsi suhet keha ja vaimu vahel. Ta eeldab, et superpositsioonilise ajuseisundiga seondub mitu omaette elamussubjekti, kuid ei püüa seda eeldust mitte kuidagi põhjendada. On selge, et vastus sellele küsimusele sõltub oluliselt meie teadvuseteooriast. Midagi sarnast ütleb ka Penrose (1989):

Eelkõige ei mõista ma, miks peaks teadvusega olend olema teadlik ainult “ühest” lineaarse superpositsiooni alternatiivist. Missugune teadvuse omadus keelab, et kellegi “teadlikkus” sisaldaks surnud kassi ja elava kassi ahvatlevat lineaarkombinatsiooni? Minule tundub, et hulgimaailmade kontseptsiooni ühildamine tegelike vaatlusandmetega eeldaks mingit teadvuseteooriat. (p. 296)

Tõepoolest, seda kvantmehhaanika keskset küsimust võib mõista kui füüsikaliste protsesside ja elamuse vahelise seose küsimust. Kvantmehhaanikas on kesksel kohal pilt, mille järgi mikroskoopilist reaalsust kirjeldab superpositsiooniline lainefunktsioon, mis muutub ajas vastavalt Schrödingeri võrrandile. Kuid meie elamuses on maailm diskreetne. Keskne küsimus on: kuidas see saab nii olla? Eri interpretatsioonid vastavad sellele erinevalt. Mõned (nagu Bohmi oma) eitavad esimest eeldust, kinnitades, et reaalsus on diskreetne isegi alustasemel. Mõned postuleerivad alusprintsiipe (mõõtmispostulaat, GRW kollapsiseadus), mille varal toimub üleminek superpositsioonilisest seisundist diskreetsesse. Mõned teooriad (need, mis valivad 2. variandi — püüda saada mõõtmispostulaat tasuta) püüavad selgitada, kuidas superpositsioonilistest mikroskoopilistest seisunditest võib statistiliselt välja kasvada diskreetne makroskoopiline reaalsus. ​**[lk.349]**​ Need kolm viimast strateegiat on kõik kaudsed — nad püüavad seletada elamuse diskreetsust elamust kandva makroskoopilise reaalsuse diskreetsusega.

Teine võimalik strateegia on otsida elamuse küsimusele otsest vastust. Schrödingeri võrrandi primaati tõsiselt võttes on keskne küsimus järgmine: miks on meie elamused just niisugused, nagu nad on, kui maailma füüsikaline struktuur on niisugune, nagu ta on? Aga see ei ole midagi muud kui küsimus, kuidas sünnib mingitest kindlatest füüsikalistest struktuuridest elamus. Teisiti öeldes, see on just see küsimus, mille üle ma olen arutlenud kogu käesolevas raamatus ja millele teadvuseteooria peaks suutma vastata.

Juhul, kui sellele küsimusele vastamiseks tuleb postuleerida mingi *ad hoc*​ teadvuseteooria, kahaneb Everetti interpretatsiooni köitvus tunduvalt. Selle interpretatsiooni parimaks omaduseks on alati olnud lihtsus, ent uute ja meelevaldsete psühhofüüsikaseaduste postuleerimise järel oleks ta samavõrd *ad hoc*​, nagu seda on Bohmi interpretatsioon. Ent kui sellele küsimusele suudab vastata sõltumatutele alustele toetuv teadvuseteooria, siis hakkab Everetti interpretatsioon tõepoolest atraktiivset ilmet omandama.

Teadvuseteooria, mida mina siin propageerinud olen, suudab anda sellele küsimusele õigekujulise vastuse. Selgub, et teooria ennustuse järgi peaks superpositsioonilise seisundiga seostuma mitu omaette subjekti, kelle elamused on diskreetsed. Et seda mõista, nimetagem *maksimaalseks fenomenaalseks seisundiks*​ fenomenaalset seisundit, mis iseloomustab kogu subjekti elamust mingil ajahetkel, ja *maksimaalseks füüsikaliseks seisundiks*​ füüsikalist seisundit, mis iseloomustab täielikult süsteemi seesmist füüsikalist seisundit samal ajahetkel. Järelduse tegemiseks piisab järgmisest *superpositsiooniprintsiibist*​:

Kui teooria ennustab, et süsteemis, mille maksimaalne füüsikaline seisund on ​*P*​, sünnib maksimaalne fenomenaalne seisund ​*E*​, siis ennustab ta ka, et süsteemis, mille seisundiks on seisundi ​*P* superpositsioon mingi selle seisundi suhtes ortogonaalse füüsikalise seisundiga, sünnib samuti fenomenaalne seisund ​*E*​.

Kui see printsiip paika peab, sünnivad ortogonaalsete füüsikaliste seisundite superpositsioonist vähemasti need maksimaalsed fenomenaalsed seisundid, mida oleks sünnitanud igaüks neist füüsikalistest seisunditest omaette. Just seda nõuab Everetti interpretatsioon. Kui aju on “ülespoole näitavana tajutud osuti” seisundi ja “allapoole näitavana tajutud osuti” seisundi superpositsioonis, siis sünnib sellest vähemalt kaks elamussubjekti, millest ühe elamuses näitab osuti üles ja teises alla. (Loomulikult on need kaks eri subjekti, kuna mõlemad fenomenaalsed seisundid on subjekti maksimaalsed fenomenaalsed seisundid.) Nii on see ka ​**[lk.350]**​ üldjuhul. Superpositsioonis sünnib alati subjektide ansambel, nagu Everetti interpretatsioon nõuabki.

Seega tuleb näidata, et kirjeldatud teooria eeldab superpositsiooniprintsiipi. Kõige lihtsam viis selleks on vaadelda [9. peatükis](#ORG.9.3) esitatud mudelit ning eelkõige väidet, et teadvus sünnib õigekujulise arvutuse implementatsioonist. Kehtestamaks selle väite abil superpositsiooniprintsiipi, tuleb tõestada, et kui süsteem, mis implementeerib mingit arvutust, on maksimaalses füüsikalises seisundis ​*P*​, siis implementeerib sama arvutust ka süsteem, mille seisundiks on seisundi ​*P* superpositsioon seisundi ​*P* suhtes ortogonaalsete füüsikaliste seisunditega.

Niisiis, olgu esialgne süsteem (mis on maksimaalses füüsikalises seisundis ​*P*​) arvutussüsteemi ​*C* implementatsioon. See tähendab, et süsteemi füüsikaliste osaseisundite omavahelised põhjuslikud suhted saab seada vastavusse arvutussüsteemi ​*C* formaalsete osaseisundite vaheliste formaalsete suhetega.[[408]](#footnote-409) Siis võimaldab sama vastavuse mingi versioon implementeerida süsteemi ​*C* ka superpositsioonilises süsteemis. Igale esialgse süsteemi osaseisundile ​*S* võib superpositsioonilises süsteemis leida vastava osaseisundi ​*S′* ilmse projektsioonisuhte abil: superpositsiooniline süsteem on ​*S′*​, kui süsteem, mis tekib tema projitseerimisel seisundile ​*P* vastavale hüpertasandile, on ​*S*​. Sellest, et superpositsioonilise süsteemi seisundiks on seisundi ​*P* ja selle seisundi suhtes ortogonaalsete seisundite superpositsioon, järeldub, et kui algne süsteem on seisundis ​*S*​, siis on superpositsiooniline süsteem seisundis ​*S′*​. Ja kuna Schrödingeri võrrand on lineaarne, siis järeldub sellest ka, et osaseisundite ​*S′* vahelised seisundisiirdesuhted peegeldavad täpselt esialgsete osaseisundite ​*S* vahelisi suhteid. Me teame, et viimased peegeldavad omakorda täpselt süsteemi ​*C* osaseisundite vahelisi formaalseid suhteid. Sellest järeldub, et süsteem ​*C* on implementeeritav ka superpositsioonilises süsteemis. Kuid see ongi tulemus, milleni me jõuda tahtsime. Organisatsioonilise invariantsuse printsiibist tuleneb, et kui esialgses süsteemis sünnib elamussubjekt, siis sünnib superpositsioonilises süsteemis sellest subjektist kvalitatiivselt eristamatu elamussubjekt.

Superpositsiooniprintsiibi kasuks võib leida argumente ka informatsiooni topeltaspektiteooria abiga: me võime oletada, et kogu relevantne informatsioon, mis sisaldub esialgses füüsikalises seisundis, on olemas ka superpositsioonilises seisundis. Ent topeltaspektiteooria mitmetivõetavuse tõttu on see argument vähem selge kui eelmine, mistõttu ma ei hakka teda siinkohal käsitlema. Oluline on see, et nii või teisiti ennustab minu osaline teadvuseteooria tulemust, mida nõuab Everetti interpretatsioon: ta ennustab, et kui maailm ongi hiiglaslikus superpositsioonis, leidub ikkagi subjekte, kelle elamuses maailm on diskreetne.

Kui rohkem raskusi ei ole, siis järeldub sellest, et Schrödingeri võrrandi ja sõltumatule alusele toetuva teadvuseteooria ühendus võib anda maailmast pildi, mida me tegelikult näemegi. See tähendab, et ainus füüsikaprintsiip, mida kvantmehhaanika vajab, on Schrödingeri võrrand, ning et mõõtmispostulaat ja muud alusprintsiibid on vaid tarbetu ballast. ​**[lk.351]**​ Muidugi on tarvis ka psühhofüüsikaprintsiipe, kuid neid vajame me ju igal juhul, ja selgub, et need printsiibid, mis on usutavad sõltumatutel põhjustel, suudavad ära teha kogu töö, mida siin tarvis läheb. Kõik see moodustab ühtekokku tugeva argumendi Everetti interpretatsiooni tõsiseltvõetavuse kasuks.

### 5. Vastuväiteid Everetti interpretatsioonile

Everetti interpretatsiooni on kirjanduses sageli rünnatud, kusjuures mitte kõik etteheited ei ole ühtviisi kaalukad. Ma rühmitan nad mitmesse klassi.

#### Vastuväited “maailmade hargnemise” alusel

Paljud vastuväited tulenevad Everetti vaate tõlgendamisest — õigusega või mitte — hargmaailmade vaatena. See on arusaadav, kui arvestada, et seda vaadet nimetatakse sageli ka hulgimaailmade interpretatsiooniks. Näiteks Bell (1976) väidab, et mõõtmise mõiste ebaselguse tõttu ei ole arusaadav, millal “hargnemine” toimuma peaks, ning et puudub põhjus eelistada üht baasi teistele. On selge, et need vastuväited ei kehti siinse interpretatsiooni kohta, mis ei nõua ei objektiivset “harunemist” ega eelisbaasi valimist. Sarnases vaimus arutleb Hughes (1989), kes võtab sõna “ontoloogilise laviini” vastu hargnemisprotsessis, Healey (1984) aga tähendab, et uute maailmade loomine rikub massi ja energia jäävuse seadust! Kahju, et Everetti vaate ilmselt vigane tõlgendus hargmaailmade vaatena on nii populaarseks saanud, huvitavamale interpretatsioonile aga ei ole osutatud väärilist tähelepanu.

#### Vastuväited eelisbaasile

Mõned vastuväited hargmaailmade interpretatsioonile toetuvad asjaolule, et see interpretatsioon nõuab eelisbaasi. Kuid seda nõuavad ka mõned vastuväited ainumaailma versiooni vastu. Muu hulgas kerkib küsimus: miks peaksid superpositsioonilise ajuseisundiga seonduvad vaimud vastama ainult selle seisundi lahutusele eelisbaasi järgi?[[409]](#footnote-410) Miks ei sünni vaime teistest lahutustest või isegi kogu superpositsioonilisest seisundist kui tervikust? See on mõistlik vastuväide Everetti enda versioonile, mis näib sellist kanoonilist lahutust eeldavat. Kuid sama vastuväidet ei saa esitada minu versioonile, mille järgi superpositsioon loob vastavad diskreetse elamuse subjektid, ilma vähimagi vajaduseta postuleerida eelisbaasi. Ja mul pole olnud vajadust ka eelduse järele, et need on ainsad vaimud, mis säärasest superpositsioonilisest süsteemist sündida võivad.

#### ​**[lk.352]**​ Aga superpositsioonilised vaimud?

Kerkib küsimus: “Kas ühe kindla superpositsiooni juurde kuulub ka teisi vaime?”. Vastus on “võib-olla”. Kui aktsepteerida informatsiooni topeltaspektiteooriat, siis me juba teame, et sellises süsteemis võib madaltaseme protsessidega kaasneda elamusi. Samuti võib elamussubjekte seonduda superpositsiooniliste infotöötlusprotsesside struktuuriga. Võib-olla seondub vaime ka süsteemi teistsuguste lahutustega, või kuulub kogu superpositsioonilise süsteemi juurde hoopis üksainus superpositsiooniline vaim. Sääraste vaimude olemasolu sõltub teadvuseteooria üksikasjadest, kuid on raske näha, kuidas võiks nende olemasolu raskusi valmistada.

Superpositsiooniliste vaimude võimalikkusele saaks üles ehitada argumendi minu teooria vastu, küsides: miks minu vaim ei ole superpositsiooniline? Vastus: sellepärast, et ma olen see, kes ma olen. Teooria ennustab mittesuperpositsiooniliste vaimude olemasolu, ja minu teadvus osutub lihtsalt üheks neist. Küsida, miks minu vaim ei ole üks superpositsioonilistest vaimudest, on sama mis küsida, miks ma ei ole hiir. See lihtsalt kuulub minu eksistentsi toore indeksikaalsuse juurde. Hiirevaimud on olemas ja võib-olla on olemas ka superpositsioonilised vaimud, kuid minu vaim ei kuulu nende hulka. Vastuväide: miks ei ole mul ligipääsu superpositsioonilistele vaimudele, näiteks mälestustele superpositsioonilistest elamustest? Vastus: teooria ennustab, et kõnealusetele diskreetsetele vaimudele avaneb maailm kõigiti diskreetsetena, neil puudub otsene ligipääs superpositsiooni teistele komponentidele. Näiteks kõik nende mälestused on mälestused diskreetsetest vaatlustest.

Igal juhul võib oletada, et kõigist superpositsioonilise süsteemi juurde kuuluvatest vaimudest pakuvad huvi ainult tavalised diskreetsed vaimud. Need on keerukad ja koherentsed ning nende elamused peegeldavad ratsionaalsete protsesside struktuuri. Kõik teised süsteemiga seonduvad vaimud on suhteliselt mittekoherentsed ja neil puudub kuigivõrd huvitav struktuur. Seda järeldust toetab Gell-Manni ja Hartle'i (1990) ja teiste autorite “dekoherentsus”-mudel. Selle mudeli järgi leidub lainesarnastes keerukates adaptiivsetes süsteemides huvitavaid struktuure “loomuliku” lahutuse teel saadud komponentide sees; süsteem “dekohereerub” loomulikul teel kindlates suundades. Seega võib ratsionaalsetes süsteemides koherentset kognitiivset struktuuri leida ainult selle loomuliku lahutuse komponentides, ja ainult nendes võib sündida keerukas, koherentne vaim. Teisi säärastes süsteemides leiduvaid elamussubjekte ei saa pidada isikuteks.

####   Vastuväited isikusamasuse alusel

Terve hulk intuitiivselt ärevaks tegevaid seiku seondub vaatleja isikusamasusega. Võtame vaimu ​*M*​₁, mis minu mäletamist mööda oli eile samal ajal olemas. Täna hargneb sellest vaimust superpositsiooni eri “harudes” ​**[lk.353]**​ palju vaime. Minu vaim ​*M*​₂ on vaid üks nendest. Ma võiksin väga hästi küsida: miks sattusin ma just sellesse harru, mitte mõnesse teise? Hofstadter (1985b) väljendab seda mõtet nii:

Miks liigub minu ühtne minatunne edasi mööda just seda juhuslikku haru, mitte mõnd teist? Mis seaduse alusel toimuvad need juhuslikud valikud, mis tõstavad teiste seast esile just selle haru, mida mina tunnen end järgivat? Miks ei lähe minu minatunne pärast hargnemist kaasa teiste minadega, et järgida teisi trajektoore? Mis seob minasuse selle keha selle vaatepunkti külge, mis liigub praegusel ajahetkel mööda just seda universumi haru?

Vastuseks peame taas viitama toorele indeksikaalsusele: minu vaim ongi see vaim, ja kogu lugu. Tunne ütleb, et siin peab olema mängus midagi sügavamat ning et asjaolu, et eilsest vaimust ​*M*​₁ on tänaseks saanud just vaim ​*M*​₂, aga mitte mõni noist teistest, on mingil moel oluline fakt maailma kohta. Objektiivselt vaadates aga ei ole harus ​*M₂* midagi erilist. Isegi vaimu ​*M*​₁ seisukohalt on kõik tänased vaimud võrdväärsed. Ükski neist ei ole vaimu ​*M*​₁ ainus õigusjärgne pärija; kõik nad kannavad vaimu ​*M*​₁ “minasust” võrdsel määral. Vaim ​*M*​₂ tundub erilisena ainult sellelt ühelt seisukohalt — vaimu ​*M*​₂ seisukohalt. (Muidugi tundub minu superpositsioonilistele teisikutele omaendi seisukohalt samamoodi.) Vaimu ​*M*​₂ eelisroll on vaid järjekordne indeksikaalne fenomen, samasugune nagu see, et ma olen David Chalmers, aga mitte Rolf Harris. See vaim on siin, mitte seal. See on samavõrd mõistatuslik fakt nagu indeksikaalsed faktid ikka ega too maailma asümmeetriat juurde.

Intuitsioon ütleb meile väga selgesti, et isikusamasus peab alati olema faktiline: kui minu praegusest seisundist hargneb palju vaime, siis peab leiduma fakt, mis ütleb, milline neist olen mina. Kuid seda ideed on mõjuvalt kritiseerinud Parfit (1984), kes näitab veenvalt, et isikusamasuse faktiline külg ammendub faktidega psüühika järjepidevuse, mälu jms. kohta. Kui selle analüüsiga nõustuda, siis võivad kõik homsed vaimud pretendeerida minasusele võrdväärselt ja pole fakti, mille alusel neid eristada saaks. Selles järelduses on midagi ebameeldivat, kuna ta teeb isikuidentsuse selgepiirilise “voo” illusoorseks. Kuid Parfiti analüüs annab põhjust uskuda, et see selgepiiriline voog oligi kogu aeg illusoorne.

#### Tõenäosuste interpreteerimine

Kõige olulisem vastuväide Everetti interpretatsioonile on, et see ei suuda mõtestada tõenäosusi, mida ennustab mõõtmispostulaat.[[410]](#footnote-411) Antud juhul võib mõõtmispostulaat meile öelda, et mingi kindel mõõtmine annab tõenäosusega 0.9 “üles”- ja tõenäosusega 0.1 “alla”-spinni. Everetti ​**[lk.354]**​ interpretatsiooni järgi on nii osuti kui ka vaatleja ajuseisund tegelikult superpositsioonis, millesse kuulub (vähemalt) kaks elamussubjekti. Ühel neist on elamuses osuti, mis näitab “üles”, teisel osuti, mis näitab “alla”. Täpselt sama lugu oleks olnud juhul, kui tõenäosused oleksid olnud 50:50. Muidugi koondub juhul 90:10 suurem osa superpositsioonilise lainefunktsiooni amplituudist tulemusele “üles” vastava ajuseisundi vahemikku, kuid mis on sellel tegemist tõenäosustega?[[411]](#footnote-412)

Everett postuleerib selle küsimuse lahendamiseks mõõtmispostulaadi abil saadud tõenäosustele (s.t. lainefunktsiooni vastavate osade amplituudide ruutudele) vastavas vaatlejate ruumis mõõdu​[[412]](#footnote-413), mille abil ta püüab näidata, et enamiku vaatlejate (s.t. vaatlejate osahulga, mille mõõt on 1) mälestus vaatlusest on kooskõlas sagedustega, mida ennustavad mõõtmispostulaadi abil saadud tõenäosused. Näiteks vaatlejate hulgas, kes on sooritanud äsjakirjeldatud mõõtmise mitu korda, mäletab enamik neist, et osuti näit oli 90% juhtudest “üles” ja 10% juhtudest “alla”. Seega on tõenäosustele roll leitud. Tekib aga küsimus: millega seda vaatluste ruumi mõõtu põhjendada? Sama ruumi teisiti mõõtes saaksime hoopis teised sagedused. Näiteks kui me omistaksime kahele superpositsioonilisele vaatlejale (olenemata amplituudidest) alati võrdsed mõõdud, siis mäletaks enamik vaatlejaid, et “üles”- ja“alla”-näitude suhe oli 50:50. Ei Schrödingeri võrrand ega psühhofüüsika seadused ei garanteeri kummagi mõõdu “õigsust”.

Albert ja Loewer (1988) heidavad samast murest lahtisaamiseks mõõdud üldse kõrvale. Selle asemele postuleerivad nad radikaalsemad psühhofüüsikaseadused, mille järgi iga ajuseisund kuulub lõpmatule hulgale vaimudele. Iga vaimuga, mille postuleeris eelmine kontseptsioon, seab see teooria vastavusse kvantitatiivselt identsete vaimude lõpmatu ansambli. Ning seal, kus Everetti teooria ennustab vaimu hargnemist kaheks vaimuks, ütleb see teooria, et iga vaim liigub kas ühte või teise suunda, ja mõlema liikumise tõenäosus on leitav mõõtmispostulaadi abil. Seega siirdub iga mõõtmiseelsele ajuseisundile vastav vaim tõenäosusega 90 % seisundisse, kus ta tajub, et osuti näit on “üles”, ja tõenäosusega 10 % seisundisse, kus ta tajub, et näit on “alla”. Nii jäävad kvantmehhaanika matemaatilise aparaadi tõenäosuslikud ennustused kehtima.

Siin on lihtsust juba selgelt vähem. Uutel psühhofüüsikaseadustel puudub igasugune iseseisev põhjendus, ja lisaks sellele vajab teooria ka “intrapsüühilisi” seadusi, mis reguleerivad vaimude ajalist muutumist. Nende *ad hoc* postulaatide tarvituselevõtmisega toob see teooria ohvriks mõned Everetti interpretatsiooni peamised eelised. Neid intrapsüühilisi seadusi võib kahtlasteks pidada ka sellepärast, et nad postuleerivad fundamentaalseid redutseerimatuid fakte ajalise isikusamasuse kohta. Ei ole ​**[lk.355]**​ selge, mida neist faktidest arvata tuleks. Muu hulgas tähendaks nende aktsepteerimine, et Parfiti esitatud isikusamasuse analüüsist tuleb loobuda. Need faktid ei ole füüsikaliste faktide suhtes superventsed isegi mitte looduslikult, ja see ajab metafüüsilise pildi keeruliseks. Seda interpretatsiooni ei saa võimalusena maha kanda, kuid tema hind on kõrge.

Teine võimalus oleks loobuda lisakeerukustest ja vaadata, kas tõenäosusi ei saaks alal hoida muul viisil. Ahvatlev on näha selles indeksikaalsuse probleemi. Miks sattusin ma kõigist lainefunktsiooni piirkondadest just, sellesse, kus minu mälestus vastab matemaatilisele ennustusele, aga mitte mõnesse teise? Üks võimalus on võtta seda lihtsalt kui toorest indeksikaalset fakti: mingi hulk vaime peab selles piirkonnas ju olema, ja mina olengi juhtumisi üks neist. Kuid see vastus ei tundu päris rahuldavana, sest niiviisi osutub arvutustulemuste tähelepanuväärne regulaarsus tohutuks kokkusattumuseks. Meil tuleb leida viis, kuidas näidata, et see ei ole nii.

Isegi kui tunnistada, et see, et ma sattusin just siia, juhuseks, peitub juba selles tunnistuseski vaikimisi idee mingist mõõdust üle vaimude ruumi. Selle idee mõte oleks, et minu tõenäosus saada viimaks üht, aga mitte teist tüüpi vaimuks, oli aprioorselt suurem, võib-olla näiteks teise tüübi suhteliselt laia leviku tõttu. Selline mõõt figureerib vaikivalt paljudes meie arutlustes maailma kohta. Kui ma tuletan mingitest andmetest induktiivselt mingi järelduse, siis ma tean, et mõne teise vaatleja jaoks, kes on sarnases episteemilises positsioonis, ei tarvitse minu järeldus kehtida, kuid eeldan, et ta enamiku selliste vaatlejate puhul kehtib, isegi kui igasse klassi kuulub lõpmatu arv vaatlejaid. Teisiti öeldes, ma eeldan, et mul oli aprioorselt suurem tõenäosus sattuda ühte klassi kui teise. Selline mõttekäik eeldab vaikivalt mingit mõõtu üle vaimude ruumi.

Seega pole võimatu, et tõenäosusi saab põhjendada mingi taolise mõõdu selgesõnalise sissetoomise teel. Kindlasti koondub suurem osa lainefunktsiooni amplituudist piirkondadesse, kus vaatlejate mälestused vastavad matemaatilistele ennustustele. Võib-olla satub minu vaim piirkonda, kus amplituud on suur, suurema tõenäosusega kui niisugusesse, kus amplituud on väike. Ja kui me eeldame, et minu aprioorne tõenäosus osutuda just selleks, aga mitte mõneks teiseks vaimuks, on võrdeline lainefunktsiooni vastava osa amplituudi ruuduga, siis järeldub sellest, et minu mälestused jäävad peaaegu kindlasti vahemikesse, mida kvantmehhaanilised arvutused ennustavadki.

Kuid millele see mõõt objektiivselt vastab? Kas teda tuleb võtta kui “minade” jaotuse alusfakti? Kas teda on kuidagi võimalik põhjendada kui kanoonilist mõõtu​[[413]](#footnote-414) selles ruumis? Need on rasked küsimused, mis puutuvad tihedalt kokku otse indeksikaalsuse müsteeriumiga: miks olen ma sattunud olema just see isik, mitte mõni teine? See ongi üks peamisi müsteeriume ja pole sugugi selge, kuidas sellele küsimusele vastata tuleks. Sellegipoolest tundub, et idee mõõdust üle vaimude ruumi üht-teist tõotab ​**[lk.356]**​ ning et teda võib tarvis minna ka teistel eesmärkidel, näiteks induktsiooni põhjendamiseks. Seni aga jääb tõenäosuste interpreteerimine Everetti interpretatsiooni kõige tõsisemaks raskuseks.

### 6. Kokkuvõte

Tuleb tunnistada, et Everetti interpretatsiooni on peaaegu võimatu uskuda. Everett postuleerib, et maailm on palju suurem, kui me iial arvanud oleme. Maailm on tegelikult hiiglaslik aegade algusest saadik eri teid kulgenud seisundite superpositsioon, ja elamuses avaneb sellest meile vaid tibatilluke osaseisund. Samuti postuleerib see interpretatsioon, et minu tulevik ei ole deterministlik: juba minuti jooksul ilmub suur hulk vaime, kellest kõigil on võrdne õigus nõuda, et neid peetaks minuks. Sellest, kui ma kirjutasin eelmise lause, ongi juba minut möödas — kes teab, mida teevad praegu kõik need teised vaimud?

Teisest küljest on nüüdseks selge, et kõik kvantmehhaanika interpretatsioonid ongi mingil määral pöörased. See on kvantmehhaanika fundamentaalne paradoks. Kolmeks kõige tähtsamaks inter­pretat­siooniks võib vististi pidada Wigneri interpretatsiooni, mille järgi kollapsi kutsub esile teadvus, Bohmi mittelokaalsete peitmuutujate interpretatsiooni ja Everetti interpretatsiooni. Kui vaadelda, mis järeldusteni need interpretatsioonid viivad, siis need on järgmised: Wigneri interpretatsioonist järeldub, et makroskoopilised objektid on sageli superpositsioonilistes seisundites, enne kui nad vaatleja juhusliku pilgu toimel kollabeeruvad; Bohmi interpretatsiooni järgi sõltub iga osakese trajektoor universumis iga teise osakese seisundist; ning Everetti interpretatsioon ütleb, et maailm on palju suurem, kui me iial arvanud oleksime.

Nendest kolmest interpretatsioonist on Bohmi oma vahest kõige vähem pöörane ja Everetti oma kõige pöörasem, Wigneri kontseptsioon jääb nende vahepeale. Seevastu teoreetilise kvaliteedi põhjal on järjestus vastupidine. Bohmi interpretatsioon on mitterahuldav oma keerulisuse ja eklektilisuse tõttu. Wigneri interpretatsiooni teevad üpris elegantseks kaks dünaamilist alusseadust, mis peegeldavad kvantmehhaanika arvutusmeetodit, kui õnnestub kõik üksikasjad välja töötada. Kuid Everetti vaade on teistest tunduvalt lihtsam. Selles pole muid postulaate peale Schrödingeri võrrandi — printsiibi, mida tunnustavad kõik kvantmehhaanika interpretatsioonid. Peale selle on Everetti interpretatsioonil kaks eelist, mis teistel interpretatsioonidel puuduvad: esiteks on ta täiesti lokaalne teooria, ja teiseks on ta otseselt ühitatav relatiivsusteooriaga.

Samuti tasub mainida, et Everetti interpretatsiooni intuitsioonivastaseid elemente võib leida ka mõlemas teises interpretatsioonis. Wigneri järgi tuleb leppida sellega, et universumi ajaline muutumine on kulgenud hiiglaslikus Everetti superpositsioonis, nii et selles on võinud leiduda superpositsioonilisi kinnistähti ja superpositsioonilisi kive, kui mitte superpositsioonilisi kasse — ja seda vähemalt esimese teadvusega olendi ilmumiseni, ​**[lk.357]**​ kes kutsus esile lainefunktsiooni kollapsi. Bohmi järgi figureerib Everetti kollabeerumata lainefunktsioon osakeste asendeid suunava “pilootlaine” kujul. Kogu teistes komponentides leiduv struktuur sisaldub seega ikkagi maailma seisundis​[[414]](#footnote-415), kuigi suurem osa sellest struktuurist ei ole osakeste ajalisele muutumisele relevantne. Asjaolu, et ka need kontseptsioonid annavad keskse rolli kollabeerumata lainefunktsioonile, lubab öelda, et Everetti interpretatsioon ei olegi nendega võrreldes nii ebausutav.

Loomulikult ei saa välistada, et kunagi töötatakse välja uus teooria, mis on nii usutavuselt kui ka teoreetiliselt ladususelt kõigist senistest teooriatest üle. Kuid see ei tundu eriti tõenäoline. Eksperimentaalsete anomaaliate täieliku puudumise tõttu võib arvata, et kvantmehhaanika arvutusmeetodid jäävad ennustusliku teooriana püsima. Ja sellisel juhul ei maksa oodata, et probleemile tooksid lahenduse empiirilised tegurid. Pole võimatu, et kontseptuaalse töö käigus sünnib mõni uus ja parem interpretatsioon, kuid samas võivad kontseptuaalse ruumi kõige perspektiivsemad nišid olla nüüdseks juba hõivatud. Sellisel juhul peamegi võib-olla leppima valikute spektriga, mis ei erine kuigi palju praegusest; valikutesse võib tulla olulisi täiustusi, kuid eelised ja puudused jäävad kvalitatiivselt samasugusteks. Everetti interpretatsioon tundub neist valikutest mitmes mõttes kõige ahvatlevamana, kuid samas on just seda interpretatsiooni kõige raskem omaks võtta.

Mõned seisukohad, mida ma olen käesolevas töös põhjendada püüdnud, on intuitsioonivastased. Ma panin keha–vaimu dualismile kaua vastu, kuid olen nüüdseks jõudnud punkti, kus ma seda aktsepteerin, kusjuures mitte ainult kui ainukest vastuvõetavat, vaid ka kui iseenesest rahuldavat seisukohta. Alati on võimalik, et ma olen midagi segi ajanud või et leidub mingi uus ja radikaalne võimalus, mida ma pole märganud, kuid ma võin ebamugavust tundmata öelda, et minu arvates peab dualism väga tõenäoliselt paika. Olen ühe võimalusena vaadelnud ka mingit laadi panpsühhismi. Nagu keha–vaimu dualism, tundub ka see esialgu intuitsioonivastasena, kuid ajapikku hakkab tema intuitsioonivastasus kahanema. Ma ei oska öelda, kas panpsühhism on õige või väär, kuid ta on vähemasti intellektuaalselt ahvatlev ja, kui järele mõelda, siis ka mitte liiga pöörane, et teda vastuvõetavaks pidada.

Everetti interpretatsiooni pöörasus on teist suurusjärku. Intellektuaalselt on see kvantmehhaanika kõikvõimalikest interpretatsioonidest kõige ahvatlevam, kuid ma tunnistan, et südamepõhjas ei suuda ma teda uskuda. Kui Jumal sunniks mind panema oma elu panti nende vaadete tõesuse või vääruse peale, mida ma olen püüdnud põhjendada, julgeksin olla üsna kindel elamuse fundamentaalsuses ja vähem kindel elamuse kõikjalolus. Everetti interpretatsiooni suhtes jääksin aga kahevahele ja viimaks vist siiski ei julgeks sellele panust teha.[[415]](#footnote-416) Võib-olla on see interpretatsioon lihtsalt liiga kummaline, et teda uskuda. Siiski ei ole selge, kui suurt rõhku neile intuitiivsetele kahtlustele tuleks kokkuvõttes panna. Everetti interpretatsioon on lihtne ja elegantne ning ennustab vaatlejate olemasolu, kes näevad maailma just nii, nagu näen maailma mina. Kas sellest ei piisa? Võib-olla ei hakkagi me iial suutma seda seisukohta emotsionaalselt omaks võtta, kuid me peaksime vähemalt võtma tõsiselt võimalust, et ta osutub tõeseks.

### ​**[lk.391]**​ Bibliograafia[[416]](#footnote-417)

Ackerman, D. 1990. *A Natural History of the Senses*​. New York: Random House.

Adams, R. M. 1974. Theories of actuality. *Nous*​ 8:211–31.

Akins, K. 1993. What is it like to be boring and myopic? In B. Dahlbom, ed., ​*Dennett and His Critics*​. Oxford: Blackwell.

Albert, D. 1992. *Quantum Mechanics and Experience*​. Cambridge, Mass: Harvard University Press.

Albert, D., and B. Loewer. 1988. Interpreting the many-worlds interpretation. *Synthese*​ 77:195–213.

————————. 1989. Two no-collapse interpretations of quantum mechanics. *Nous*​ 23: 169–86.

————————. 1990. Wanted dead or alive: Two attempts to solve Schrodinger’s paradox. *PSA 1990*​, vol. 1, pp. 277–85.

Alexander, S. 1920. *Space, Time, and Deity*​. London: Macmillan.

Armstrong, D. M. 1968. *A Materialist Theory of the Mind*​. London: Routledge and Kegan Paul.

————————. 1973. *Belief, Truth, and Knowledge*​. Cambridge: Cambridge University Press.

————————. 1981. What is consciousness? In ​*The Nature of Mind*​. Ithaca, N.Y.: Cornell University Press.

————————. 1982. Metaphysics and supervenience. *Critica*​ 42:3–17.

————————. 1983. *What Is a Law of Nature?* Cambridge: Cambridge University Press.

————————. 1990. *A Combinatorial Theory of Possibility*​. Cambridge: Cambridge University Press.

Austin, D. F. 1990. *What’s the Meaning of “This”?* Ithaca, N.Y.: Cornell University Press.

Baars, B. J. 1988. *A Cognitive Theory of Consciousness*​. Cambridge: Cambridge University Press.

Bacon, J. 1986. Supervenience, necessary coextension, and reducibility. *Philosophical Studies*​ 49:163–76.

Barwise, J., and J. Perry. 1983. *Situations and Attitudes*​. Cambridge, Mass.: MIT Press.

Bateson, G. 1972. *Steps to an Ecology of Mind*​. San Francisco: Chandler.

Bealer, G. 1994. Mental properties. *Journal of Philosophy*​ 91:185–208.

Bell, J. S. 1964. On the Einstein–Podolsky–Rosen paradox. *Physics*​ 1:195–200. [Reprinted in Bell 1987b]

————————. 1976. The measurement theory of Everett and de Broglie’s pilot wave. In M. Flato, ed., ​*Quantum Mechanics, Determinism, Causality, and Particles*​. Dordrecht: Reidel. [Reprinted in Bell 1987b]

​**[lk.392]**​ ————————. 1981. Quantum mechanics for cosmologists. In C. Isham, R. Penrose, and D. Sciama, eds., ​*Quantum Gravity*​. Vol. 2. Oxford: Oxford University Press. [Reprinted in Bell 1987b]

————————. 1987a. Are there quantum jumps? In ​*Schrodinger: Centenary of a Polymath*​. Cambridge: Cambridge University Press.

————————. 1987b. *Speakable and Unspeakable in Quantum Mechanics*​. Cambridge: Cambridge University Press.

Bigelow, J., and R. Pargetter. 1990. Acquaintance with qualia. *Theoria*​ 56:129–47.

Bisiach, E. 1988. The (haunted) brain and consciousness. In A. Marcel and E. Bisiach, eds., ​*Consciousness in Contemporary Science*​. Oxford: Oxford University Press.

Blackburn, S. 1971. Moral realism. In J. Casey, ed., ​*Morality and Moral Reasoning*​. London: Methuen.

————————. 1985. Supervenience revisited. In I. Hacking, ed., ​*Exercises in Analysis: Essays by Students of Casimir Lewy*​. Cambridge: Cambridge University Press.

————————. 1990. Filling in space. *Analysis*​ 50:62–65.

Block, N. 1978. Troubles with functionalism. In C. W. Savage, ed., ​*Perception and Cognition: Issues in the Foundation of Psychology*​. Minneapolis: University of Minnesota Press. [Reprinted in N. Block, ed., ​*Readings in the Philosophy of Psychology*​. Vol. 1. Cambridge, Mass.: Harvard University Press, 1980]

————————. 1980. What is functionalism? In N. Block, ed., ​*Readings in the Philosophy of Psychology*​. Vol. 1. Cambridge, Mass.: Harvard University Press.

————————. 1981. Psychologism and behaviorism. *Philosophical Review*​ 90:5–43.

————————. 1990. Inverted earth. *Philosophical Perspectives*​ 4:53–79.

————————. 1995. On a confusion about a function of consciousness. *Behavioral and Brain Sciences*​ 18:227–47.

Boden, M. 1988. Escaping from the Chinese Room. In ​*Computer Models of Mind*​. Cambridge: Cambridge University Press.

Bogen, J. 1981. Agony in the schools. *Canadian Journal of Philosophy*​ 11:1–21.

Bohm, D. 1952. A suggested interpretation of quantum mechanics in terms of “hidden variables,” pts. 1 and 2. *Physical Review*​ 85:166–93.

————————. 1980. *Wholeness and the Implicate Order*​. London: Routledge.

Bohm, D., and B. Hiley. 1993. *The Undivided Universe: An Ontological Interpretation of Quantum Theory*​. London: Routledge.

Boring, E. G. 1942. *Sensation and Perception in the History of Experimental Psychology*​. New York: Appleton–Century–Crofts.

Boyd, R. N. 1980. Materialism without reductionism: What physicalism does not entail. In N. Block, ed., ​*Readings in the Philosophy of Psychology*​. Vol. 1. Cambridge, Mass.: Harvard University Press.

————————. 1988. How to be a moral realist. In G. Sayre-McCord, ed., ​*Essays on Moral Realism*​. Ithaca, N.Y.: Cornell University Press.

Brindley, G. S. 1960. *Physiology of the Retina and Visual Pathway*​. London: Edward Arnold.

Brink, D. 1989. *Moral Realism and the Foundations of Ethics*​. Cambridge: Cambridge University Press.

Broad, C. D. 1925. *Mind and Its Place in Nature*​. London: Routledge and Kegan Paul.

Brooks, D. H. M. 1994. How to perform a reduction. *Philosophy and Phenomenological Research*​ 54:803–14.

​**[lk.393]**​ Byrne, A. 1993. The emergent mind. Ph.D. diss., Princeton University.

Campbell, K. K. 1970. *Body and Mind*​. New York: Doubleday.

Carroll, J. W. 1990. The Humean tradition. *Philosophical Review*​ 99:185–219.

————————. 1994. *Laws of Nature*​. Cambridge: Cambridge University Press.

Carruthers, P. 1992. Consciousness and concepts. *Proceedings of the Aristotelian Society*​, suppl., 66:41–59.

Chalmers, D. J. 1990. Consciousness and cognition. Technical Report 38, Center for Research on Concepts and Cognition, Indiana University.

————————. 1994a. On implementing a computation. *Minds and Machines*​ 4:391–402.

————————. 1994b. A computational foundation for the study of cognition. PNP Technical Report 94-03, Washington University.

————————. 1994c. The components of content. PNP Technical Report 94-04, Washington University, [[http:​//​ling.vcsc.edu​/~chalmers/​papers/content.html](http:///%E2%80%8B//%E2%80%8Bling.vcsc.edu%E2%80%8B/~chalmers/%E2%80%8Bpapers/content.html)]

————————. 1995a. Does a rock implement every finite state automation? *Synthese*​.[[417]](#footnote-418)

————————. 1995b. Facing up to the problem of consciousness. *Journal of Consciousness Studies*​ 2:200–219. [Also in S. Hameroff, A. Kaszniak, and A. Scott, eds., ​*Toward a Science of Consciousness*​. Cambridge, Mass.: MIT Press, 1996]

————————. 1995c. Minds, machines, and mathematics. *PSYCHE*​ 2:1.

————————. 1995d. The puzzle of conscious experience. *Scientific American*​ 273:80–86.

Cheney, D. L., and R. M. Seyfarth. 1990. *How Monkeys See the World*​. Chicago: University of Chicago Press.

Chisholm, R. 1957. *Perceiving*​. Ithaca, N.Y.: Cornell University Press.

Churchland, P. M. 1985. Reduction, qualia and the direct introspection of brain states. *Journal of Philosophy*​ 82:8–28.

————————. 1995. *The Engine of Reason, the Seat of the Soul: A Philosophical Journey into the Brain*​. Cambridge, Mass.: MIT Press.

Churchland, P. M., and P. S. Churchland. 1981. Functionalism, qualia and intentionality. *Philosophical Topics*​ 12:121–32.

Churchland, P. S. 1986. *Neurophilosophy: Toward a Unified Science of the MindBrain*​. Cambridge, Mass.: MIT Press.

————————. 1988. The significance of neuroscience for philosophy. *Trends in the Neurosciences*​ 11:304–7.

Clark, A. 1986. Psychofunctionalism and chauvinism. *Philosophy of Science*​ 53: 535–59.

————————. 1993. *Sensory Qualities*​. Oxford: Oxford University Press.

Cole, D. 1990. Functionalism and inverted spectra. *Synthese*​ 82:207–22.

Conee, E. 1985a. Physicalism and phenomenal properties. *Philosophical Quarterly*​ 35:296–302.

————————. 1985b. The possibility of absent qualia. *Philosophical Review*​ 94:345–66.

Cowey, A., and P. Stoerig. 1992. Reflections on blindsight. In D. Milner and M. Rugg, eds., *The Neuropsychology of Consciousness*​. London: Academic Press.

Crane, T. 1992. The nonconceptual content of experience. In T. Crane, ed., ​*The Contents of Experience*​. Cambridge: Cambridge University Press.

Crick, F. H. C. 1994. *The Astonishing Hypothesis: The Scientific Search for the Soul*​. New York: Scribner.

Crick, F. H. C., and C. Koch. 1990. Towards a neurobiological theory of consciousness. *Seminars in the Neurosciences*​ 2:263–75.

Cuda, T. 1985. Against neural chauvinism. *Philosophical Studies*​ 48:111–27.

​**[lk.394]**​ Cussins, A. 1990. The connectionist construction of concepts. In M. Boden, ed., ​*The Philosophy of Artificial Intelligence*​. Oxford: Oxford University Press.

Daneri, A., A. Loinger, and G. M. Prosperi. 1962. Quantum theory of measurement and ergodicity conditions. *Nuclear Physics*​ 33:297–319. [Reprinted in Wheeler and Zurek 1983]

Davidson, D. 1970. Mental events. In L. Foster and J. Swanson, eds., ​*Experience and Theory*​. London: Duckworth.

Davies, M. K., and I. L. Humberstone. 1980. Two notions of necessity. *Philosophical Studies*​ 38:1–30.

Dennett, D. C. 1969. *Content and Consciousness*​. London: Routledge and Kegan Paul.

————————. 1978a. *Brainstorms*​. Cambridge, Mass.: MIT Press.

————————. 1978b. Are dreams experiences? In Dennett 1978a.

————————. 1978c. Toward a cognitive theory of consciousness. In Dennett 1978a.

————————. 1978d. Where am I? In Dennett 1978a.

————————. 1979. On the absence of phenomenology. In D. Gustafson and B. Tapscott, eds., ​*Body, Mind, and Method*​. Dordrecht: Kluwer.

————————. 1987. *The Intentional Stance*​. Cambridge, Mass.: MIT Press.

————————. 1988. Quining qualia. In A. Marcel and E. Bisiach, eds., ​*Consciousness in Contemporary Science*​. Oxford: Oxford University Press.

————————. 1991. *Consciousness Explained*​. Boston: Little, Brown.

————————. 1993a. Back from the drawing board. In B. Dahlbom, ed., ​*Dennett and His Critics*​. Oxford: BlackweU.

————————. 1993b. The message is: There is no medium. *Philosophy and Phenomenological Research*​ 53:919–31.

Descartes, R. 1984. *The Philosophical Writings of Descartes*​. Translated by J. Cottingham, R. Stoothoff, and D. Murdoch. Cambridge: Cambridge University Press.

DeWitt, B. S. 1970. Quantum mechanics and reality. *Physics Today*​ 23:30–35. [Reprinted in DeWitt and Graham 1973]

————————. 1971. The many-universes interpretation of quantum mechanics. In B. d’Espagnat, ed., ​*Foundations of Quantum Mechanics*​. New York: Academic Press. [Reprinted in DeWitt and Graham 1973]

DeWitt, B. S., and N. Graham, eds., 1973. *The Many-Worlds Interpretation of Quantum Mechanics*​. Princeton: Princeton University Press.

Dretske, F. I. 1977. Laws of nature. *Philosophy of Science*​ 44:248–68.

————————. 1981. *Knowledge and the Flow of Information*​. Cambridge, Mass.: MIT Press.

————————. 1995. *Naturalizing the Mind*​. Cambridge, Mass.: MIT Press.

Dreyfus, H. 1972. *What Computers Can’t Do*​. New York: Harper & Row.

Eccles, J. C. 1986. Do mental events cause neural events analogously to the probability fields of quantum mechanics? *Proceedings of the Royal Society of London*​ B227:411–28.

Edelman, G. 1989. *The Remembered Present: A Biological Theory of Consciousness*​. New York: Basic Books.

————————. 1992. *Bright Air, Brilliant Fire*​. New York: Basic Books.

Elitzur, A. 1989. Consciousness and the incompleteness of the physical explanation of behavior. *Journal of Mind and Behavior*​ 10:1–20.

Evans, G. 1979. Reference and contingency. *The Monist*​ 62:161–89.

​**[lk.395]**​ ————————. 1982. *The Varieties of Reference*​. Oxford: Oxford University Press.

Everett, H. 1957. “Relative-state” formulations of quantum mechanics. *Reviews of Modern Physics*​ 29:454–62. [Reprinted in Wheeler and Zurek 1983]

————————. 1973. The theory of the universal wave function. In B. S. deWitt and N. Graham, eds., *The Many-Worlds Interpretation of Quantum Mechanics*​. Princeton: Princeton University Press.

Farah, M. 1994. Visual perception and visual awareness after brain damage: A tutorial overview. In C. Umilta and M. Moscovitch, eds., ​*Conscious and Nonconscious Information Processing: Attention and Performance*​ 15. Cambridge, Mass.: MIT Press.

Farrell, B. A. 1950. Experience. *Mind*​ 59:170–98.

Feigl, H. 1958. The “mental” and the “physical.” In H. Feigl, M. Scriven, and G. Maxwell, eds., ​*Concepts, Theories, and the Mind–Body Problem*​. Minnesota Studies in the Philosophy of Science , vol. 2. Minneapolis: University of Minnesota Press.

Feldman, F. 1974. Kripke on the identity theory. *Journal of Philosophy*​ 71:665–76.

Field, H. 1973. Theory change and the indeterminacy of reference. *Journal of Philosophy*​ 40:762–81.

Flanagan, O. 1992. *Consciousness Reconsidered*​. Cambridge, Mass.: MIT Press.

Fodor, J. A. 1980. Searle on what only brains can do. *Behavioral and Brain Sciences*​ 3:431–32.

————————. 1987. *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*​. Cambridge, Mass.: MIT Press.

————————. 1992. The big idea: Can there be a science of mind? *Times Literary Supplement*​, July 3, pp. 5–7.

Forrest, P. 1986. Ways worlds could be. *Australasian Journal of Philosophy*​ 64:15–24.

Foss, J. 1989. On the logic of what it is like to be a conscious subject. Australasian Journal of Philosophy 67:305–20.

Foster, J. 1991. *The Immaterial Self: A Defence of the Cartesian Dualism Conception of Mind*​. London: Routledge.

Fredkin, E. 1990. Digital mechanics. *Physica*​ D45:254–70.

Geach, P. 1957. *Mental Acts*​. London: Routledge and Kegan Paul.

Gell-Mann, M., and J. B. Hartle. 1990. Quantum mechanics in the light of quantum cosmology. In W. Zurek, ed., ​*Complexity, Entropy, and the Physics of Information*​. Redwood City, Calif.: Addison–Wesley.

Gert, B. 1965. Imagination and verifiability. *Philosophical Studies*​ 16:44–47.

Ghirardi, G. C., A. Rimini, and T. Weber. 1986. Unified dynamics for microscopic and macroscopic systems. *Physical Review*​ D34:470.

Goldman, A. 1986. *Epistemology and Cognition*​. Cambridge, Mass.: Harvard University Press.

————————. 1993. The psychology of folk psychology. *Behavioral and Brain Sciences*​ 16:15–28.

Gunderson, K. 1970. Asymmetries and mind–body perplexities. In M. Radner and S. Winokur, eds., ​*Analyses of Theories and Methods of Physics and Psychology*​. Minnesota Studies in the Philosophy of Science, vol. 4. Minneapolis: University of Minnesota Press.

Hameroff, S. R. 1994. Quantum coherence in microtubules: A neural basis for an emergent consciousness? *Journal of Consciousness Studies*​ 1:91–118.

​**[lk.396]**​ Hardin, C. L. 1987. Qualia and materialism: Closing the explanatory gap. *Philosophy and Phenomenological Research*​ 48:281–98.

————————. 1988. *Color for Philosophers: Unweaving the Rainbow*​. Indianapolis: Hackett.

Hare, R. M. 1952. *The Language of Morals*​. Oxford: Clarendon Press.

————————. 1984. Supervenience. *Proceedings of the Aristotelian Society*​, suppl., 58:1–16.

Harman, G. 1982. Conceptual role semantics. *Notre Dame Journal of Formal Logic*​ 28:242–56.

————————. 1990. The intrinsic quality of experience. *Philosophical Perspectives*​ 4:31–52.

Harnad, S. 1989. Minds, machines and Searle. *Journal of Experimental and Theoretical Artificial Intelligence*​ 1:5–25.

Harrison, B. 1967. On describing colors. *Inquiry*​ 10:38–52.

————————. 1973. *Form and Content*​. Oxford: Blackwell.

Harrison, J. 1981a. Three philosophical fairy stories. *Ratio*​ 23:63–67.

————————. 1981b. Gulliver’s adventures in Fairyland. *Ratio*​ 23:158–64.

Haugeland, J. 1980. Programs, causal powers, and intentionality. *Behavioral and Brain Sciences*​ 4:432–33.

————————. 1982. Weak supervenience. *American Philosophical Quarterly*​ 19:93–103.

Healey, R. A. 1984. How many worlds? *Nous*​ 18:591–616.

Heil, J. 1992. *The Nature of True Minds*​. Cambridge: Cambridge University Press.

Hellman, G., and F. Thompson. 1975. Physicalism: Ontology, determination and reduction. *Journal of Philosophy*​ 72:551–64.

Hill, C. S. 1991. *Sensations: A Defense of Type Materialism*​. Cambridge: Cambridge University Press.

Hodgson, D. 1988. *The Mind Matters: Consciousness and Choice in a Quantum World*​. Oxford: Oxford University Press.

Hofstadter, D. R. 1979. *Gödel, Escher, Bach: an Eternal Golden Braid*​. New York: Basic Books.

————————. 1981. Reflections on Searle. In D. R. Hofstadter and D. C. Dennett, eds., ​*The Mind’s I*​. New York: Basic Books.

————————. 1985a. Who shoves whom around inside the careenium? In ​*Metamagical Themas*​. New York: Basic Books

————————. 1985b. Heisenberg’s Uncertainty Principle and the many-worlds interpretation of quantum mechanics. In ​*Metamagical Themas*​. New York: Basic Books.

Honderich, T. 1981. Psychophysical law-like connections and their problems. *Inquiry*​ 24:277–303.

Horgan, T. 1978. Supervenient bridge laws. *Philosophy of Science*​ 45:227–49.

————————. 1982. Supervenience and microphysics. *Pacific Philosophical Quarterly*​ 63: 29–43.

————————. 1984a. Functionalism, qualia, and the inverted spectrum. *Philosophy and Phenomenological Research*​ 44:453–69.

————————. 1984b. Jackson on physical information and qualia. *Philosophical Quarterly*​ 34:147–83.

————————. 1984c. Supervenience and cosmic hermeneutics. *Southern Journal of Philosophy*​, suppl., 22:19–38.

————————. 1987. Supervenient qualia. *Philosophical Review*​ 96:491–520.

————————. 1993. From supervenience to superdupervenience: Meeting the demands of a material world. *Mind*​ 102:555–86.

Horgan, T., and M. Timmons. 1992a. Troubles for new wave moral semantics: The “open question argument” revived. *Philosophical Papers*​.

​**[lk.397]**​ ————————. 1992b. Trouble on moral twin earth: Moral queerness revived. *Synthese*​ 92:223–60.

Horst, S. 1995. Phenomenology and psychophysics. Manuscript, Wesleyan University.

Hughes, R. I. G. 1989. *The Structure and Interpretation of Quantum Mechanics*​. Cambridge, Mass.: Harvard University Press.

Humphrey, N. 1992. *A History of the Mind: Evolution and the Birth of Consciousness*​. New York: Simon and Schuster.

Huxley, T. 1874. On the hypothesis that animals are automata. In ​*Collected Essays*​. London, 1893–94.

Jackendoff, R. 1987. *Consciousness and the Computational Mind*​. Cambridge, Mass.: MIT Press.

Jackson, F. 1977. *Perception*​. Cambridge: Cambridge University Press.

————————. 1980. A note on physicalism and heat. *Australasian Journal of Philosophy*​ 58:26–34.

————————. 1982. Epiphenomenal qualia. *Philosophical Quarterly*​ 32:127–36.

————————. 1993. Armchair metaphysics. In J. O’Leary-Hawthorne and M. Michael, eds., ​*Philosophy in Mind*​. Dordrecht: Kluwer.

————————. 1994. Finding the mind in the natural world. In R. Casati, B. Smith, and G. White, eds., ​*Philosophy and the Cognitive Sciences*​. Vienna: Holder–Pichler–Tempsky.

————————. 1995. Postscript to “What Mary didn’t know.” In P. K. Moser and J. D. Trout, eds., ​*Contemporary Materialism*​. London: Routledge.

Jacoby, H. 1990. Empirical functionalism and conceivability arguments. *Philosophical Psychology*​ 2:271–82.

Jaynes, J. 1976. *The Origins of Consciousness in the Breakdown of the Bicameral Mind*​. Boston: Houghton Mifflin.

Johnson-Laird, P. 1983. A computational analysis of consciousness. *Cognition and Brain Theory*​ 6:499–508.

Kaplan, D. 1979. *Dthat*​. In P. Cole, ed., ​*Syntax and Semantics*​. New York: Academic Press.

————————. 1989. Demonstratives. In J. Almog, J. Perry, and H. Wettstein, ed., ​*Themes from Kaplan*​. New York: Oxford University Press.

Kim, J. 1978. Supervenience and nomological incommensurables. *American Philosophical Quarterly*​ 15:149–56.

————————. 1984. Concepts of upervenience. *Philosophy and Phenomenological Research*​ 45:153–76.

————————. 1985. Psychophysical laws. In B. McLaughlin and E. LePore, eds., ​*Action and Events*​. Oxford: Blackwell.

————————. 1989. Mechanism, purpose, and explanatory exclusion. *Philosophical Perspectives*​ 3:77–108.

————————. 1993. *Supervenience and Mind*​. Cambridge: Cambridge University Press.

Kirk, R. 1974. Zombies versus materialists. *Aristotelian Society*​ 48(suppl.):13552.

————————. 1979. From physical explicability to full-blooded materialism. *Philosophical Quarterly*​ 29:229–37.

————————. 1992. Consciousness and concepts. *Proceedings of the Aristotelian Society*​ 66(suppl.):23–40.

————————. 1994. *Raw Feeling: A Philosophical Account of the Essence of Consciousness*​. Oxford: Oxford University Press.

​**[lk.398]**​ Korb, K. 1991. Searle’s AI program. *Journal of Experimental and Theoretical Artificial Intelligence*​ 3:283–96

Kripke, S. A. 1971. Identity and necessity. In M. Munitz, ed., ​*Identity and Individuation*​. New York: New York University Press.

————————. 1972. Naming and necessity. In G. Harman and D. Davidson, eds., ​*The Semantics of Natural Language*​. Dordrecht: Reidel. [Reprinted as Kripke 1980]

————————. 1980. *Naming and Necessity*​. Cambridge, Mass.: Harvard University Press.

————————. 1982. *Wittgenstein on Rule-Following and Private Language*​. Cambridge, Mass.: Harvard University Press.

Lahav, R., and N. Shanks. 1992. How to be a scientifically respectable “property dualist.” *Journal of Mind and Behavior*​ 13:211–32.

Langton, C. G. 1989. *Artificial Life: The Proceedings of an Interdisciplinary Workshop on the Synthesis and Simulation of Living Systems*​. Redwood City, Calif.: Addison–Wesley.

Leckey, M. 1993. The universe as a computer: A model for prespace metaphysics. Manuscript. Philosophy Department, Monash University.

Levine, J. 1983. Materialism and qualia: The explanatory gap. *Pacific Philosophical Quarterly*​ 64:354–61.

————————. 1988. Absent and inverted qualia revisited. *Mind and Language*​ 3:271–87.

————————. 1991. Cool red. *Philosophical Psychology*​ 4:27–40.

————————. 1993. On leaving out what it’s like. In M. Davies and G. Humphreys, eds., ​*Consciousness: Psychological and Philosophical Essays*​. Oxford: Blackwell.

Lewis, D. 1966. An argument for the identity theory. *Journal of Philosophy*​ 63:17–25.

————————. 1972. Psychophysical and theoretical identifications. *Australasian Journal of Philosophy*​ 50:249–58.

————————. 1973. *Counterfactuals*​. Cambridge, Mass.: Harvard University Press.

————————. 1974. Radical interpretation. *Synthese*​ 23:331–44.

————————. 1979. Attitudes *de dicto* and *de se*​. *Philosophical Review*​ 88:513–45.

————————. 1983a. Extrinsic properties. *Philosophical Studies*​ 44:197–200.

————————. 1983b. New work for a theory of universals. *Australasian Journal of Philosophy*​ 61:343–77.

————————. 1986a. *On the Plurality of Worlds*​. Oxford: Blackwell.

————————. 1986b. *Philosophical Papers*​. Vol. 2. New York: Oxford University Press.

————————. 1990. What experience teaches. In W. Lycan, ed., ​*Mind and Cognition*​. Oxford: Blackwell.

————————. 1994. Reduction of mind. In S. Guttenplan, ed., ​*A Companion to the Philosophy of Mind*​. Oxford: Blackwell.

Libet, B. 1993. The neural time factor in conscious and unconscious events. In ​*Experimental and Theoretical Studies of Consciousness*​. Ciba Foundation Symposium 174. New York: Wiley.

Loar, B. 1990. Phenomenal states. *Philosophical Perspectives*​ 4:81–108.

Lockwood, M. 1989. *Mind, Brain, and the Quantum*​. Oxford: Blackwell.

————————. 1992. The grain problem. In H. Robinson, ed., ​*Objections to Physicalism*​. Oxford: Oxford University Press.

Logothetis, N., and J. D. Schall. 1989. Neuronal correlates of subjective visual perception. *Science*​ 245:761–63.

London, F., and E. Bauer. 1939. The theory of observation in quantum mechanics (in French). *Actualités scientifiques et industrielles*​, no. 775. [English translation in Wheeler and Zurek 1983]

​**[lk.399]**​ Lucas, J. R. 1961. Minds, machines and Godel. *Philosophy*​ 36:112–27.

Lycan, W. G. 1973. Inverted spectrum. *Ratio*​ 15:315–19.

————————. 1987. *Consciousness*​. Cambridge, Mass.: MIT Press.

————————. 1995. A limited defense of phenomenal information. In T. Metzinger, ed., ​*Conscious Experience*​. Paderborn: Schoningh.

————————. 1996. *Consciousness and Experience*​. Cambridge, Mass.: MIT Press.

Mackay, D. M. 1969. *Information, Mechanism, and Meaning*​. Cambridge, Mass.: MIT Press.

Mackie, J. L. 1974. *The Cement of the Universe*​. Oxford: Oxford University Press.

————————. 1977. *Ethics: Inventing Right and Wrong*​. Harmondsworth: Penguin Books.

Marks, L. E. 1978. *The Unity of the Senses: Interrelations among the Modalities*​. New York: Academic Press.

Matzke, D., ed. 1993. *Proceedings of the 1992 Workshop on Physics and Computation*​. Los Alamitos, Calif.: IEEE Computer Society Press.

————————, ed. 1995. *Proceedings of the 1994 Workshop on Physics and Computation*​. Los Alamitos, Calif.: IEEE Computer Society Press.

Maxwell, G. 1978. Rigid designators and mind–brain identity. In C. W. Savage, ed., ​*Perception and Cognition: Issues in the Foundations of Psychology*​. Minnesota Studies in the Philosophy of Science, vol. 9. Minneapolis: University of Minnesota Press.

McCarthy, J. 1979. Ascribing mental qualities to machines. In M. Ringle, ed., ​*Philosophical Perspectives in Artificial Intelligence*​. Atlantic Highlands, N.J.: Humanities Press.

McDowell, J. 1994. *Mind and World*​. Cambridge, Mass.: Harvard University Press.

McGinn, C. 1977. Anomalous monism and Kripke’s Cartesian intuitions. *Analysis*​ 2:78–80.

————————. 1989. Can we solve the mind–body problem? *Mind*​ 98:349–66.

McLaughlin, B. P. 1992. The rise and fall of the British emergentists. In A. Beckermann, H. Flohr, and J. Kim, eds., ​*Emergence or Reduction? Prospects for Nonreductive Physicalism*​. Berlin: De Gruyter.

————————. 1995. Varieties of supervenience. In E. E. Savellos and U. D. Yalcin, eds., ​*Supervenience: New Essays*​. Cambridge: Cambridge University Press.

McMullen, C. 1985. “Knowing what it’s like” and the essential indexical. *Philosophical Studies*​ 48:211–33.

Meehl, P.E., and W. Sellars. 1956. The concept of emergence. In H. Feigl and M. Scriven, eds., ​*The Foundations of Science and the Concept of Psychology and Psychoanalysis*​. Minnesota Studies in the Philosophy of Science, vol. 1. Minneapolis: University of Minnesota Press.

Molnar, G. 1969. Kneale’s argument revisited. *Philosophical Review*​ 78:79–89.

Moore, G. E. 1922. *Philosophical Studies*​. London: Routledge and Kegan Paul.

Müller, G. E. 1896. Zur Psychophysik der Gesichtsempfindungen. *Zeitschrift fur Psychologic und Physiologic der Sinnesorgane*​ 10:1–82.

Nagel, T. 1970. Armstrong on the mind. *Philosophical Review*​ 79:394–403.

————————. 1974. What is it like to be a bat? *Philosophical Review*​ 4:435–50.

————————. 1983. The objective self. In C. Ginet and S. Shoemaker, eds., ​*Knowledge and Mind: Philosophical Essays*​. New York: Oxford University Press.

————————. 1986. *The View from Nowhere*​. New York: Oxford University Press.

Natsoulas, T. 1978. Consciousness. *American Psychologist*​ 33:906–14.

Nelkin, N. 1989. Unconscious sensations. *Philosophical Psychology*​ 2:129–41.

​**[lk.400]**​ ————————. 1993. What is consciousness? *Philosophy of Science*​ 60:419–34.

Nemirow, L. 1990. Physicalism and the cognitive role of acquaintance. In W. Lycan, ed., ​*Mind and Cognition*​. Oxford: Blackwell.

Newell, A. 1992. SOAR as a unified theory of cognition: Issues and explanations. *Behavioral and Brain Sciences*​ 15:464–92.

Newton, N. 1989. Machine understanding and the Chinese Room. *Philosophical Psychology*​ 2:207–15.

Nida-Rümelin, M. 1995. What Mary couldn’t know: Belief about phenomenal states. In T. Metzinger, ed., ​*Conscious Experience*​. Paderborn: Schoningh.

————————. 1996. Pseudonormal vision: An actual case of qualia inversion? *Philosophical Studies*​.

Papineau, D. 1993. *Philosophical Naturalism*​. Oxford: Blackwell.

Parfit, D. 1984. *Reasons and Persons*​. Oxford: Oxford University Press.

Peacocke, C. 1992. Scenarios, concepts, and perception. In T. Crane, ed., ​*The Contents of Experience*​. Cambridge: Cambridge University Press.

Penrose, R. 1987. Quantum physics and conscious thought. In B. Hiley and Peat, eds., ​*Quantum Implications: Essays in Honor of David Bohm*​. New York: Methuen.

————————. 1989. *The Emperor’s New Mind*​. Oxford: Oxford University Press.

————————. 1994. *Shadows of the Mind*​. Oxford: Oxford University Press.

Perry, J. 1979. The problem of the essential indexical. *Nous*​ 13:3–21.

Petrie, B. 1987. Global supervenience and reduction. *Philosophy and Phenomenological Research*​ 48:119–30.

Place, U. T. 1956. Is consciousness a brain process? British Journal of Psychology 47:44–50.

Plantinga, A. 1976. Actualism and possible worlds. *Theoria*​ 42:139–60.

Putnam, H. 1960. Minds and machines. In S. Hook, ed., ​*Dimensions of Mind*​. New York: New York University Press.

————————. 1975. The meaning of “meaning.” In K. Gunderson, ed., ​*Language, Mind, and Knowledge*​. Minneapolis: University of Minnesota Press.

————————. 1981. *Reason, Truth, and History*​. Cambridge: Cambridge University Press.

————————. 1983. Possibility and necessity. In ​*Philosophical Papers*​. Vol. 3. Cambridge: Cambridge University Press.

————————. 1988. *Representation and Reality*​. Cambridge, Mass.: MIT Press.

Pylyshyn, Z. 1980. The “causal power” of machines. *Behavioral and Brain Sciences*​ 3:442–44.

Quine, W. V. 1951. Two dogmas of empiricism. *Philosophical Review*​ 60:20–43.

————————. 1969. Propositional objects. In ​*Ontological Relativity and Other Essays*​. New York: Columbia University Press.

Rensink, R. A., J. K. O’Regan, and J. J. Clark. 1995. Image flicker is as good as saccades in making large scene changes invisible. *Perception*​ 24 (suppl.):26–27.

Rey, G. 1982. A reason for doubting the existence of consciousness. In R. Davidson, S. Schwartz, and D. Shapiro, eds., ​*Consciousness and Self-Regulation*​. Vol 3. New York: Plenum.

————————. 1986. What’s really going on in Searle’s “Chinese Room.” *Philosophical Studies*​ 50:169–85.

————————. 1992. Sensational sentences reversed. *Philosophical Studies*​ 68:289–319.

Reynolds, C. 1987. Flocks, herds, and schools: A distributed behavioral model. *Computer Graphics*​ 21:25–34.

​**[lk.401]**​ Robinson, H. 1976. The mind–body problem in contemporary philosophy. *Zygon*​ 11:346–60.

————————. 1982. *Matter and Sense*​. Cambridge: Cambridge University Press.

Robinson, W. S. 1988. *Brains and People: An Essay on Mentality and Its Causal Conditions*​. Philadelphia: Temple University Press.

Rosenberg, G. H. 1996. Consciousness and causation: Clues toward a double-aspect theory. Manuscript, Indiana University.

Rosenthal, D. M. 1996. A theory of consciousness. In N. Block, O. Flanagan, and G. Guzeldere, eds., ​*The Nature of Consciousness*​. Cambridge, Mass.: MIT Press.

Russell, B. 1927. *The Analysis of Matter*​. London: Kegan Paul.

Ryle, G. 1949. *The Concept of Mind*​. London: Hutchinson.

Savage, C. W. 1970. *The Measurement of Sensation*​. Berkeley: University of California Press.

Savitt, S. 1982. Searle’s demon and the brain simulator reply. *Behavioral and Brain Sciences*​ 5:342–43.

Sayre, K. M. 1976. *Cybernetics and the Philosophy of Mind*​. Atlantic Highlands, N.J.: Humanities Press.

Sayre-McCord, G. 1988. Introduction: The many moral realisms. In G. Sayre-McCord, ed., ​*Essays on Moral Realism*​. Ithaca, N.Y.: Cornell University Press.

Schacter, D. L. 1989. On the relation between memory and consciousness: Dissociable interactions and conscious experience. In H. Roediger and F. Craik, eds., ​*Varieties of Memory and Consciousness: Essays in Honor of Endel Tulving*​. Hillsdale, N.J.: Erlbaum.

Schlick, M. 1932. Positivism and Realism. *Erkenntnis*​ 3.

————————. 1938. Form and content: An introduction to philosophical thinking. In ​*Gesammelte Aufsätze*​ 1926–1936. Vienna: Gerold [Reprinted in H. L. Mulder and B. van de Velde-Schlick, eds., ​*Philosophical Papers*​. Dordrecht: Reidel, 1979]

Seager, W. E. 1988. Weak supervenience and materialism. *Philosophy and Phenomenological Research*​ 48:697–709.

————————. 1991. *Metaphysics of Consciousness*​. London: Routledge.

Searle, J. R. 1980. Minds, brains, and programs. *Behavioral and Brain Sciences*​ 3: 417–24.

————————. 1984. *Minds, Brains and Science*​. Cambridge, Mass.: Harvard University Press.

————————. 1990a. Consciousness, explanatory inversion and cognitive science. *Behavioral and Brain Sciences*​ 13:585–642.

————————. 1990b. Is the brain a digital computer? *Proceedings and Addresses of the American Philosophical Association*​ 64:21–37.

————————. 1992. *The Rediscovery of the Mind*​. Cambridge, Mass.: MIT Press.

Sellars, W. 1956. Empiricism and the philosophy of mind. In H. Feigl and M. Scriven, eds., ​*The Foundations of Science and the Concepts of Psychology and Psychoanalysis*​. Minnesota Studies in the Philosophy of Science, vol. 1. Minneapolis: University of Minnesota Press.

————————. 1965. The identity approach to the mind–body problem. *Review of Metaphysics*​ 18:430–51.

————————. 1981. Is consciousness physical? *Monist*​ 64:66–90.

Shallice, T. 1972. Dual functions of consciousness. *Psychological Review*​ 79:383–93.

————————. 1988a. Information-processing models of consciousness: Possibilities and​**[lk.402]**​ problems. In A. Marcel and E. Bisiach, eds., ​*Consciousness in Contemporary Science*​. Oxford: Oxford University Press.

————————. 1988b. *From Neuropsychology to Mental Structure*​. Cambridge: Cambridge University Press.

Shannon, C. E. 1948. A mathematical theory of communication. *Bell Systems Technical Journal*​ 27:379–423. [Reprinted in C. E. Shannon and W. Weaver, ​*The Mathematical Theory of Communication*​. Urbana: University of Illinois Press, 1949]

Shepard, R. N. 1993. On the physical basis, linguistic representation, and conscious experience of colors. In G. Harman, ed., ​*Conceptions of the Human Mind: Essays in Honor of George A. Miller*​. Hillsdale, N.J.: Erlbaum.

Shoemaker, S. 1975a. Functionalism and qualia. *Philosophical Studies*​ 27:291–315.

————————. 1975b. Phenomenal similarity. *Critica*​ 7:3–37.

————————. 1980. Causality and properties. In P. van Inwagen, ed., ​*Time and Cause*​. Dordecht: Reidel.

————————. 1981. Some varieties of functionalism. *Philosophical Topics*​ 12:93–119.

————————. 1982. The inverted spectrum. *Journal of Philosophy*​ 79:357–81.

Sidelle, A. 1989. *Necessity, essence, and individuation*​. Ithaca, N.Y.: Cornell University Press.

————————. 1992. Rigidity, ontology, and semantic structure. J​*ournal of Philosophy*​ 8:410–30.

Siewert, C. 1993. What Dennett can’t imagine and why. *Inquiry*​ 36:93–112.

————————. 1994. Understanding consciousness. Ph.D. diss. University of California, Berkeley. [Forthcoming as a book from Princeton University Press]

Skyrms, B. 1980. *Causal Necessity*​. New Haven: Yale University Press.

Smart, J. J. C. 1959. Sensations and brain processes. *Philosophical Review*​ 68:141–56.

Sperling, G. 1960. The information available in brief visual presentations. *Psychological Monographs*​ 74.

Sperry, R. W. 1969. A modified concept of consciousness. *Psychological Review*​ 76:532–36.

————————. 1992. Turnabout on consciousness: A mentalist view. *Journal of Mind and Behavior*​ 13:259–80.

Sprigge, T. L. S. 1971. Final causes. *Proceedings of the Aristotelian Society*​ 45 (suppl.):149–70.

————————. 1994. Consciousness. *Synthese*​ 98:73–93.

Squires, E. 1990. *Conscious Mind in the Physical World*​. Bristol: Hilger.

Stalnaker, R. 1976. Possible worlds. *Nous*​ 10:65–75.

————————. 1978. Assertion. In P. Cole, ed., ​*Syntax and Semantics: Pragmatics*​. Vol. 9. New York: Academic Press.

Stapp, H. P. 1993. *Mind, Matter, and Quantum Mechanics*​. Berlin: SpringerVerlag.

Stevens, S. S. 1975. *Psychophysics: Introduction to Its Perceptual, Neural, and Social Prospects*​. New York: Wiley.

Sutherland, N. S., ed. 1989. *The International Dictionary of Psychology*​. New York: Continuum.

Swoyer, C. 1982. The nature of natural laws. *Australasian Journal of Philosophy*​ 60:203–23.

Teller, D. Y. 1984. Linking propositions. *Vision Research*​ 24:1233–46.

​**[lk.403]**​ ————————. 1990. The domain of visual science. In L. Spillman, and J. S. Werner, eds., ​*Visual Perception: The Neurophysiological Foundations*​. New York: Academic Press.

Teller, P. 1984. A poor man’s guide to supervenience and determination. *Southern Journal of Philosophy*​, suppl., 22:137–62.

————————. 1992. A contemporary look at emergence. In A. Beckermann, H. Flohr, and J. Kim, eds., ​*Emergence or Reduction? Prospects for Nonreductive Physicalism*​. Berlin: De Gruyter.

Thagard, P. 1986. The emergence of meaning: An escape from Searle’s Chinese Room. *Behaviorism*​ 14:139–46.

Thompson, E. 1992. Novel colours. *Philosophical Studies*​ 68:321–49.

Tienson, J. L. 1987. Brains are not conscious. *Philosophical Papers*​ 16:187–93.

Tooley, M. 1977. The nature of laws. *Canadian Journal of Philosophy*​ 7:667–98.

————————. 1987. *Causation: A Realist Approach*​. Oxford: Oxford University Press.

Tye, M. 1986. The subjective qualities of experience. *Mind*​ 95:1–17.

————————. 1992. Visual qualia and visual content. In T. Crane, ed., ​*The Contents of Experience*​. Cambridge: Cambridge University Press.

————————. 1993. Blindsight, the absent qualia hypothesis, and the mystery of consciousness. In C. Hookway, ed., ​*Philosophy and the Cognitive Sciences*​. Cambridge: Cambridge University Press.

————————. 1995. *Ten Problems of Consciousness*​. Cambridge, Mass.: MIT Press.

van Cleve, J. 1990. Mind-dust or magic? Panpsychism versus emergence. *Philosophical Perspectives*​ 4:215–26.

van Gulick, R. 1988. A functionalist plea for self-consciousness. *Philosophical Review*​ 97:149–88.

————————. 1992. Nonreductive materialism and the nature of intertheoretical constraint. In ​*A. Beckermann, H. Flohr, and J. Kim, eds., Emergence or Reduction? Prospects for Nonreductive Physicalism*​. Berlin: De Gruyter.

————————. 1993. Understanding the phenomenal mind: Are we all just armadillos? In M. Davies and G. Humphreys, eds., ​*Consciousness: A Mind and Language Reader*​. Oxford: Blackwell.

Velmans, M. 1991. Is human information processing conscious? *Behavioral and Brain Sciences*​ 14:651–69.

Weinberg, S. 1992. *Dreams of a Final Theory*​. New York: Pantheon Books.

Weiskrantz, L. 1986. *Blindsight: A Case Study and Implications*​. Oxford: Oxford University Press.

————————. 1992. Introduction: Dissociated issues. In D. Milner and M. Rugg, eds., ​*The Neuropsychology of Consciousness*​. London: Academic Press.

Wheeler, J. A. 1990. Information, physics, quantum: The search for links. In W. H. Zurek, ed., ​*Complexity, Entropy, and the Physics of Information*​. Redwood City, Calif.: Addison–Wesley.

————————. 1994. It from bit. In ​*At Home in the Universe*​. Woodbury, N.Y.: American Institute of Physics Press.

Wheeler, J. A., and W. H. Zurek. 1983. *Quantum Theory and Measurement*​. Princeton: Princeton University Press.

White, S. L. 1986. Curse of the qualia. *Synthese*​ 68:333–68.

Wigner, E. P. 1961. Remarks on the mind–body question. In I. J. Good, ed., ​*The Scientist Speculates*​. New York: Basic Books.

​**[lk.404]**​ Wilkes, K. V. 1984. Is consciousness important? *British Journal for the Philosophy of Science*​ 35:223–43.

Wilson, M. 1982. Predicate meets property. *Philosophical Review*​ 91:549–89.

————————. 1985. What is this thing called “pain”? The philosophy of science behind the contemporary debate. *Pacific Philosophical Quarterly*​ 66:227–67.

Winograd, T. 1972. *Understanding Natural Language*​. New York: Academic Press.

Wittgenstein, L. 1953. *Philosophical Investigations*​. London: Macmillan.

————————. 1968. Notes for lectures on “private experience” and “sense data.” *Philosophical Review*​ 77.

Wright, R. 1988. *Three Scientists and Their Gods*​. New York: Times Books.

Yablo, S. 1993. Is conceivability a guide to possibility? *Philosophy and Phenomenological Research*​ 53:1–42.

Zuboff, A. 1994. What is a mind? In P. A. French, T. E. Uehling, and H. K. Wettstein, eds., ​*Philosophical Naturalism*​. Midwest Studies in Philosophy, vol. 19. Notre Dame, Ind.: University of Notre Dame Press.

Zurek, W. H. 1990. *Complexity, Entropy, and the Physics of Information*​. Redwood City, Calif.: Addison–Wesley.

1. “Nick’s English Hut” — Indiana ülikooli tudengite pubi Bloomingtonis. [↑](#footnote-ref-2)
2. Kuidas seda eesti keeles öelda? Jutt on [loengusarjast nimega “John Locke Lectures”](https://www.philosophy.ox.ac.uk/john-locke-lectures-test), mis on nimetatud Locke’i järgi, kuid loengute teemaks ei tarvitse olla John Locke. [↑](#footnote-ref-3)
3. “fortune cookie” — kas eesti keeles on talutavat vastet? [↑](#footnote-ref-4)
4. “My Zombie Valentine” — ebamugav tõlkida, eriti kui raamat ise pole kättesaadav; siinse variandi eeskujuks ja õigustuseks on venelase vaste “Моя подруга зомби” ja kaks sisukirjeldust (kaaneteksti?), mida võib leida Internetis paljudest kohtadest: 1) “Sixth-grader Joey Ocean is sure there is something very special about Xia Dedd, the new girl in school.” 2) “Something is very strange about the new girl in Joey Ocean’s class. Her name is Xia Dedd. She never speaks, and her hair blows in the breeze-even inside. There is one thing Xia can do: “mind talk” with Joey. He learns that Dr. Samedi is experimenting on humans and Xia is one of his unfortunate subjects. Together they must find a way to stop the evil doctor before it’s too late for Xia. And possibly, humankind!” [↑](#footnote-ref-5)
5. ​**AM 1.1**​ Vt. Nagel 1974. Selle väljendi esmakordset tarvitust filosoofilisetes kontekstides omistatakse tavaliselt Farrellile (1950). Vt. ka Sprigge 1971. [↑](#footnote-ref-6)
6. ​**AM 1.2**​ Terminit “kvaal” on mõistetud mitut moodi. Mina tarvitan seda, nagu mulle tundub, kanoonilisel viisil, mõistes kvaalide all vaimsete seisundite omadusi, mis iseloomustavad neid tunde järgi neis olla. Selle termini kasutamisega ei taha ma juba ette valida seisukohta allpool arutamisele tulevates probleemides: kas me võime kvaalide tundmises ka eksida, kas nad on intensionaalsed omadused, jne. Kvaal võib olla nii “seesmiste” vaimsete seisundite kui ka aistingute omadus. Sageli on mugav rääkida kvaalidest nii, nagu nad oleksid omadused, mida realiseerib otseselt subjekt, aga mitte nii, nagu neid realiseeriksid tema vaimuseisundid; säärasest praktikast ei sünni mingit kahju, ja teda õigustab asjaolu, et kvaalid vastavadki oma olemuselt vaimuseisundite tüüpidele. [↑](#footnote-ref-7)
7. “spatiotemporal manifold” — miks tarvitab autor siin nii keerulist väljendit; kas see tähendab siin midagi spetsiifilist? (Nagu näitab [Google’i otsing​](https://www.google.com/search?q=spatiotemporal+manifold), paistab väljendil “spatiotemporal manifold” olevat mingites kontekstides tõepoolest mingi omaette tähendus, aga minu haridusest ei piisa, et aru saada, missugune nimelt.)  — (Nimi)sõna “manifold” kordub allpool veel, üle 10 korra.  — Definitsioone Internetist: [​matemaatika](https://mathworld.wolfram.com/Manifold.html): “A manifold is a topological space that is locally Euclidean (i.e., around every point, there is a neighborhood that is topologically the same as the open unit ball in Rⁿ).”; [filosoofia](https://www.dictionary.com/browse/manifold): “(in Kantian epistemology) the totality of discrete items of experience as presented to the mind; the constituents of a sensory experience.”  — See filosoofiline definitsioon võiks ju siia ka sobida, aga tundub, et kui otsustada, et autor on silmas pidanud seda, siis “muutkond” küll hea vaste ei ole. [↑](#footnote-ref-8)
8. ​**AM 1.3**​ Ma tarvitan siin raamatus väljendeid nagu “punane aisting”, “roheline elamus” jms. Loomulikult ei taha ma öelda, nagu realiseeruksid teadvuselamustes endis samasugused värvused nagu välismaailma objektidel (õuntel, puudel). Seesuguseid ütlusi saab alati ümber sõnastada, öeldes näiteks “seesugune elamus, mis tekib mul tavaliselt (aktuaalses maailmas) punast objekti vaadates” jne., kuid lühem väljendusviis mõjub loomulikumalt. [↑](#footnote-ref-9)
9. “middle C”; “neljanda oktavi” sain, võrreldes andmeid oktavite nimetuste kohta [siin (eesti)](https://et.wikipedia.org/wiki/Oktav_%28helik%C3%B5rguspiirkond%29) ja [siin (inglise)](https://en.wikipedia.org/wiki/C_%28musical_note%29%22%20%5Cl%20%22Middle_C). [↑](#footnote-ref-10)
10. Mis kääne? “Korteksi​**le**​” või “korteksi​**sse**​”? [↑](#footnote-ref-11)
11. Kas ja kuidas seda tõlkida? (Sääraseid märkeid tuleb allpool kaks tükki veel.) [↑](#footnote-ref-12)
12. ​**AM 1.4**​ Vala 12 tassitäit kuivatatud ​aedube​ vette, millele on lisatud 4 supilusikatäit soola. Keeda oad pehmeks ja jahuta külmas vees. Lisa 2 tükeldatud punast piprakauna, 5 tükeldatud rohelist piprakauna, 2 tükeldatud suurt sibulat, 3 tassitäit rosinaid ja näputäis hakitud koriandrilehti. Valmista 1.5 tassitäiest maisiõlist, 0.75 tassitäiest veiniäädikast, 4 supilusikatäiest suhkrust, 1 supilusikatäiest soolast, 4 supilusikatäiest mustast piprast, 5 supilusikatäiest karripulbrist ja poolest supilusikatäiest jahvatatud nelgist kaste ja vala salatile. Serveeri külmalt. (Tänan retsepti eest Lisa Thomast ja kohvikut “Encore”.) [↑](#footnote-ref-13)
13. ​**AM 1.5**​ Lugeja, kes soovib oma teadvuselamustesse põhjalikumalt süveneda, leiab mitmeks päevaks väga mitmesuguste spetsiifiliste elamuste üle rikkalikku mõtisklusainet Ackermani raamatust “Meelte looduslugu” [​*A Natural History of the Senses*​] (1990). [↑](#footnote-ref-14)
14. ​**AM 1.6**​ Huvitaval kombel jättis Descartes vaimsete nähtuste kategooriast sageli välja aistingud, ühendades need hoopis kehaga, mistõttu tema järgi tuleb välja, et mitte iga fenomenaalset seisundit (vähemalt nii, nagu mina sellest mõistest aru saan) ei saa pidada ühtlasi ka vaimu seisundiks. [↑](#footnote-ref-15)
15. ​**AM 1.7**​ See Ryle’i vaadete tavapärane tõlgendus ei võta täielikult arvesse kõiki peensusi, kuid on vähemasti kasulik fiktsioon. [↑](#footnote-ref-16)
16. ​**AM 1.8**​ Funktsionalismil on ka teisi vorme, näiteks see, mille töötas välja Putnam (1960). Ma ei hakka neid siin vaatlema, kuna neid on pakutud välja empiiriliste hüpoteesidena, mitte vahenditena vaimsete mõistete analüüsimiseks. [↑](#footnote-ref-17)
17. ​**AM 1.9**​ ​**[lk.360]**​  Nagel (1970) esitab Armstrongile sarnase vastuväite [​seoses​](#fn-23reference) teiste vaimude probleemiga. [↑](#footnote-ref-18)
18. ​**AM 1.10**​ Searle’i mõttekäik toetub eeldusele, et ilma teadvuseta ei oleks võimalik seletada intentsionaalsuse *aspektuaalset kuju* — näiteks juhul, kui keegi usub midagi Veenuse kohta tema Koidutähe-aspektis, kuid mitte tema Ehatähe-aspektis. Minu arvates pole ilmne, et aspektuaalset kuju pole võimalik seletada ühelgi muul viisil. Võiks koguni öelda, et midagi selletaolist võib täheldada tänapäeva arvutites: neil võib olla informatsiooni minu kohta minu mingis ühes aspektis, kuid mitte teises. (Näiteks võib arvuti teada minu nime, aga mitte minu isikukoodi). Sellele võidakse vastu väita, et see on vaid “justkui”-aspektuaalne kuju, sest tõeline aspektuaalne kuju saab olla üksnes fenomenaalne; kuid tundub, et sellega kaotaks argument kogu huvitavuse. [↑](#footnote-ref-19)
19. Tõlkisin “social security number” “isikukoodiks”. [↑](#footnote-ref-20)
20. ​**AM 1.11**​ Säärasele seisukohale viitavad mõningad Lockwoodi (1989) ja Nageli (1986) märkused, kuigi ma pole kindel, et kumbki neist sellist otsustust heaks kiidaks. [↑](#footnote-ref-21)
21. “… that the manifest phenomena in the vicinity that need explaining” — ma ei saa aru, mida peab siin tähendama see “vicinity”. Ja “manifest” tundub siin ülearu olevat, sest kas pole “phenomenon” juba definitsiooni järgi “manifest”? [↑](#footnote-ref-22)
22. ​**AM 1.12**​ Mõni osutaks siin võib-olla argumentidele, mida esitas Kripke (1982) näitamaks, et uskumuse sisu ei tulene psühholoogilistest ja fenomenaalstest omadustest. Nende argumentide üle võib vaielda, kuid igal juhul väärib tähelepanu, et järeldused, milleni nad välja viivad, ei seisne selles, et see sisu on vaimu element, mida ei saa enam millelegi edasi taandada, vaid selles, et sisu on ise määratlematu. Kogu asi on põhimõtteliselt selles, et põhitekstis esitatud kaalutluste põhjal on meil täielik alus uskuda, et vaimul ei ole enam mingit kolmandat, sõltumatult varieeritavat aspekti, ja see tähendab, et küsimustele, millele ei ole võimalik vastata esimest kaht liiki elementide kaudu, ei olegi võimalik vastata. [↑](#footnote-ref-23)
23. [Filosoofilised uurimused](https://www.ester.ee/record%3Db2023219%2Aest), lk. 246. (Jätsin ära (ühekordsed) jutumärgid kahe sõna ümbert eestikeelses tõlkes, mis Chalmersi tsiteeringus puuduvad.) [↑](#footnote-ref-24)
24. ​**AM 1.13**​ Sel spetsiifiliste fenomenaalsete mõistete “temaatiliselt neutraalsel” analüüsil ei puudu sarnasus analüüsidega, mida on esitanud Place (1956) ja Smart (1959). Oranži värvuse elamus​ on väga jämedalt öeldes sedaliiki elamus, mida tavaliselt kutsub esile apelsini nägemine​. Place: “Öeldes järelkujutise kohta, et see on roheline …, ütleme, et meil on umbes samasugune elamus nagu see, mida meis normaalselt põhjustab roheliste valguslaigu vaatamine ja mida me oleme õppinud säärasena kirjeldama” (lk. 49); Smart: “Kui keegi ütleb: “Ma näen kollakasoranži järelkujutust”, siis ütleb ta midagi niisugust: “Minuga on toimumas midagi sarnast sellega, mis toimub, kui mul on silmad lahti, ma olen ärkvel, toas on valge ja minu ees on apelsin” ”. (lk. 150). Ent kuna elamuse mõiste on jäänud analüüsimata, siis saa ainuüksi selle analüüsi varal ilma pikemata samastada fenomenaalseid ja füüsilisi seisundeid nii, nagu seda tahavad teha Place ja Smart. Smart jätab sellest raskusest vabanemiseks analüüsist välja “elamuse”, asendades selle mitmemõttelise fraasiga “on toimumas”. Kui võtta seda väljendit küllalt avaralt, nii et ta kataks igatliiki seisundeid, siis on Smarti analüüs puudulik; paremini tabaksime märki, võttes neid kitsalt kui mingi läbielamise kirjeldust, kuid sellisel juhul on analüüs järelduse põhjendamiseks ebapiisav. [↑](#footnote-ref-25)
25. “To be an orange experience … is to be the kind of experience that is generally caused by oranges” — esimene “orange” on siin kahemõtteline, võib tähendada nii oranži [värvust] kui ka apelsini. [↑](#footnote-ref-26)
26. “not so much as splitting the mind-body problem as factoring it into two separate parts” — ei saa aru: mis seal vahet on?! [↑](#footnote-ref-27)
27. ​**AM 1.14**​ Jackendoff teeb vahet “fenomenaalse vaimu” ja “informaatilise vaimu” vahel. See eristus ühtib enam-vähem minu vahetegemisega fenomenaalsete ja psühholoogiliste nähtuste vahel, kuigi ma ei tahaks hakata siin arutlema küsimuse üle, kas psühholoogilised protsessid on informaatilised. [↑](#footnote-ref-28)
28. Mul oli siin kiusatus pealkirja (originaalis “Varieties of psychological consciousness”) kallal omavolitseda, sest ta on eksitav: nagu ütleb juba jao esimene lausegi, ei tule juttu mitte ühe asja eri kujudest, vaid ühe väljendi eri tähendustest. [↑](#footnote-ref-29)
29. ​**AM 1.15**​ Nelkini (1989) järgi on olemas vähemalt​ kolme liiki teadvust: CN (teadvus Nageli mõttes), C1 (esimest järku infotöötlusseisund) ja C2 (teist järku otsene loogikateheteta ligipääs teistele teadvusseisunditele). Teises artiklis (Nelkin 1993) eristab ta analoogiliselt kolme sarnast mõistet: fenomenaalsust, intentsionaalsust ja introspekteeritavust. Bisiach (1988) kõneleb teadvusest C1 (fenomenaalne elamus) ja teadvusest C2 (süsteemi osade või protsesside suhtlus süsteemi teiste osade või protsessidega). Natsoulas (1978) ​**[lk.361]**​ eristab sõnal “teadvus” tervet hulka tähendusi. Dennett (1969) teeb vahet kaht liiki “teadlikkusel”: esimene seostub sõnaliste teavitustega, teine üldisemalt käitumise juhtimisega, kuigi kumbki neist pole päris selgelt fenomenaalne mõiste. [↑](#footnote-ref-30)
30. Lisasin selle “vähemalt” omalt poolt, vt. [originaalartiklit](https://philpapers.org/rec/NELUS). [↑](#footnote-ref-31)
31. Puhas oletus, mul õnnestus Nelkini artiklist kätte saada ainult resümee. [↑](#footnote-ref-32)
32. ​**AM 1.16**​ Rosenthal teeb selget vahet teadvuse ja “sensoorse kvaliteedi” vahel; ta ütleb, et tema teooria käsitleb ainult esimest, ja sellest võiks järeldada, et ta fenomenaalsete aspektidega ei tegele. Kuid ta ütleb ka, et mingi seisund on teadvuslik, kui on olemas tunne olla selles seisundis, ja selle põhjal võiks arvata, et tema jutt käib ikkagi fenomenaalsest teadvusest. Ent tema kontseptsioonis ei ole kuigi palju sellist, mis aitaks kaasa fenomenaalse teadvuse seletamisele. Miks peaks kõrgemat järku mõtetest mingi seisundi kohta järelduma, et on olemas tunne olla selles seisundis? Kuigi Rosenthal möönab, et praktikas esinevad need kaks fenomeni nähtavasti koos, ei anna ta sellele küsimusele vastust. [↑](#footnote-ref-33)
33. ​**AM 1.17**​ Üheks erandiks on psühhofüüsika, mis võib nähtavasti heita valgust teadvuselamuste mitmesugustele iseärasustele, kuigi ta täielikku seletust ei anna. Sellest tuleb pikemalt juttu [6. peatükis](#id_128psychophysics). [↑](#footnote-ref-34)
34. ​**AM 2.1**​ Superventsuse ideega tuli esmakordselt välja Moore (1922). Sõna “superventsus” [​*supervenience*​] tarvitas trükisõnas esmakordselt Hare (1952). Davidson (1970) rakendas seda mõistet esmakordselt keha–vaimu probleemile. Hiljem töötasid superventsuse üksikasjaliku teooria välja Kim (1978, 1984, 1993), Morgan (1982, 1984c, 1993), Hellman ja Thompson (1975) jt. [↑](#footnote-ref-35)
35. ​**AM 2.2**​ “A-fakt” tähendab lühidalt mingi A-omaduse realiseeringut. Omaduste asemel faktidest kõnelemine ei muuda mõttekäikude sisu, kuid teeb nad ladusamaks, sest kõike, mida me ütleme faktide kohta, saab ümber sõnastada omaduste ja nende üheaegse realiseerumise mustrite kaudu; vajadusel annan ma märkustes lisaseletusi. Tuleb kohe tähendada, et minu tõlgenduses ei sõltu A-fakt sellest, missugune indiviid teda realiseerib; oluline on vaid vastava omaduse realiseerumine. \_Kui mingi A-fakti koosseisu kuuluks ühe või teise indiviidi identsus, siis tuleneksid kõigist A-faktidest faktid selle indiviidi olemuslike omaduste kohta, ja superventsuse definitsioon viiks intuitsioonivastaste järeldusteni.\_ [↑](#footnote-ref-36)
36. “If the identity of an individual were partly constitutive of an A-fact, then any A-fact would entail facts about that individual’s essential properties, in which case the definition of supervenience would lead to counterintuitive consequences.” — ei saa aru, missugusteni! (Kui autor ei pea just silmas tulemusi, millest on juttu artiklis [Ingvar Johansson: Identity Puzzles and Supervenient Identities](https://citeseerx.ist.psu.edu/document?doi=da49f64ffd1ff360216644855029a2981545d3a8) (lk. 8, 22, 25–6), aga niipalju kui mina aru saan, saab sealt järeldada ainult niipalju, et indiviidi mõiste on liiga laialivalguv ja kahtlane, et teda superventsuse üle arutledes üldse kasutada saaks.) [↑](#footnote-ref-37)
37. “numerical identity”: [definitsioon](https://pressbooks.online.ucf.edu/introductiontophilosophy/chapter/personal-identity/): “Numerical identity is a term that philosophers use to describe an object being the very same object. It is contrasted with qualitative identity which simply means that an object has all the same properties or qualities. An object is only numerically identical to itself but can be qualitatively identical to other objects.”  — Oletan, et sõnal “numerical” võib olla ka käibiv eestikeelne vaste, aga Internetist ma seda ei leidnud. [↑](#footnote-ref-38)
38. ​**AM 2.3**​ Ma eeldan siin võib-olla pisut meelevaldselt, et indiviididel on ajas ja ruumis täpsed piirid, nii et nende füüsilised omadused realiseeruvad mingis kindlas aja- ja ruumipiirkonnas. Kui me peaksime lokaalsest superventsusest kõneldes lugema füüsiliselt identseteks objekte, mille ruumiline asukoht on erinev, siis tuleks arvata superventsusbaasist välja kõik omadused, mis puudutavad objekti absoluutset asukohta ajas ja ruumis (kuigi vajadust vaadelda ruumi eri kohtades paiknevaid objekte saaks vältida, vaadeldes ainult võimalikke objekte ühes ja samas asukohas). Peale selle kõnelen ma alati nii, nagu realiseeruksid nii madaltaseme kui ka kõrgtaseme omadused sama tüüpi indiviidides: näiteks lauas realiseeruvad mikrofüüsikalised omadused, kuna teda määratleb nende omaduste jaotus​. Võib-olla oleks õigem ja täpsem kõnelda mikrofüüsikaliste omaduste realiseerumisest ainult mikrofüüsiliste objektide puhul, kuid minu sõnastusviis teeb pildi lihtsamaks. Ja nagunii ei ole tõeliselt kesksete küsimuste puhul oluline mitte lokaalne, vaid globaalne superventsus. [↑](#footnote-ref-39)
39. Ei saanud aru! [↑](#footnote-ref-40)
40. ​**AM 2.4**​ On võimalik mitut moodi täpselt defineerida, mida tähendab, et kaks maailma on mingilt omaduste kogumilt identsed; meie arutluskäik sellest oluliselt ei muutu. ​**[lk.362]**​ Võib-olla on kõige parem öelda, et kaks maailma on oma A-omadustelt identsed, kui on võimalik korraldada üksühene vastavus mõlemas maailmas A-omadusi realiseerivate indiviidiklasside vahel, nii et kahes teineteisele vastavas indiviidis realiseeruvad alati samad A-omadused. Säärasel juhul on globaalseks superventsuseks tarvilik, et kujutused, mille raames kaht maailma võib pidada nii A- kui ka B-omadustelt eristamatuteks, oleksid omavahel ühilduvad, või teisiti öeldes, et A-kujutus annaks igale indiviidile paarikuks sama indiviidi mis B-kujutus. Globaalse superventsuse definitsioon kõlaks siis järgmiselt: Kaks maailma, mis on (mingi kujutuse järgi) A-identsed, on (sama kujutuse ​mingi ülemkujutuse järgi​) alati ka B-identsed.

Tavalisem viis on postuleerida, et A-identsed maailmad peavad sisaldama täpselt samu ja täpselt samade omadustega indiviide, kuid nagu tähendab McLaughlin (1995), on see ebamõistlikult tugev eeldus, kuna ta tagab, et sellised asjad nagu maailma kardinaalsus​ ja indiviidide olemuslikud omadused on superventsed üldse kõikvõimalike omaduste suhtes. Minu pakutud definitsioon laveerib sellest takistusest mööda, hoolitsedes selle eest, et determineerivas suhtes ei osale miski peale A-omaduste mustrite​. [↑](#footnote-ref-41)
41. “under an extension of that mapping” — Siin oleks mul kange tahtmine omavolitseda: mulle jääb arusaamatuks, miks autoril siin seda “extensionit” vaja oli, ja see tundub isegi lausa eksitav (just sellepärast, et paneb lugejat ilmaaegu mõtlema, miks seda “extensionit” vaja on): on ju indiviidide hulk nii A- kui ka B-kujutuse korral täpselt sama. Iseasi, kui B-kujutuses figureeriks indiviide, millel pole A-kujutuses kohta. Seega oleks selgem ja sirgjoonelisem, kui autor oleks öelnud “under that mapping”, s.t. “sama kujutuse järgi”. [↑](#footnote-ref-42)
42. “cardinality” — mis see on?? [↑](#footnote-ref-43)
43. ​**AM 2.5**​ Ühe erandiga: Jumal poleks saanud luua maailma, mis poleks Jumala loodud, kuigi maailm, mis ei oleks Jumala loodud, on vististi loogiliselt võimalik! Selletaolisi komplikatsioone ma edaspidi ei vaatle. [↑](#footnote-ref-44)
44. ​**AM 2.6**​ Sedaliiki võimalikkuse seos tuletatavusega formaalsetes süsteemides on üpris raskesti tabatav. Leidub argumente väite kasuks, et iga konkreetse formaalse süsteemi aksioomide ja tuletusreeglite aluseks on, vastupidi, juba olemas olev arusaam loogilisest võimalikkusest ja paratamatusest. [↑](#footnote-ref-45)
45. ​**AM 2.7**​ Intuitiivne arusaam looduslikust võimalikkusest eelneb kontseptuaalselt selle võimalikkuse definitsioonile loodusseaduste kaudu: seaduspärasust võib pidada seaduseks just siis, kui ta peab paika kõigis olukordades, mis võiksid looduses kujuneda — see tähendab, kõigis olukordades, mis on looduslikult võimalikud nende sõnade intuitiivses tähenduses. Mõnikord formuleeritakse seda nii: selleks, et mingit seaduspärasust saaks pidada seaduseks, peab ta lisaks aktuaalsetele olukordadele kehtima ka kontrafaktilistes olukordades, ning otsustamaks, missugused kontrafaktilised olukorrad on relevantsed, on tarvis fundamentaalsemat arusaama looduslikust võimalikkusest. [↑](#footnote-ref-46)
46. ​**AM 2.8**​ Samaliiki paratamatuse tähistamiseks tarvitatakse sageli ka väljendeid “füüsiline paratamatus” ja “põhjuslik paratamatus”, kuid ma ei taha midagi ette otsustada küsimuses, kas kõik loodusseadused on füüsilised või põhjuslikud. [↑](#footnote-ref-47)
47. ​**AM 2.9**​ Seda olulist erinevust loogilise ja loodusliku superventsuse vahel vaikitakse kirjanduses sageli maha või ignoreeritakse, jättes superventsussuhete modaalsuse sageli täpsustamata. Loodusliku (või nomoloogilise) superventsuse üle ilma loogilise superventsuseta arutleb van Cleve (1990), kes püüab selle varal selgitada üht emergentsuse erikuju. Seager (1991) eristab sarnaselt nn. ​*konstitutiivset superventsust*​ ja *korrelatiivset superventsust*​. Need vastavad otseselt loogilisele ja looduslikule superventsusele, kuigi Seager analüüsib neid mõisteid mõnevõrra teisiti. [↑](#footnote-ref-48)
48. ​**AM 2.10**​ *Nõrk superventsus* nõuab ainult seda, et olukord, kus poleks “ühtki B-erinevust ilma A-erinevuseta”, kehtiks ainult ühe maailma piires, mitte kõigis maailmades. (Täpsemalt vt. Kim 1984.) Oma modaalse nõrkuse tõttu on see suhe enamikuks otstarveteks liiga nõrk. Parimal juhul võib ta mängida mingit rolli kontseptuaalsetes kitsendustes mittefaktilistele arutlustele (nagu artiklis Hare 1984), ehkki, nagu osutab Horgan (1993), kaasnevad nendegi kitsendustega sõltuvused, milles figureerib mitu maailma. Seager (1988) kasutab nõrga superventsuse mõistet, väljendamaks mingit süstemaatilist, kuid mitte rangelt paratamatut korrelatsiooni ühe maailma piires. Kuid looduslik superventsus teeniks seda eesmärki palju paremini. [↑](#footnote-ref-49)
49. ​**AM 2.11**​ ​**[lk.363]**​ Globaalne loogiline superventsus ilma lokaalsete regulaarsusteta on loogiliselt koherentne mõiste, kui lähtuda seaduste mitte-hume’likust kontseptsioonist, ent kui toetuda Hume’i kontseptsioonile (mille aluseks on reeglipärad), siis ta seda võib-olla ei ole. Ent isegi mitte-hume’likku seisukohta tunnistades on raske mõista, missuguseid andmeid võiks pidada tõendiks niisuguste suhete kohta. [↑](#footnote-ref-50)
50. Hume’i arvates on loodusseadused lihtsalt mingid kõikjal ja alati korduvad mustrid ja ei midagi muud; vastaspoole arvates on loodusseaduste taga (või kohal, või all) mingi kehtestav jõud või võim. [↑](#footnote-ref-51)
51. Mis mõttes, mille definitsiooni? Tundub, et autor ei pea silmas inglitega maailma või bioloogiliste omaduste definitsiooni; üle jääb superventsus, aga sellisel juhul miks ütleb ta “​**meie** definitsiooni”? [Superventsuse mõistet selgitades](#ORG.2.1) nimetab ta superventsust lihtsalt filosoofiliseks mõisteks ja loetleb selle tarvituselevõtjaid ja väljatöötajaid ega ütle kusagil, et tal oleks selle kohta mingi oma definitsioon. Superventsuse definitsiooni täpsustab ta alles käesolevas ajajaos (ülejärgmises lõigus), aga seda teebki ta selleks, et siin lõigus mainitud murest lahti saada. [↑](#footnote-ref-52)
52. ​**AM 2.12**​ Horgan (1982), Jackson (1994), and Lewis (1983b) arutlevad sarnase raskuse üle materialismi defineerimisel​. [↑](#footnote-ref-53)
53. “denning materialism” — ei oska arvata, mida see võiks tähendada, ja ei leidnud ka Internetist midagi; panin esialgu “defineerimisel” venelase eeskujul, kellel on “в контексте определения материализма”.  — Veel üks võimalus, mis pähe tuleb, on, et “den” võiks olla lühend sõnast “denounce” (“hukka mõistma, taunima”), aga mulle ei tundu, et autori mõte selles oleks. [↑](#footnote-ref-54)
54. ​**AM 2.13**​ Seda uut definitsiooni saab täpsemini formuleerida [märkuse 2.4](#fn-2-4) vaimus. Olgu $B\left(W\right)$ B-omadustega indiviidide klass maailmas $W$. Me võime öelda, et maailm $W’$ on maailma $W$ B-​*ülemmaailm*​, kui leidub niisugune ühene vastavus $f:B\left(w\right)\rightarrow B\left(W’\right)$ (s.t. üksühene vastavus hulga $B\left(W\right)$ ja hulga $B\left(W’\right)$ mingi osahulga vahel), et iga indiviidi $a\in B\left(W\right)$ puhul realiseeruvad indiviidis $f\left(a\right)$ kõik omadused, mis realiseeruvad indiviidis $a$. Siis on B-omadused maailmas $W$ A-omaduste suhtes loogiliselt superventsed, kui iga maailmast $W$ A-eristamatu maailm on selle maailma B-ülemmaailm, kusjuures B-vastavused peavad olema jällegi A-vastavuste [​ülemhulgad​](#fn-34extension).

Mõistmaks, miks säärane kitsendus tarvilik on, kujutlegem, et meie maailmas on loenduv-lõpmatu arv pühholoogiliselt identseid vaime, milledest üks on realiseerunud ektoplasmas ja ülejäänud füüsiliselt. Intuitsioon ütleb, et selles maailmas ei ole psühholoogilised omadused füüsikaliste omaduste suhtes superventsed, kuid iga sellest maailmast füüsiliselt eristamatu maailm on psühholoogiliselt selle maailma ülemmaailm. Kuigi me ootaksime, et ektoplasmata maailm annaks tunnistust superventsuse vastu, on selle maailma psühholoogia ja meie maailma psühholoogia vahel üksühene vastavus. Raskus on selles, et see vastavus ei laiene füüsilisele maailmale, kuna ühe füüsilise entiteedi vasteks on ektoplasmaatiline entiteet; seepärast ongi kirjeldatud lisakitsendus vajalik. [↑](#footnote-ref-55)
55. Kus? Kas selleks erandiks on teadvuselamus? [↑](#footnote-ref-56)
56. “(in a certain sense)” — mis mõttes “mingis mõttes”? Meie maailmas on ta ju siis järglasteta igas mõttes. [↑](#footnote-ref-57)
57. “that contains​ ​*W* as a proper part”.  — Omavolitsesin: jätsin “proper” tõlkimata (kaks korda selles põhiteksti lõigus ja kaks korda [märkuses 2.14](#fn-2-14); vastasel korral oleks “osahulga” asemel tulnud panna “pärisosahulk”); see “proper” ei ole küll vale, kuid on täiesti ülearu ja teeb mõtet ainult segasemaks. (On ju ilmne, et kui miski kehtib maailmas ​*W*​, siis ta sellessamas maailmas ​*W*​ kehtib.) [↑](#footnote-ref-58)
58. ​**AM 2.14**​ Selles definitsioonis võib maailmade vahelist sisaldumissuhet võtta kui primitiivmõistet. Lewis (1983a) ja Jackson (1993) on tähendanud, et on viljatu sääraseid mõisteid lõpmatuseni aina edasi analüüsida. Algmõistetest me nagunii ei pääse, ja sisaldumissuhe on vististi nii selge, kui üks suhe üldse olla saab. Võimalik oleks ka kõnelda maailmade asemel, mis sisaldavad maailma​ ​*W* oma osahulgana, maailmadest, mis sisaldavad oma osahulgana maailma *W* kvalitatiivset duplikaati; see töötab sama hästi. [↑](#footnote-ref-59)
59. ​**AM 2.15**​ Paneme tähele, et selle definitsiooni kohaselt leidub positiivseid fakte, mis pole positiivsete omaduste realiseeringud. Mõtleme näiteks, missugune võiks olla omaduse “olla järglasteta või känguru” realiseering. Võib-olla tuleks positiivseid fakte defineerida rangemini kui positiivsete omaduste realiseeringuid. Ent niipalju kui ma otsustada oskan, ei too ka nõrgem definitsioon kaasa mingeid soovimatuid tagajärgi. [↑](#footnote-ref-60)
60. Omavolitsesin, sõnastasin selgemini; originaalitruum oleks midagi umbes niisugust: “iga loogiliselt võimaliku maailma ​*W* puhul, mis on meie maailmast A-eristamatu, kehtivad kõik B-faktid, mis kehtivad meie maailmas, ka maailmas ​*W*​.” [↑](#footnote-ref-61)
61. ​**AM 2.16**​ Küllap on soovitav, et omaduste loogiline superventsus meie maailmas tuleneks mingist seadusest. Kui maailma ajalugu oleks (näiteks mõnede teistsuguste juhuslike fluktuatsioonide tõttu) kujunenud teisiti, nii et meie maailmas lei­duks mittefüüsilisi elusaid ingleid, loodusseadused aga kehtiksid endiselt, siis oleks asjaolu, et bioloogilised omadused on füüsikaliste omaduste suhtes loogiliselt superventsed, vaid ajalooline juhus. Sisukama ja huvitavama metafüüsilise teesi saame, asendades loogilise superventsuse definitsioonides viited meie maailmale ja aktuaalsetele indiviididele viidetega looduslikult võimalikele maailmadele ja indiviididele. Sellisel juhul langeksid säärased stsenaariumid mängust välja. Lisaboonusena saaksime võimaluse teha kindlaks, kas omadused, mida ei realiseeri ükski indiviid — näiteks omadus olla miilikõrgune pilvelõhkuja —, on loogiliselt superventsed või mitte. Senise definitsiooni järgi on kõik säärased omadused superventsed, kuid see on sisutühi superventsus.

Nii jõuame (tavalisi kitsendusi kehtima jättes) järgmise definitsioonini: B-omadused on A-omaduste suhtes loogiliselt superventsed parajasti siis, kui iga loogiliselt võimalik olukord on iga temast A-eristamatu looduslikult võimaliku olukorra [B-ülemolukord](#fn-35-definition1). ​**[lk.364]**​ Või lühemalt: iga looduslikult võimaliku olukorra B-faktid tulenevad tema A-faktidest. See muudatus ei mõjuta kuigivõrd põhiteksti arutluskäiku, seepärast [jätsin](#fn-34-omit) ta sealt lihtsuse mõttes välja. Kuid kogu põhitekstis toodud arutlust on kerge sellele rangemale definitsioonile vastavaks ümber sõnastada, kaotamata midagi neis kohtades esitatud väidete veenvusest: tarvitseb lihtsalt asendada kogu tekstis “meie maailm” “kõigi looduslikult võimalike maailmadega”.

Tulemus meenutab “tugeva” lokaalse superventsuse tavapärast definitsiooni (Kim 1984), kus figureerib kaks modaalset operaatorit. Selle definitsiooni järgi on B-omadused A-omaduste suhtes superventsed, kui iga B-omaduse​ ​*F* ja iga selle omadusega indiviidi *x* puhul leidub paratamatult niisugune A-omadus ​*G*​, et indiviidil *x* on omadus ​*G*​, ja et kui ükskõik millisel indiviidil *y* on omadus ​*G*​, siis on tal paratamatult ka omadus ​*F*​. (Indiviidi​ ​*G*​ A-omadustena võib vajadusel vaadelda lihtsamate A-omaduste komplekse.) Raskus inglimaailmaga näitab, et äsjases definitsioonis tuleb esimest modaalset operaatorit alati mõista loodusliku paratamatuse mõttes, ka siis, kui teine väljendab loogilist paratamatust. Globaalse superventsuse tavapärase definitsiooniga (mille järgi iga A-eristamatu maailm on ka B-eristamatu) on lugu keerulisem — seda tuleb täpsustada ülalöeldud viisil. Vajadusel saaks eelmistega paralleelsel viisil defineerida ka “metafüüsilise” superventsuse. Loodusliku superventsuse juhul inglitega raskusi ei tekigi, sest pole põhjust uskuda, nagu oleks ektoplasma looduslikult võimalik; seega võib loodusliku superventsuse puhul jääda rahule sirgjoonelise definitsiooniga. [↑](#footnote-ref-62)
62. ​**AM 2.17**​ ​Võib-olla tuleks loogilise superventsuse definitsioonile anda tugevam kuju​(\*)​; sellisel juhul on materialism tõene, kui kõik kõigi looduslikult võimalike maailmade positiivsed faktid tulenevad nende maailmade füüsikalistest faktidest. Vaatleme näiteks ektoplasmata maailma, kus mittefüüsiline ektoplasma on siiski looduslikult võimalik: ta oleks võinud tekkida evolutsiooni käigus, kui mõned juhuslikud fluktuatsioonid oleksid olnud teistsugused. Võib vist öelda, et niisuguses maailmas oleks materialism väär, või et ta oleks tõene ainult mingis nõrgas mõttes.

(\*) [​*T.R.:*​] Ma ei saa aru, mida autor siin mõtleb. Originaalis on tegelikult “Arguably, we should use **the** stronger definition of logical supervenience …” — määrava artikliga, otsekui oleks jutt mingist tugevamast definitsioonist, mida lugeja juba teab. Kuid senist teksti uurides lugeja (vähemalt kui lugeja olen mina) mingit teist ja tugevamat definitsiooni ei leia. [↑](#footnote-ref-63)
63. Omavolitsesin, lihtsustasin sõnatust; originaalitruum oleks “… kui iga loogiliselt võimaliku ja meie maailmast eristamatu maailma *W* puhul kehtivad kõik positiivsed faktid, mis kehtivad meie maailmas, ka maailmas ​*W*​”. [↑](#footnote-ref-64)
64. ​**AM 2.18**​ Selleks, et see definitsioon oleks eelmisega samaväärne, peab kehtima tõepärane printsiip, mis ütleb, et kui maailm​ ​*B* on maailma *A* pärisülemmaailm, siis kehtib temas vähemalt üks fakt, mis ei kehti maailmas ​*A*​, või teisiti öeldes, on olemas fakte, mis kehtivad maailmas *B* ja kõigis temast mahukamates maailmades, kuid ei kehti maailmas​ ​*A*​. [↑](#footnote-ref-65)
65. ​**AM 2.19**​ See definitsioon on ka enam-vähem võrdväärne Horgani (1982) ja Lewise (1983b) definitsioonidega, kuid nendest erinevalt ei toetu ta mõnevõrra ähmasele “võõra [​*alien*​] omaduse” mõistele, mis on võetud tarvitusele selleks, et välistada relevantsete võimalike maailmade hulgast ektoplasmamaailmad. [↑](#footnote-ref-66)
66. ​**AM 2.20**​ Filosoofilises kirjanduses osutatakse mitmesele realiseeritavusele sageli kui peamisele takistusele “reduktsiooni” teel. Ent nagu näitab Brooks (1994), ei ole mitmesel realiseeritavusel kuigi palju tegemist reduktiivsete seletuste tegeliku rolliga teaduses. Näiteks bioloogias on tiibu võimalik realiseerida väga mitmel viisil, kuid sellegipoolest pakuvad bioloogid reduktiivseid seletusi. Veel enamgi, isegi paljud sageli redutseeritavuse õpikunäideteks peetavad füüsikalised fenomenide (kasvõi temperatuur) võivad, nagu osutavad Wilson (1985) ja Churchland (1986), tegelikkuses realiseeruda mitmeti. [↑](#footnote-ref-67)
67. [tõlkemärkus:] Nii on see inglise keeles, kus mõlema kohta on üksainus sõna “perception”. Eesti keeles on siin kummagi kohta oma sõna: vastavalt “tajumine” ja “taju”. [↑](#footnote-ref-68)
68. Omavolitsesin: originaalis on “The objections to these notions”, kuid järgnevas kolmes punktis on juttu ainult kontseptuaalsetest tõdedest (nagu võiski arvata lõigu esimese lause põhjal), aga mitte empiirilistest tõdedest. [↑](#footnote-ref-69)
69. “utility 800” — mida see tähendab? Ja mida tähendab järgnev “metabolizes with efficiency 75 percent”? [↑](#footnote-ref-70)
70. On ikka, näiteks [kumeruse](https://et.wikipedia.org/wiki/Kumer_hulknurk) mõiste abil. [↑](#footnote-ref-71)
71. Mida see tähendab, kas see on termin? Väljend “background statement” on inglise keeles küll täiesti olemas ja kasutusel, aga [hoopis teises tähenduses](https://ieltsliz.com/background-statement-practice/). Selgust ei toonud mulle ka Quine’i originaalartikli ([“Two Dogmas of Empiricism”](http://www.ditext.com/quine/quine.html)) lugemine — mulle jäi täiesti arusaamatuks, kuidas on sealt võimalik välja lugeda seda, mida Chalmers ütleb seal öeldud olevat.  — Venelasel on siinkohal “базовые положения”, võtsin malli.. [↑](#footnote-ref-72)
72. Kuidas on eesti keeles “antecedent” ja “consequent”? Internetist olen leidnud: 1) “antetsendent” ja “konsekvent”; 2) “eesliige” ja “tagaliige”. Millegipärast mulle tundub, et eesti keeles peaks nende mõistete kohta ka paremaid sõnu leiduma, need kaks paari ei meeldi mulle kumbki. [↑](#footnote-ref-73)
73. ​**AM 2.21**​ Sellele võidakse vastata, et sellest, mis oleks mõiste “vesi” referent, kui aktuaalseks maailmaks osutuks XYZ-maailm, ei tohiks kõneldagi, sest säärasel juhul oleks sõna, mis kõlab nagu “vesi”, juba hoopis teine sõna! Lugeja, keda see asjaolu häirib, võib mõiste “vesi” referendi asemel kõnelda sõnaga “vesi” samakõlalise sõna referendist; veelgi parem oleks vaadelda neid stsenaariume kui *episteemilisi* võimalikkusi (avaras mõttes) ja konditsionaale kui episteemilisi konditsionaale; nii saame üle murest sõnade olemuslike omaduste pärast. Nii või teisiti on jutu üldine mõte — mõiste osutuse sõltumine aktuaalses maailmas ​**[lk.365]**​ sellest, missuguseks see maailm osutub — selge, ükskõik kuidas seda stsenaariumi kirjeldada. Põhitekstis lähen ma sellest nüansist enamasti vaikides mööda. [↑](#footnote-ref-74)
74. “… maps the XYZ world to XYZ, and the H2O world to H2O” [↑](#footnote-ref-75)
75. ​**AM 2.22**​ Mitte kõik autorid ei ole kindlad, et Kripkel ja Putnamil on õigus, kui nad kinnitavad, et vesi on tingimata H₂O ning et XYZ ei ole vesi; kahtlusi väljendab näiteks Lewis (1994). Meie keelepraktika jääks vististi suuremas osas samasuguseks ka siis, kui me osutaksime sõnaga “vesi” kontrafaktilistes maailmades vesijale ollusele. Asjaolu, et seda küsimust saab kergesti vaadelda mõlemat pidi, viib mõttele, et mingi fenomeni, näiteks vee seletamisel ei ole sekundaarset intensiooni puudutavad seigad tegelikult nii olulised. Tõepoolest, tõsiasi, et sellistes olukordades võib alati öelda “vesi” asemel “vesijas ollus”, viib mõttele, et aposterioorsest paratamatusest ei olene seletamise, füsikalismi jne. vaatlemisel arvatavasti kuigi palju. Just seesugustel põhjustel läheb Siewert (1994) aposterioorse paratamatuse küsimusest vaikides mööda. Minul oli kiusatus teha sedasama, kuid lõpuks otsustasin, et see kahemõõtmeline mudel on juba iseenesestki huvitav. [↑](#footnote-ref-76)
76. ​**AM 2.23**​ Muidugi võib olla piirjuhte, kus küsimusele, kas mõiste vastab mingile olukorrale, ei leidugi kindlapiirilist vastust, juhul kui maailm, millest me kõneleme, osutub aktuaalseks maailmaks. Sellest pole midagi katki: me võime lubada ebamäärasust primaarses intensioonis, samuti nagu me mõnikord lubame osutuse ebamäärasust omaenda maailmas. Võib ka juhtuda, et mingil mõistel on mõnes maailmas mitu võrdselt head kandidaat referenti (nagu võib-olla massi mõistele relativistlikus maailmas, kus see mõiste võib viidata kas seisumassile või relativistlikule massile); nii nagu me sallisime mitmele eri referendile viitamist säärastel aktuaalsetel juhtudel, ei tohiks me keelata seda ka primaarsele intensioonile.

Mõnel piirjuhul võib ühe või teise objekti paigutamine mingi mõiste ekstensiooni alla sõltuda isegi mitmesugustest juhuslikest ajaloolistest teguritest. Selle kohta annab virgutavat mõtlemisainet Wilsoni artikkel (1982), mis käsitleb teiste näidete seas hüpoteetilist juhtumit druiididega: kui nad näevad lennukit endist esmakordselt üle lendamas, liigitavad nad ta arvatavasti linnuks, ent kui nad leiavad džunglist allakukkunud lennuki ja pole ühtki lennukit varem näinud, siis nad teda vist linnuks ei pea. Me võime nende kahe stsenaariumi kohta öelda, et need on aktuaalse maailma realiseerumise kaks võimalikku viisi — nii jääb primaarne intensioon kindlakujuliseks ja detailseks ent võime ka öelda, et niisugustel juhtudel on määramatus juba primaarse intensiooni tuumikalal​. Nii või teisiti ei ole väike laialivalguvus primaarse intensiooni äärealadel sugugi vastuolus sellega, kuidas ma oma mudelit rakendan. [↑](#footnote-ref-77)
77. Omavolitsesin: originaalis “general relativity”, aga see on autori ebatäpsus: massi ja energia ekvivalentsus (E = mc²) kehtib ka juba erirelatiivsusteoorias. [↑](#footnote-ref-78)
78. ​**AM 2.24**​ Vt. Field 1973. Usun, et analüüs primaarsete intensioonide kaudu lubab pidada tähendusnihkeid palju harvemaks nähtuseks, kui tavaliselt arvatakse. Näiteks ei anna relatiivsusteooria meile põhjust arvata, nagu oleks termini “mass” primaarne intensioon eelmise sajandiga võrreldes muutunud, kuigi kahtlemata on muutunud meie uskumused aktuaalse maailma kohta. (Sarnastel põhjustel võiks proovida analüüsi primaarsete intensioonide kaudu pidada argumendiks “tähendusholismi” vastu mõtlemises.) Kui primaarsete intensioonide “arengu” käigus üldse midagi toimub, siis arvatavasti midagi hoopis peenemat, nagu näiteks Wilsoni (1982) näidetes, kus tuum jääb samaks, kuid äärealadel võivad liigituspraktikad ajalooliste juhuste tõttu muutuda. Kuid see kõik väärib palju põhjalikumat käsitlust. [↑](#footnote-ref-79)
79. “meaning holism”, vt. [PhilPapers](https://philpapers.org/browse/meaning-holism). [↑](#footnote-ref-80)
80. *Deskriptiivne osutusteooria* ehk *deskriptivistlik nimeteooria* ütleb, et nime semantiline sisu on identne kirjeldusega, mida kõnelejad temaga seostavad. (Vt. [Wikipedia](https://en.wikipedia.org/wiki/Descriptivist_theory_of_names)). [↑](#footnote-ref-81)
81. Läbiv raskus: kuidas kirjeldada lühidalt funktsiooni, mille määramispiirkonnaks on X ja muutumispiirkonnaks Y? Originaalis on siin “functions from possible worlds to extensions”, ka vene keeles saab hästi hakkama: “функций от возможных миров к экстенсионалам”. Eesti keeles ei tundu säärast elegantset ja kompaktset keelendit leiduvat; alternatiivideks on säärased lohisevad sõnastused nagu “funktsioon, mille määramispiirkonnaks/argumentideks on X ja muutumispiirkonnaks​/väärtusteks on Y”, aga need on juba nii kohmakad, et hägustaksid kogu lause mõtte. [↑](#footnote-ref-82)
82. *Põhjusliku osutusteooria* (vt. [Wikipedia](https://en.wikipedia.org/wiki/Causal_theory_of_reference)) järgi fikseeritakse termini tähendus esialgselt nimetamisaktiga, ja järgnevatele kasutustele annab osutuse põhjuslik ahel, mis kulgeb nimetamisaktist kasutuseni. [↑](#footnote-ref-83)
83. ​**AM 2.25**​ Mõned erinevused: 1) Kaplani “sisu” vastab väga täpselt sekundaarsele intensioonile, “karakterit” aga kirjeldab ta kui funktsiooni kontekstide hulgast sisude hulka, erinevalt primaarsest intensioonist, mis on funktsioon kontekstide hulgast ekstensioonide hulka. Pärast mõiste rigidiseerimist aga saab tema karakterist otseselt tuletada primaarse intensiooni ja ​**[lk.366]**​ ümberpöördult. Mina kõnelen sümmeetria ja lihtsuse huvides esimesest​. 2) Kaplan käsitleb oma kontseptsioonis indeksikaalseid ja demonstratiivseid sõnu nagu “mina” ja “see”, kuid ei laienda käsitlust looduslikele sõnadele nagu “vesi”, kuna tema arvates osutab sõna “vesi” kõigis kontekstides H₂O-le (samakõlaline sõna Teisikmaal on lihtsalt teine sõna) ja tema arvates ei kuulu referendi fikseerimise protsess siinsel juhul semantikasse, vaid “metasemantikasse”. Nagu varemgi, ei olene minu eesmärgid kuigi palju sellest, kas referendi fikseerimisega peaks tegelema metasemantika või semantika; oluline on vaid see, et referendi fikseering sõltub mingil viisil asjade seisust aktuaalses maailmas. [↑](#footnote-ref-84)
84. Ei saa aru: mis on siin see “esimene” (“former”)? [↑](#footnote-ref-85)
85. ​**AM 2.26**​ Võib tunduda, et primaarne intensioon on täpselt defineeritav ainult niisuguste võimalike maailmade puhul, mille keskpunktiks on indiviidid, kes mõtlevad mingit kindlat mõtet või ütlevad mingeid kindlaid lausungeid. Arvan siiski, et primaarne intensioon on loomulikul viisil laiendatav mahukamale maailmade klassile: me võime jääda omaenda maailma kuuluva mõiste juurde ja vaadelda, kuidas see rakendub teistele maailmadele, kui vaadelda neid aktuaalsetena (vt. Chalmers 1994c), kuigi mõnes maailmas võib mõiste referent jääda määratlematuks. Kuid järgnevaid mõttekäike see oluliselt ei muuda. [↑](#footnote-ref-86)
86. ​**AM 2.27**​ Täpsustame, et rangelt võttes osutab primaarne intensioon vedelikule meie ajaloolises keskkonnas: kui ma reisin Teisikmaale ja ütlen seal “vesi”, on referendiks endiselt H₂O. [↑](#footnote-ref-87)
87. Autor tähistab funktsiooni argumente sõnaga “parameters” — ebaõnnestunud sõnakasutus, sest tavaliselt tähistatakse sõnaga “parameter” funktsiooni *parameetrit*​ (s.t. mingit funktsiooni definitsioonis figureerivat konstanti, millel võib olla mitmesuguseid väärtusi, kuid millele antakse mingi pikema mõttekäigu vältel üks kindel väärtus); segadust suurendavad veelgi *parameetrilised funktsioonid* (vt. [Parametric equation](https://en.wikipedia.org/wiki/Parametric%20equation)​) (funktsiooni parameetriline esitusviis), mis on veel hoopis kolmas asi. Üldkäibivad sõnad on [inglise keeles](https://en.wikipedia.org/wiki/Function_%28mathematics%29) “argument” ja “input value”, [eesti keeles](https://et.wikipedia.org/wiki/Funktsioon_%28matemaatika%29) “argument” ja (uuemal ajal) “sisendväärtus”; ega viimasel ju ka midagi viga ei ole, aga valisin siiski “argumendi” kui harjumuspärasema nii minule kui usutavasti ka lugejale. [↑](#footnote-ref-88)
88. Lisasin selle sulgudes sõna omalt poolt, et lugeja ei satuks segadusse: mis mõttes “teised”, mis need “ühed” siis olid? (Indeksikaalsed mõisted said “ühtedeks” juba neli lõiku tagasi.) [↑](#footnote-ref-89)
89. “truth in virtue of meaning” — kuidas paremini tõlkida seda “in virtue of”? [↑](#footnote-ref-90)
90. Omavolitsesin, tegin voolujoonelisemaks; originaalilähedasem ja kohmakam oleks “… on lause S paratamatult tõene, kui S osutuks tõeseks olenemata sellest, missuguseks osutub aktuaalne maailm”. [↑](#footnote-ref-91)
91. ​**AM 2.28**​ Käesoleva jao teise ja kolmanda mõttekäigu — s.t. Quine’i empiirilise revideeritavuse ja Kripke aposterioorse paratamatuse — omavaheline seos on keeruline ja huvitav. Nagu tähendab Kripke, lubab tema mudel lahendada mõned Quine’i tõstatatud probleemid, kuid mitte kõiki. Kripke analüüs annab seletuse intensioonide aposterioorsetele revideerimistele ja seega ka muutustele “tähenduse” mõnedes aspektides. Kahemõõtmeline analüüs on aga kooskõlas lausetele aktuaalses maailmas leitud tõeväärtuste kirjeldusega üheainsa intensiooni abil ega seleta seega võimalust, millest kõneleb Quine — võimalust, et piisava empiirilise andmestiku valguses võivad mõned aprioorseteks peetud kontseptuaalsed tõed osutada aktuaalses maailmas vääraks. Minule tundub, et säärased arvatavad kontseptuaalsed tõed lihtsalt ei olegi kontseptuaalsed tõed, kuigi nad võivad olla sellele headeks lähendusteks. [↑](#footnote-ref-92)
92. ​**AM 2.29**​ Kahemõõtmelise mudeli rakendamisel mõtte sisu kindlakstegemiseks tuleb esile üks peensus: mõnikord võib mõtlemine tunnustada potentsiaalse keskkonnana isegi niisugust keskpunktiga maailma, mis ei sisalda sellesamast mõttest endast ainsatki eksemplari. Näiteks kui ma mõtlen “ma olen koomas”, tunnustan ma keskpunktiga maailmu, mille keskpunktis asuv indiviid on koomas, ükskõik kas sel indiviidil on selles maailmas mõtteid või mitte. Seega tuleb mõtete primaarsete intensioonide ja primaarpropositsioonide defineerimisel ilmutada rohkem ettevaatust, kui ilmutasin mina siin keelt käsitledes. [↑](#footnote-ref-93)
93. ​**AM 2.30**​ Maailmu tuleb käsitada prelingvistiliselt, võib-olla baaskvaliteetide jaotustena. Arvatavasti on kõige parem mitte käsitada maailmu kui mingite lausete kogumeid, kuna laused on juba maailma kirjeldused ja me oleme näinud, et nad võivad kirjeldada maailma rohkem kui ühel viisil. Kui vaadelda maailma lausete kogumitena, lähevad need erinevused kaduma. Võib-olla saaks maailmu vaadelda kui propositsioonide kogumeid (Adams 1974) — kui mõista propositsioone sobival viisil, kas maksimaalsete omadustena (Stalnaker 1976), asjaseisudena maailmas (Plantinga 1976), struktuursete universaalidena (Forrest 1986) või konkreetsete objektidena, mis on analoogilised meie oma maailmaga (Lewis 1986a). Võib-olla saaks võtta selle mõiste ka lihtsalt elementaarmõisteks. Igal juhul on võimalikest maailmadest kõnelemiseks sama palju või sama vähe alust kui kõnelemiseks võimalikkusest ja paratamatusest üldse. Nagu matemaatilised mõisted, nii ka neid modaalseid mõisteid võib kasuga rakendada juba siis, kui rahuldavat ontoloogilist analüüsi veel ei olegi.

​**[lk.367]**​ Siin ja edaspidi käsitlen ma maailmu alati “kvalitatiivselt” ja jätan “sellesuse” [​*haecceity*​] küsimused kõrvale. Teisiti öeldes: ma loen kaks maailma identseteks, kui nad on kvalitatiivselt identsed, ja mind ei huvita küsimus, kas nendes maailmades võivad inimestel olla erinevad “identsused”. (Mõned autorid on püüdnud näidata, et võivad.) Need maailmade omavahelise identsuse küsimused viivad paljude huvitavate probleemideni, mis on aga minu võimalike maailmade mudeli rakendamise seisukohalt enamikus irrelevantsed. [↑](#footnote-ref-94)
94. Jätsin vahepealt tõlkimata sõnakese “or”, mis tundub olevat autori näpuviga: “if propositions are understood appropriately, **or** as maximal properties …, or as states of affairs …, or …” [↑](#footnote-ref-95)
95. “haecceity” — ei leia head vastet. “Sellesusest” loogilisem oleks “seesus”, aga see kõlaks kole halvasti, ärataks kohatuid assotsiatsioone ja ei oleks lugejale vist ka esimesel hoobil arusaadav. [↑](#footnote-ref-96)
96. Autoril on siin “in the first sense; and … in the second”, s.t. “esimeses mõttes; ja … teises mõttes”, aga see kippus mulle iga kord uuesti lugedes esimesel pilgul arusaamatuks jääma; panin “mõtete” asemele “intensioonid”, et jutt oleks selgem. [↑](#footnote-ref-97)
97. ​**AM 2.31**​ Tähenduse ja intensiooni samasust võidakse eitada näiteks matemaatiliste terminite puhul. Sageli kinnitatakse, et näiteks lause “algarvude hulk on lõpmatu” ei ole tõene tähenduse põhjal, olgugi ta kõigis võimalikes maailmades tõene; niiviisi ütlejad ei ole tähenduse ja intensiooni samastamisega arvatavasti nõus.

Veel üks tees, millega kõik ei tarvitse nõustuda, on, et näiteks sõna “vesi” primaarne intensioon on üks osa tema tähendusest; võib-olla arvavad nad, et sõna tähenduse ammendab tema osutus ning et primaarsed intensioonid ei kuulu semantika, vaid pragmaatika alla. Mõned jällegi ei ole nõus pidama tähenduse osaks sekundaarset intensiooni. Nii või teisiti ei toetu ma kusagil sõna “tähendus” nii- või teistsugusele tarvitusviisile. Mind huvitab intensioonipõhine tõde olenemata sellest, kas intensioon on või ei ole samane tähendusega. [↑](#footnote-ref-98)
98. ​**AM 2.32**​ See mõeldavuse definitsioon on lähedane definitsioonile, mille andis Yablo (1993): lause *P* on mõeldav, kui on võimalik ette kujutada maailma, mille varal lauset *P* verifitseerida saab. Erinevus on selles, et Yablo lisaklausel “mille varal lauset *P* saab verifitseerida” ei välista võimalust kirjeldada mõeldavaid olukordi ekslikult, mistõttu see mõeldavuse kuju sobib laitmatult parimal juhul vaid võimalikkuse määratlemiseks. Minu definitsiooni korral see segaduste allikas kaob. Muidugi ilmub sama raskus nüüd teises kohas: lünk esimesel pilgul mõeldavateks peetavate lausete ja päriselt mõeldavate lausete vahel on nüüd suurem, mistõttu mõeldavuse üle otsustades tuleb olla ettevaatlikum. [↑](#footnote-ref-99)
99. ​**AM 2.33**​ Seda mõttekäiku saab vististi rakendada Descartes’i mõttele, et kuna ta võib mõelda, nagu tal ei oleks keha, siis on võimalik, et ta on kehatu, ja seega, et ta ei ole füüsiline objekt (kuna iga füüsiline entiteet on paratamatult kehaline). Lause “mul ei ole keha” võib olla 1-mõeldav ja seega 1-võimalik, kuid sellest ei järeldu, et ta oleks 2-mõeldav või 2-võimalik. Seevastu kui ma ütlen, et lause “mul on keha” on paratamatu juhul, kui ma olen füüsiline objekt, siis see on 2-, mitte 1-paratamatus. (Descartes’i “mina”-mõiste primaarne intensioon osutab suvalises maailmas indiviidile selle maailma keskpunktis; sekundaarne intensioon osutab kõigis maailmades Descartes’ile.) [↑](#footnote-ref-100)
100. ​**AM 2.34**​ Võidakse öelda, et metafüüsilises paratamatuses pole midagi eriliselt “metafüüsilist”. See tähendaks, et ta on vaid üks kontseptuaalse paratamatuse erikujusid aposterioorse semantilise nihkega, mis tuleneb meie mõistete kahemõõtmelisusest. Arusaama kohta, et aposterioorne paratamatus tuleneb samavõrd konventsioonidest kui metafüüsikast, vt. Putnam 1983 ja Sidelle 1989, 1992. [↑](#footnote-ref-101)
101. Omavolitsesin: originaalis on “… implies conceivability of the statement (in the relevant sense)”; mul andis mõelda, mida see sulgudes märkus tähendab: kõigepealt, kas see käib üldse “conceivability” või “statement’i” kohta (millele ta on tekstis lähemal). Kui minu arusaamine oli õige (nagu ma loodan), siis on minu tõlge selgem. [↑](#footnote-ref-102)
102. Omavolitsesin: originaalis on “neither of which is dependent on *a posteriori* facts **in this case**​”. See “case” tähendab juhtu, kus “lause on loogiliselt võimalik või paratamatu oma primaarsete intensioonide järgi” (vt. lõigu algust); jätsin “in this case” ära, sest mulle isiklikult on ta mõjunud pigem segastavalt kui selgitavalt: lugedes “in this case”, tekib mul kõigepealt küsimus, mis “juhtudest” autor räägib, ja kulub mõni sekund, enne kui ma saan aru, et ta peab silmas primaarseid intensioone, millest on ju nagunii jutt. Kui “in this case” ära jätta, siis jääb kõik arusaadavaks, ja lisaks jääb ära eksitav küsimus. (Äkki ma eksin ja autor mõtleb midagi muud?) [↑](#footnote-ref-103)
103. “tablehood” — tahaks panna “laudsus”, aga kardan, et see võtaks lugeja õhku ahmima. [↑](#footnote-ref-104)
104. ​**AM 2.35**​ Horgan (1984c) arendab ja põhjendab seisukohta, et kõik kõrgtaseme faktid on mikrofüüsikaliste faktide suhtes loogiliselt superventsed. Tema sõnastuse järgi seovad neid fakte mikrofüüsikaga “semantilised kitsendused”, nii et maailmas ei ole midagi peale mikrofüüsika ja “kosmilise hermeneutika”. Samas väldib ta väga silmatorkavalt teadvuselamuse probleemi. Loogilise superventsuse teesi mitmesuguste variantide pooldajate hulka kuuluvad veel Jackson (1993), Kirk (1974) ja Lewis (1994). [↑](#footnote-ref-105)
105. ​**AM 2.36**​ Argumente selle kasuks, et faktid abstraktsete entiteetide kohta on füüsika suhtes loogiliselt superventsed, toob Armstrong (1982). [↑](#footnote-ref-106)
106. Siin on küll midagi valesti: need kaks ei ole ju samaväärsed! Kuna autori mõtteviisiga sobib palju ennemini teine võimalus, siis on esimese võimaluse sõnastuses arvatavasti mingi viga. [↑](#footnote-ref-107)
107. Mis mõttes? Nagu põhjusliku seose olemasolu neli lõiku tagasi? [↑](#footnote-ref-108)
108. Missugustel, kus? [↑](#footnote-ref-109)
109. “tissues” — ei oska teisiti tõlkida; venelasel on “из ткани”, aga see ei sobi originaali mitmusega. [↑](#footnote-ref-110)
110. ​**AM 2.37**​ Teadvuselamus annab sellesse primaarsesse intensiooni arvatavasti oma panuse, juhul kui mingi osa vee individuatsioonist moodustab elamus, mida ta esile kutsub. Mingi osa on ​**[lk.368]**​ kindlasti ka indeksikaalsusel, nagu näitab sõna “meie” sisaldumine ühendis “läbipaistev, joogikõlblik vedelik meie keskkonnas”. Need faktid ei õõnesta loogilist superventsust, kui teadvuselamuste ja indeksikaalsuse panust mitte arvestada. [↑](#footnote-ref-111)
111. “We can run things either way.” — ei saa päriselt aru. [↑](#footnote-ref-112)
112. “caching out” (venelasel “избавлением”) — kas nii on õige? (Kas see, et klammerdaja lööb klambreid ja et klambreid lüüakse klammerdajaga, on põhjuslikud rollid?) [↑](#footnote-ref-113)
113. ​**AM 2.38**​ Nagu see toimub Ramsey-lausete[fn:51Ramsey] meetodi puhul teoreetiliste terminite rakendamiseks; vt. Lewis 1972. [↑](#footnote-ref-114)
114. “caching out” (venelasel “избавлением”) — kas nii on õige? (Kas see, et klammerdaja lööb klambreid ja et klambreid lüüakse klammerdajaga, on põhjuslikud rollid?) [↑](#footnote-ref-115)
115. ​**AM 2.39**​ Näite autor on Ned Block, kes pakkus selle välja nuputamisülesandena. [↑](#footnote-ref-116)
116. ​**AM 2.40**​ Sarnase väite, mille järgi superventsus eeldab analüüsitavust, on esitanud Jackson (1993) ja Lewis (1994). [↑](#footnote-ref-117)
117. ​**AM 2.41**​ Ühe alternatiivse tõepärase käsitluse järgi peab objekt, et olla punane, kalduma põhjustama vaid otsustusi punase värvuse kohta. Selline tõlgendus vabastaks meid äsjaöeldud raskusest, kuna neid otsustusi võib üsna kindlalt pidada füüsika suhtes loogiliselt superventseteks. [↑](#footnote-ref-118)
118. Oletus, originaal on pisut ebamäärane (“that this sort of thing counts”); oletan, et “this sort of thing” tähendab seletust, kus mainitud viimane samm on puudu. (Venelase tõlge on veelgi segasem ja, nagu mulle tundub, ka mitte õige: “На практике наши требования к объяснению достаточно лояльны для того, чтобы это имело значение.”) [↑](#footnote-ref-119)
119. ​**AM 2.42**​ Võib-olla täpsustusega, et kui mina usun midagi Bill Clintoni kohta, siis usub minu teisik midagi Clintoni teisiku kohta. Nagu ikka, võib need maailmadevahelise identsuse küsimused kõrvale jätta. [↑](#footnote-ref-120)
120. ​**AM 2.43**​ Kõige lähemale jõuab sellisele argumendile Kripke (1982) versioon Wittgensteini mõttekäigust; väide seisneb põhimõtteliselt selles, et füüsikalistest ja fenomenaalsetest faktidest ei ole võimalik tuletada intentsionaalseid fakte, kuna vajalikku tuletuskäiku ei saa pakkuda ei intentsionaalsete mõistete füüsiline, funktsionaalne ega fenomenaalne analüüs. Tema argumendid (eriti need, mida ta esitab funktsionaalse analüüsi vastu) on vaieldavad, kuid, nagu ülal tähendatud, ei seisne mõttekäigu lõppjäreldus nii või teisiti mitte selles, et intentsionaalsed faktid on sõltumatu lisandus, vaid et nad ei olegi rangelt võttes faktid.

Kui nõustuda Kripke–Wittgensteini argumendiga seesuguse tuletatavuse vastu, satub intentsionaalsus samasugusesse olukorda nagu moraal (vt. [allpool](#ORG.2.5.1.3)). Mõlemal juhul 1) ei ole B-faktid nähtavasti  A-faktidest kontseptuaalselt tuletatavad, ent 2) kui meie maailmas on B-fakte, siis kehtivad nad ka igas meie maailmast A-eristamatus mõeldavas maailmas. Ainus mõistlik järeldus on, et B-fakte ei olegi rangelt võttes olemas ning et B-atributsioone tuleb käsitleda kuidagi deflatiivses vaimus​. Võimalus, et B-faktid on fundamentaalsed lisafaktid, langeb ära mõeldavuskaalutlustel​, milledest tuleneb, et kui B-faktid üldse realiseeruvad, siis peab neil olema aprioorne seos A-faktidega. [↑](#footnote-ref-121)
121. “B-attributions must be treated in some deflationary way” — sellest kohast ma ei saanud aru, ja allikat Kripke 1982 (raamat, “Wittgenstein on Rule-Following and Private Language”) ma Internetist kätte ei saanud. [↑](#footnote-ref-122)
122. ​**AM 2.44**​ Kui subjektiivsed moraalsed faktid on olemas, siis on moraalsetel atributsioonidel selgepiirilised tõesustingimused, mis aga olenevad atribueerijast​. Säärasel juhul on moraalimõistetel olemas indeksikaalne primaarne intensioon ning kehtib loogiline superventsus, kui mitte arvestada indeksikaalsust. Seda analüüsi toetavad “subjektivistliku moraalse realismi” pooldajad (SayreMcCord 1989), kelle järgi “hea” tähendab midagi umbes niisugust nagu “hea minule” või “hea minu kogukonnale”. Oma subjektiivsuse tõttu on see siiski väga nõrka liiki realism. Näiteks tuleneb sellest vaatest, et kaks inimest, kes vaidlevad selle üle, mis on hea, ei pruugigi olla erinval arvamusel. [↑](#footnote-ref-123)
123. Originaalis “Similarly”, aga mis on siin millega sarnane? (Ja üldse, ma ei saa järgnevast lausest aru.) [↑](#footnote-ref-124)
124. ​**AM 2.45**​ Kripke (1972) argumendid, näiteks need, mis puudutavad nime “Gödel” osutust erinevates olukordades, viivad mõttele, et nime lausumisega kaasnevat primaarset intensiooni ei saa üldjuhul väljendada lühida kirjeldusega. Neist võiks järeldada ka, et et seda primaarset intensiooni ei saa väljendada ühegi lõpliku kirjeldusega, kuigi selles ma ei ole enam nii kindel. (Kindlasti näitavad nad aga seda, et iga selline kirjeldus peab sisaldama metalingvistilist elementi ning tingimust, mis nõuab kohast põhjuslikku seost agendiga.) Kuid neis argumentides ei ole midagi, mis viitaks võimalusele, et nimedel (nagu neid igal konkreetsel juhul lausutakse) primaarne intensioon üldse puudub. Tõepoolest, Kripke rajabki oma mõttekäigud vaatlusele, kuidas nime osutus sõltub ​**[lk.369]**​ sellest, missuguseks osutub aktuaalne maailm, s.t. nime primaarse intensiooni kindlakstegemisele mitmesugustes keskpunktiga maailmades. [↑](#footnote-ref-125)
125. ​**AM 2.46**​ Selle küsimuse väga selget käsitlust vt. Nagel (1986). [↑](#footnote-ref-126)
126. “localized facts” — mis need on? [↑](#footnote-ref-127)
127. ​**AM 2.47**​ Miks ei ole ülaltoodud argumentidel loogilise superventsuse kasuks jõudu negatiivsete faktide puhul? Mõeldavuse argument​ ei tööta, nagu nähtub näitest inglitega. Epistemoloogiline argument ei tööta, kuna piiramatu mahuga universaalsete väidete kontrollimine valmistab ilmseid epistemoloogilisi raskusi. (Meil pole võimalik kindlalt teada, et ingleid ei ole olemas.). Analüüsitavuse argument ei tööta, kuna neid negatiivseid fakte ei saa täielikult analüüsida positiivsete faktide kaudu (kui me ei võta kasutusele teist järku fakti “ja see on kõik”). [↑](#footnote-ref-128)
128. Läbiv raskus: kuidas oleks “argument from X” eesti keeles? Esialgu tõlgin “argument from conceivability” “mõeldavuse argumendiks” ja sarnaselt edasistel juhtudel, aga kas ei saaks paremini? Ka vene keeles on olemas hea väljend “аргумент от …”, aga eesti keeles ei tundu midagi nii head olevat. [↑](#footnote-ref-129)
129. ​**AM 2.48**​ Kuidas õnnestub seadustel neist loogilise superventsuse kasuks toodud argumentidest mööda hiilida? Mõeldavuse argument ei tööta, nagu äsjatoodud näitest nägime. Epistemoloogiline argument ei tööta, kuna seaduste ja põhjuslike seoste epistemoloogia valmistab, nagu näitavad Hume’i skeptilised rünnakud, ilmseid raskusi. Analüütiline argument ei tööta, kuna seaduse kehtimine eeldab universaalset reeglipära, mis toimiks ka kontrafaktuaalide puhul, vastavaid kontrafaktuaale aga ei ole — kogu lugupidamise juures Lewise (1973) vastu — võimalik analüüsida maailmaloo partikulaarsete faktide kaudu. Võib leida kõige erinevamaid kontrafaktuaale, mille tõesus on ühildatav maailma aegruumiloo konkreetsete faktidega. [↑](#footnote-ref-130)
130. “Humean” — kuidas see eesti keeles oleks? [↑](#footnote-ref-131)
131. ​**AM 2.49**​ Hume’likke vaateid seaduste ja põhjuslikkuse kohta toetavad Lewis (1986b), Mackie (1974) ja Skyrms (1980). Vastupidiseid seisukohti esindavad Arm­strong (1982), Car­roll (1994), Dretske (1977), Mol­nar (1969) ja Tooley (1977). [↑](#footnote-ref-132)
132. “perhaps a time-slice of physical facts” — kas sai õigesti? [↑](#footnote-ref-133)
133. “underwritten by meanings” [↑](#footnote-ref-134)
134. ​**AM 2.50**​ On ka autoreid — teiste seas Armstrong (1982), Horgan (1984c), Jackson (1993), Lewis (1994) ja Nagel (1974) —, kes leiavad, et loogiline superventsus on, vastupidi, pigem reegel kui erand. [↑](#footnote-ref-135)
135. ​**AM 2.51**​ Seda teemat on põhjalikumalt käsitlenud Horgan ja Timmons (1992b). [↑](#footnote-ref-136)
136. ​**AM 3.1**​ Kirk (1974) kirjeldab värvikalt sombit ja, kirjeldades vastavaid vahepealseid juhtumeid, visandab koguni olukorra, kus me võiksime jääda uskuma, et keegi ongi aktuaalses maailmas sombiks muutunud. Campbell (1970) käsitleb samamoodi “inimese imitatsiooni”, kes on tavalise inimesega füüsikaliselt identne, kuid kellel puudub täielikult teadvuselamus. [↑](#footnote-ref-137)
137. Jätsin tõlkimata sulgudes märkuse, mis käsitleb puhtalt inglise keele siseasju (asesõnu) ja ei anna eestikeelsele lugejale midagi. [↑](#footnote-ref-138)
138. “burden of proof” — õigusliku terminina on see eesti keeles “tõendamiskohustus” (sellepealkirjaline artikkel leidub [eestikeelses Vikipeedias](https://et.wikipedia.org/wiki/T%C3%B5endamiskohustus)). [↑](#footnote-ref-139)
139. ​**AM 3.2**​ Niisugusel kaudsel teel kaitseb sombide loogilist võimalikkust Kirk (1974). [↑](#footnote-ref-140)
140. ​**AM 3.3**​ Jacoby (1990) esitab suurepärase tõestuse, et mõeldavusargumendid ei too teadvuse funktsionalistlikesse käsitlustesse suuremaid raskusi kui materialistlikesse käsitlustesse üldiselt. Tema peab seda argumendiks funktsionalistlike käsitluste kasuks, mina aga argumendiks materialistlike käsitluste vastu. [↑](#footnote-ref-141)
141. ​**AM 3.4**​ Õigupoolest ei vahetuks punane sellisel juhul mitte sinise, vaid kollasega, kuna mõlemad on oma telgede positiivsetes otstes. Need on aga ebaolulised detailid. Inimese värvusruumi keerukuste üle on huvitavalt arutlenud Hardin (1988). [↑](#footnote-ref-142)
142. ​**AM 3.5**​ Seda möönab tõepoolest ka Hardin (1987, lk. 138) ise. Ta ütleb, et selline inversioon oleks vaid “veider”, kuid mitte “kontseptuaalselt ebakoherentne”. [↑](#footnote-ref-143)
143. “The extra neurophysiological properties that are **constrained** in such a case …” — ei oska sellest aru saada (tõlkisin venelase eeskujul: “Дополнительные нейрофизиологические свойства, не реализованные в подобном случае …”); siin on justkui midagi valesti, sest jutt on ju (vt. lõigu esimest lauset) süsteemist, mille “functional organization stays constant” — kas ei peaks siis olema täpses vastavuses ju ka “neurophysiological properties”?! [↑](#footnote-ref-144)
144. ​**AM 3.6**​ Midagi sarnast kõneleb Gunderson (1970) “uurimuslikust asümmeetriast” esimese isiku ja kolmanda isiku ütluste vahel. [↑](#footnote-ref-145)
145. ​**AM 3.7**​ Thompson (1992) juhib tähelepanu sellele, et ka must-valges toas võib Mary ikkagi värvuselamusi saada, näiteks silmamunadele vajutades. Sellest raskusest ülesaamiseks tuleks ehk hoopis eeldada, et Mary on sündimisest saadik värvipime. [↑](#footnote-ref-146)
146. “change of subject” — selle kohta peaks olema mingi termin, aga ma pole seda leidnud. (Veel parem on venelase “подмена предмета”.) [↑](#footnote-ref-147)
147. ​**AM 3.8**​ ​**[lk.370]**​ Vitalismile viitavad selles kontekstis näiteks Churchland (1995, lk. 193) ja Dennett (1991, lk. 281). [↑](#footnote-ref-148)
148. “production systems”; “reegelsüsteemi” võtsin [siit](https://tepandi.ee/is-loeng.pdf) (lk. 82). [↑](#footnote-ref-149)
149. ​**AM 3.9**​ Oma raamatu teise osa lõpus lubab Dennett kolmandas osas näidata, miks võib tema funktsionaalne kontseptsioon seletada teadvuse kohta kõike, mis seletust nõuab. Kuid lubatud argumente pole kerge leida. Suure osa tema arutlustest moodustavad tähelepanekud kognitiivsete protsesside kohta, millega minu vaadetega inimene võiks rõõmuga nõustuda. Küsimus ei ole selles, kas tema arusaamine kognitiivsest töötlusest on õige, vaid kas see seletab elamust. Otsustav argument on nähtavasti dialoogis lk. 362–68, kus ta väidab (põhimõtteliselt), et seletada tuleb, kuidas asjad välja paistavad, ning et tema teooria selgitabki, kuidas asjad välja paistavad. Kuid Dennett, nagu ma 5. peatükis näidata püüan, kasutab siin väljendit “paistma” läbisegi psühholoogilises ja fenomenaalses mõttes. Teooria võib seletada meie soodumust moodustada stiimulite kohta mingeid kindlaid otsustusi, kuid need otsustused on hoopis midagi muud kui meid huvitav mõistatuslik *explanandum*​.

Argumente leidub veel 12. peatükis: 1) argument kvaalide inversiooni empiirilise võimalikkuse vastu (konstantseks jääva käitumise juures); kuid empiiriline võimatus on siin ühitatav mittereduktiivse positsiooniga​; 2) argument Jacksoni teadmisargumendi vastu (vaatlen seda [4. peatükis](#ORG.4.3.2)); 3) väide, et epifenomenalistlik vaade kvaalidele on naeruväärne (see küsimus tuleb vaatluse alla [4.](#ORG.4.4.3) ja 5.​ peatükis). Dennett ajab oma arutlustes looduslikku võimalikkust sageli segi loogilise võimalikkusega. Näiteks usub ta, et “kvalofiili” arvates ei ole arvutusmasinatel [​*computational machines*​] elamusi, ning tõendab seepärast pikalt-laialt, et säärastel masinatel võib olla samasugune teadvus nagu meil. Kuid see võimalus on mittereduktiivse positsiooniga väga hästi ühitatav; järgnevates peatükkides kaitsen ma ka ise samasugust seisukohta. [↑](#footnote-ref-150)
150. ​**AM 3.10**​ Viimasel ajal on Crick ja Koch hakanud teadvuse neuraalse aluse otsinguid laiendama teiste sagedustega võnkumistele, kuid asja mõte jääb samaks. Võnkumistel on see hea omadus, et neid on lihtne jälgida. [↑](#footnote-ref-151)
151. ​**AM 3.11**​ Tsitaat ajakirjast “Discover”, 1992, november, lk. 96. Ka Crick (1994, lk. 258) möönab võimalust, et kvaalidele ei tarvitse teaduslikku seletust leidudagi, kuigi tema väljendab end ettevaatlikumalt. [↑](#footnote-ref-152)
152. “whether they are considered as scientific observer or as subjects” — ei saa päris hästi aru. [↑](#footnote-ref-153)
153. ​**AM 3.12**​ Edelman (1989, lk. 168) väljendab seda täie selgusega. “Piisab, kui luua mudel, mis seletab nende eristamist, varieerumist ja tagajärgi. Teadlastena ei tarvitse me muretseda ontoloogiliste müsteeriumide pärast ega juurelda selle üle, miks midagi üldse olemas on või miks soe tundub soojana.” Ta toob selle kohta analoogia kvantväljateooriaga, mis annab meetodi energia ja mateeria seisundite eristamiseks, kuid ei ütle meile, miks mateeria üldse olemas on. See analoogia sobib väga hästi mittereduktiivse kontseptsiooniga, mida ma hakkan arendama [edasistes peatükkides](#ORG.6.1). [↑](#footnote-ref-154)
154. ​**AM 3.13**​ Tõepoolest, kui teadvus oleks füüsika suhtes loogiliselt superventne, siis ei oleks need “kollaps”-interpretatsioonid võimalikudki, sest puuduks igasugune alus anda teadvusele seadustes eriline koht. [↑](#footnote-ref-155)
155. ​**AM 4.1**​ Edelman on pannud oma (tema enda arvates materialistlikule) raamatule (1992) sarnase alapealkirja “Kuidas vaim saab alguse ajus”. [↑](#footnote-ref-156)
156. Originaalis “electromagnetic charge”, aga see ei ole vist päris korrektne. [↑](#footnote-ref-157)
157. “complex systems” — kas see on termin? [↑](#footnote-ref-158)
158. ​**AM 4.2**​ Searle’i seisukohta, nagu mina sellest aru saan, on palju loomulikum tõlgendada omadusdualismina kui materialismina, hoolimata tema enda arvamusest selles küsimuses. Minu arusaamist kinnitab tema väide, et fenomenaalsete seisundite põhjustajateks on aju seisundid, samuti sombimaailma argumendid ​**[lk.371]**​ ning väide, et “ajus on ainult neurofüsioloogilised protsessid ja teadvus, ei midagi muud”. Minu tõlgendust toetab ka Searle’i raamatu 8. peatükis esitatud argument intentsionaalsuse kohta. Searle väidab, et intentsionaalsus on reaalselt olemas (lk. 156), kuid et neurofüsioloogilised faktid ei saa konstitueerida intentsionaalseid fakte (lk. 157–58). Tema väitel on mõistatuse ainus lahendus selles, et teadvusel peab olema oma osa intentsionaalsuse koosseisus, kuna teadvus on ainus asi aju ontoloogias peale aju enda. Paistab, et see argument eeldab teadvuse küsimuses juba ette omadusdualismi.

5. peatükis oma ontoloogiat selgitades kinnitab Searle, et teadvus ei ole redutseeritav, kuid et sel asjaolul ei ole sügavaid järelmeid. Ta ütleb, et kuumus ja teised taolised fenomenid on redutseeritavad ainult sellepärast, et me defineerime neid (viisil, millest ma kõnelesin ​2. peatükis​) ümber nii, et fenomenaalne aspekt kaob, kuid et defineerida samal viisil ümber teadvust ennast, mis koosnebki ainuüksi oma subjektiivsest aspektist, on lihtlabaselt võimatu. Selles tundub tal olevat õigus. Nagu ma tõendasin [2. peatükis](#PARA.48-5), on kuumus ja teised sellised fenomenid reduktiivselt seletatavad ainult teadvuselamusi arvestusest välja jättes. Kuid edasi ütleb ta: “see näitab, et teadvuse mitteredutseeritavus on meie defineerimispraktikate pragmaatika triviaalne tagajärg” (lk. 122). Tundub, et need sõnad pööravad kogu asja pahupidi. Mainitud praktikad tulenevad ju, vastupidi, teadvuse mitteredutseeritavusest: kui me ei jätaks kuumuselamust arvestusest välja, oleks kuumust üldse võimatu redutseerida! Seega on mitteredutseeritavus meie praktikate allikas, mitte nende tagajärg. Selles kõiges on raske leida midagi, mis teadvuse mitteredutseeritavuse tähtsust kuidagi kahandaks. [↑](#footnote-ref-159)
159. ​**AM 4.3**​ Väga sarnaseid argumente näitamaks, et materialist ei saa toetuda aposterioorsele paratamatusele, on esitanud Jackson (1980, 1994), Lewis (1994) ja White (1986). [↑](#footnote-ref-160)
160. ​**AM 4.4**​ Midagi sarnast on öelnud Jackson (1980), kes väidab, et isegi kui aposterioorsete kaalutluste abil saaks kindlaks teha, et valu on füüsikaline omadus, jääks materialismile ikka muret tegema omadus “valu presenteerub”​.[fn:84presents] [↑](#footnote-ref-161)
161. “topic-neutral” — mida see siin peaks tähendama ja kuidas seda tõlkida?  — Muidu tähendab “topic-neutral”, kui seda öeldakse keelendite kohta, et neid saab kasutada igasuguses jutus, olenemata teemast. (Vt. näit. ingliskeelne Wikipedia, Type physicalism: “The concept of topic-neutral terms and expressions goes back to Gilbert Ryle, who identified such topic-neutral terms as "if", "or", "not", "because" and "and." If one were to hear these terms alone in the course of a conversation, it would be impossible to tell whether the topic under discussion concerned geology, physics, history, gardening, or selling pizza.”) [↑](#footnote-ref-162)
162. ​**AM 4.5**​ Ka Bealer (1994) soovitab siin strateegiana lähtumist füüsikalisest terminist, kuigi ta ei vii oma arutluskäiku loomuliku lõppjärelduseni. [↑](#footnote-ref-163)
163. ​**AM 4.6**​ Niisugust seisukohta on kirjanduses väljendatud harva. Enamik autoritest, kes apelleerivad materialismi kaitsmiseks aposterioorsele paratamatusele, toetuvad Kripke ideedele (näiteks Hill 1991; Lycan 1995; Tye 1995), ja peaaegu mitte keegi pole sel eesmärgil selge sõnaga põhjendanud metafüüsilist paratamatust kui paratamatuse tugevamat vormi. Tundub aga, et Bigelow ja Pargetter (1990), Byrne (1993), Levine (1993) ja Loar (1990) on sisuliselt siiski enam-vähem sellel seisukohal, kuigi nad seda otsesõnu välja ei ütle. Eraviisilistes jutuajamistes on samasuguseid vaateid kaitsnud Byrne, Levine ja Terry Horgan. [↑](#footnote-ref-164)
164. ​**AM 4.7**​ Mõnikord võib kuulda, et matemaatilised tõed olevat paratamatud metafüüsiliselt, kuid mitte kontseptuaalselt. See sõltub matemaatiliste mõistete ja kontseptuaalse paratamatuse analüüsi peensustest, kuid siiski valitseb suuresti üksmeel selles, et matemaatilised tõed on aprioorsed (ühe väikese mööndusega, millest tuleb juttu [põhiteksti järgmises jaos](#PARA.87-2)​). Peaasi on, et pole isegi mõeldavat maailma, kus matemaatilised tõed oleksid väärad. Seega ei tee need tõed võimalike maailmade ruumi mõeldavate maailmade hulgast kitsamaks.

Üheks näiteks metafüüsilisest superventsusest ilma aprioorse seoseta võiks pakkuda moraalset superventsust. Kuid argument tugeva metafüüsilise paratamatuse kasuks tundub olevat siin veelgi nõrgem kui elamuste puhul. Siin leidub alternatiive (antirealism, aprioorne seos), mis on palju ahvatlevamad kui vastavad alternatiivid teadvuselamuse puhul. Pealegi ei tundu maailm, mis oleks füüsiliselt ja vaimselt meie omaga identne, kuid ​**[lk.372]**​ moraalselt erinev, isegi mitte mõeldavana. Nii et ka siin ei sea moraalne superventsus võimalike maailmade ruumile lisakitsendusi. [↑](#footnote-ref-165)
165. ​**AM 4.8**​ Seda, kui veidrad on füsikalistlikud seisukohad, mis ei tunnista aprioorset ühenduslüli füüsika ja psühholoogia vahel, demonstreerib lihtsa mõttekäiguga (mis on ühtviisi rakendatav nii “tugeva metafüüsilise paratamatuse” kui ka “kognitiivsete kitsenduste” positsioonile) Jackson (1995):

Vaevalt leidub väga lihtsate organismide kohta fakte, mida ei saaks ​*a priori* tuletada piisavast informatsioonist nende füüsilise ehituse kohta ja füüsikalisest kirjeldusest nende vastasmõjude kohta keskkonnaga. Füüsikaline lugu amööbidest ja nende vastasmõjudest keskkonnaga ongi kogu lugu, mida amööbidest jutustada on. Kuid materialismi järgi teeb meid amööbidest erinevaks põhimõtteliselt ainult koostisosade ja nende omavaheliste suhete keerukus. On raske mõista, kuidas võiks seesugusest erinevusest sündida meie kohta olulisi fakte, mis käivad meie deduktsioonivõimele põhimõtteliselt üle jõu. … Mõelge bioloogiaklassist tuttavatele diagrammidele, mis illustreerivad evolutsioonilist arengut ainuraksetest organismidest vasakul kõrgemate ahvide ja inimeseni paremal: kus on selles jadas füsikalisti arvates täpne punkt, kus meid puudutavate oluliste faktide aprioorne tuletamine korraga võimatuks muutub? Või kui jutt juba sellele läks, siis kas materialist oskab meie arengus sügoodist inimeseni osutada täpsele punktile, kus hakkavad ilmuma olulised faktid meie kohta, mida ole enam võimalik dedutseerida meie füüsikalisest kirjeldusest? [↑](#footnote-ref-166)
166. Ei saa aru: minu arusaamist mööda tähendab Kripke paratamatus seda, et kui objektid *x* ja *y* on identsed, siis on nad ka **paratamatult** identsed; kust tulevad siia “empiirilised tegurid” (“empirical factors”)? [↑](#footnote-ref-167)
167. ​**AM 4.9**​ Analoogia psühhofüüsikalise superventsuse ja keerukate matemaatiliste tõdede vahel on isiklikes jutuajamistes teineteisest sõltumatult välja pakkunud John O’Leary-Hawthorne ja Barry Loewer. [↑](#footnote-ref-168)
168. ​**AM 4.10**​ Pange tähele, et see arutluskäik veenab analoogia puudumises matemaatiliste tõdedega isegi neid, kes pooldavad tugevat seisukoha, et on olemas nii sügavaid matemaatilisi tõdesid, et neid ei või ​*a priori* teada mitte ükski olendite klass. [↑](#footnote-ref-169)
169. ​**AM 4.11**​ Kummalisel kombel on kõnealune seisukoht üsna lähedane reduktsionistide, näiteks Dennetti vaadetele. Leiavad ju mõlemad, et nende oponentide intuitsiooni suunajaks on kognitiivne puudulikkus. Peamine erinevus on selles, et reduktsionisti arvates on meie hulgas neid, kes suudavad sellest puudulikkusest üle olla, siin vaadeldud vastuväite esitaja aga leiab, et see ei ole jõukohane kellelegi. Kuid ta ei saa välistada, et keegi teine (võib-olla koguni Dennett) on juba jõudnud kirgastumiseni. Ei suudakski ju kognitiivselt puudulik inimene hinnata lahendust, mille on andnud kirgastunu! [↑](#footnote-ref-170)
170. ​**AM 4.12**​ Loogilise võimalikkuse argumendi on kõige selgemini formuleerinud Kirk (1974). Seda võib leida ka raamatust Campbell 1970, artiklites Nagel 1974, Robinson 1976 ja mujalt. Minu versiooni iseärasuseks on peamiselt superventsuse mõiste kasutamine, mis annab ühtsustava mudeli, ning aposterioorse paratamatuse vaatlemine. Invertspektrite võimalikkuse argumendi kohta vt. ka Seager 1991. [↑](#footnote-ref-171)
171. ​**AM 4.13**​ Tundub, et mingis mõttes on mõistlik öelda, et otsustused “vesi on märg” ja “H₂O on märg” väljendavad erinevaid fakte, nagu ka otsustused “vesi on H₂O” ja “H₂O on H₂O”. Selles mõttes me individualiseerime fakte nende väljendusse kuuluvate sõnade primaarsete, mitte sekundaarsete intensioonide alusel. [↑](#footnote-ref-172)
172. ​**AM 4.14**​ Põhimõtteliselt sedasama ütleb Lockwood (1989, lk. 136–37). Tema sõnastab seda nii: kui kellelgi puudub teadmine, et need kaks on üks ja sama fakt erinevates presentatsioonimoodustes, siis tuleb see sellest, et tal puudub teadmine mingi(te)st olulis(t)est fakti(de)st suvalises presentatsioonimooduses. Midagi sarnast ütleb ka Conee (1985a). [↑](#footnote-ref-173)
173. “conditional from … to …” — kuidas seda eesti keeles öelda? Mõte on nähtavasti “konditsionaal kujul “kui metafüüsilised faktid …, siis fenomenaalsed faktid …” ”. [↑](#footnote-ref-174)
174. Tõstsin lause lõpuosa (“täpselt samasugune kriitika”) kolm lauset tagasi — minu arvates sai mõttekäigu kulg niimoodi loomulikum. [↑](#footnote-ref-175)
175. ​**AM 4.15**​ ​**[lk.373]**​ Loar oletab, et fenomenaalsed mõisted on *rekognitiivmõisted*​ ning kinnitab, et on mõistlik oodata, et rekognitiivmõiste *R* “toob sisse” sama omaduse, millest kõneleb teoreetiliselt määratletud omadus ​*P*​. Ta toob näiteks inimese, kes oskab ära tunda mõningaid Kalifornia kõrbes kasvavaid kaktusi, ilma et tal oleks nende kohta teoreetilisi teadmisi. Kuid tundub, et siin on midagi valesti: kui subjektil ei ole võimalik ​*a priori* teada, et *R* on ​*P*​, siis fikseeruvad *R* ja *P* referendid erinevalt, ja referente fikseerivad intensioonid võivad mõnes mõeldavas olukorras teineteisest lahku minna. Kui me just ei võta lisamehhanismina appi tugevat metafüüsilist paratamatust, siis vastab primaarsete intensioonide erinevus viidet fikseerivate omaduste erinevusele.

Ühes kohas oletab Loar, et rekognitiivmõisted viitavad oma objektile “otse”, ilma viiteid fikseerivate omaduste (primaarsete intensioonide) abita. Kuid seegi ei tundu õige olevat. Juba asjaolu, et teistsuguses mõeldavas olukorras võiks mõiste viidata ka millelegi muule (näiteks teistsugusele kaktuste hulgale), kõneleb meile, et siin on mängus oluline primaarne intensioon. Seega ei saa viitamine toimuda meid siin huvitavas mõttes “otse”. [↑](#footnote-ref-176)
176. “recognitional concepts”; mul ei õnnestu saada Internetist kätte Loari enda artiklit (“Phenomenal states”), kus ta selle mõiste esialgselt defineeris; üks definitsioon on olemas [ingliskeelses Wikipedias](https://en.wikipedia.org/wiki/Phenomenal_concept_strategy%22%20%5Cl%20%22Recognitional_concepts), aga see on vist poolik, või vähemalt ei saa ma sellest aru. [↑](#footnote-ref-177)
177. ​**AM 4.16**​ Lugeja, keda indeksikaalsuse probleem väga huvitab, võiks arvata, et keskpunktiga maailma keskpunkti asukohal on ontoloogiline tähendus, või et tavalised võimalikud maailmad võivad omavahel erineda indeksikaalsetelt faktidelt (näiteks kui oleks olemas midagi Nageli “objektiivse mina” taolist?). Need küsimused, nagu ka indeksikaalsuse ontoloogiline staatus üldse, on minu jaoks üsna segased. Põhitekstis lähtun ma sellest, et oponent eeldab, et indeksikaalsus ei too kaasa ontoloogilist lünka, ning tähendan, et analoogia lüngaga, mida me nägime fenomenaalsel juhul, ei pea paika isegi sel eeldusel. [↑](#footnote-ref-178)
178. Ei leia head vastet: “conscious experience remains out in the cold”. [↑](#footnote-ref-179)
179. ​**AM 4.17**​ Mõtted elamusest võivad indeksikaalidega sarnaneda veel ühel viisil: nimelt kui me osutame mõnele elamusele kui “sellele” elamusele. On mõeldav, et kui keegi viitab ühele kahest kvalitatiivselt identsest elamusest (nagu Austini (1990) “kahe toru” näites), võib küsimus, missugusele elamusele ta viitab, jääda vastuseta, isegi kui kõik “objektiivsed” faktid on teada. (Pange tähele, et jutt on konkreetsetest juhtudest [​*tokens*​], mitte tüüpidest.) Juhuks, kui materialist peaks soovima polemiseerida teadmiste argumendi vastu selle juhtumi abil, juhin tema tähelepanu sellele, et 1) sel juhul on episteemiliselt uudne fakt fenomenaalsetest faktidest sõltumatu (s.t. isegi kõigi fenomenaalsete faktide teadmine ei ütle meile, missugune neist on “see” elamus); 2) ta ei demonstreeri meile olukorda, kus on olemas sellest maailmast erinev mõeldav keskpunktita maailm, ja seepärast ei saa selle juhu alusel konstrueerida põhitekstis tooduga sarnast ontoloogilist argumenti; 3) parimal juhul aitab see uus fakt vaid lokaliseerida seda, mis toimub maailma keskpunktis, öeldes meile, milline entiteet on “see”, samamoodi nagu indeksikaalsed faktid ütlevad meile, milline entiteet olen mina.

Asja tegelik moraal on selles, et mõnel juhul on tarvis mahutada maailma keskpunkti rohkem informatsiooni: tähistada ei tule mitte ainult indiviidi ja aega (nagu “mina” ja “siin”)[fn:89here], vaid ka elamust (nagu “see”). Midagi sarnast kehtib vaieldamatult ka suunamäärsõnade​ puhul, nagu “vasak” ja “parem”. (Objektiivsete faktide teadmine maailma kohta ei pruugi öelda, kumb on vasak ja kumb on parem pool.) Episteemilised lüngad on kõigil neil juhtudel suhteliselt ohutut, indeksikaalset liiki: mitte ühelgi neist juhtudest ei leidu keskpunktita maailmu, milles alusfaktid kehtivad, lisafakt aga mitte. [↑](#footnote-ref-180)
180. Katseisik vaatab kummagi silmaga torusse ja näeb mõlemas torus punast täppi. Kui ta ütleb selle kohta, mida ta näeb, “see”, kummast täpist ta siis räägib? [↑](#footnote-ref-181)
181. “orienting demonstratives” — kas nii võib öelda? Eesti keeleõpetuses liigitatakse “vasak” ja “parem” lihtsalt *omadusssõnadeks*​, “vasakul” ja “paremal” tavaliselt *kohamäärsõnadeks* (aga see on hoopis laiem termin, mille alla mahuvad ka “allamäge”, “vastutuult”, “külas”, “lähemal”). [↑](#footnote-ref-182)
182. ​**AM 4.18**​ Sarnase vastuväite on esitanud Churchland (1985), kes ütleb, et Jacksoni argument kasutab sõna “teadma” mitmes eri tähenduses: Maryl on täielikud “propositsioonilised” või “lauselised” teadmised füüsiliste faktide kohta, kuid puudub “tutvuslik” teadmine punastest elamustest. Vastus on analoogiline. Sedavõrd kui Mary teadmised punastest elamustest kitsendavad ​**[lk.374]**​ maailma võimalike olemisviiside hulka, on nad faktiteadmised, ja argument töötab. (Mitte miski ei nõua, et faktiteadmised oleksid “lauselised”.) Niisiis peab ka Churchland, nagu Lewis ja Nemirowgi, nõustuma ebausutava väitega, et Mary teadmised punasest elamusest ei ütle talle midagi selle kohta, missugune on maailm. [↑](#footnote-ref-183)
183. ​**AM 4.19**​ Lycan (1995) toob üheksa (!) argumenti väite kasuks, et Mary teadmine sisaldab uut informatsiooni. [↑](#footnote-ref-184)
184. ​**AM 4.20**​ Teadmise argumendile on veel terve hulk vastuväiteid, mida ma pole siin arutanud, kuid minu vastuseid neile ei tohiks olla raske ette aimata. Et mainida veel vaid üht: Dretske (1995) väidab, et teadmisteks, mis Maryl puuduvad, on teadmised oma keskkonnast. Kui ta teaks rohkem punaste asjade koostisest, siis teaks ta, mida nimelt representeerivad punased elamused, ja seega (Dretske teooria järgi) ta teaks, missugused on punased elamused. Kummalisel kombel ei vaatle Dretske ilmset vastulauset: kui Mary teakski punaste objektide koostisest kõike, ei teaks ta ju ikkagi, mis tunne on näha punast! [↑](#footnote-ref-185)
185. ​**AM 4.21**​ Tänan Frank Jacksonit mõttevahetuse eest selle punkti üle. [↑](#footnote-ref-186)
186. *Identsustees* [​*identity thesis*​] (ka *identsusteooria* [​*identity theory*​]) on väide, et psühholoogilised seisundid on identsed mingite vastavate füüsilise seisunditega. Pikemalt näit. [siin](https://brainly.com/question/35156734). [↑](#footnote-ref-187)
187. Vististi [Group C nerve fibers](https://en.wikipedia.org/wiki/Group_C_nerve_fiber); kuidas oleks eesti keeles “C-fibers”, ja kuidas on “firing”? [↑](#footnote-ref-188)
188. “contingent”; [eestikeelne Wikipedia (artikkel “Kontingentne propositsioon”)](https://et.wikipedia.org/wiki/Kontingentne_propositsioon) ütleb: “Kontingentne propositsioon on propositsioon, mis on mõnes võimalikus maailmas tõene, mõnes väär.”  — Olen kohanud eesti keeles ka vastet “sattumuslik”. [↑](#footnote-ref-189)
189. FIXME “rigid” → “rigiidne”; “rigid designator” → “rigiidne designaator” — tähendab designaatorit, mis osutab igas võimalikus maailmas kui üldse millelegi, siis ühele ja samale objektile. [↑](#footnote-ref-190)
190. “kind” [↑](#footnote-ref-191)
191. Mul andis kõvasti mõtelda, enne kui sellele kohale pihta sain; püüdsin arusaadavamaks sõnastada, loodan, et sai õigesti. [↑](#footnote-ref-192)
192. FIXME Proovin tõlkida “token” “objektiks”. Vrd. [artiklit “Type-token distinction” ingliskeelses Wikipedias](https://en.wikipedia.org/wiki/Type%E2%80%93token_distinction). (Ega see variant mulle eriti ei meeldi, sest näiteks just siinkohal ei ole ju jutt objektidest, vaid pigem valu “juhtudest”, “kordadest” vms.) [↑](#footnote-ref-193)
193. “rely on a certain essentialism about various states”; vrd. [Kripke: Essentialism Recast](https://science.jrank.org/pages/9217/Essentialism-Kripke-Essentialism-Recast.html). Põhimõtteliselt peaksid see roosaks tehtud lauseosa tähendama enam-vähem sama, millest oli juttu üle-eelmises lõigus: et valutunne on valule kui seisundile olemuslikult omane, s.t. et valuseisundi juurde kuulub olemuslikult see, et ta on tuntav kui valu. [↑](#footnote-ref-194)
194. ​**AM 4.22**​ Kripke möönab ka ise (1972, 74. allmärkus), et identsuse puudumisest ei tarvitse materialismi kummutamiseks piisata​. Kuid ta märgib, et modaalsete argumentidega saab rünnata ka materialismi üldisemaid vorme. [↑](#footnote-ref-195)
195. Sellest osalausest (“the mere absence of identity may be a weak conclusion”; ka analoogilisest osalausest järgneva autorimärkuse teise lause algul) ma ei saanud päriselt aru, tõlgin oletamisi. Mul õnnestus Internetist kätte saada 1980. aasta väljaanne, aga selle 74. allmärkusest (ega teistest seal lähikonnas) ei õnnestunud mul välja lugeda midagi, mis neid sõnu selgitaks, kuigi vastavas põhitekstis on kehatusest ja identsusest juttu küll. [↑](#footnote-ref-196)
196. ​**AM 4.23**​ Sarnasel viisil võiksid argumendid, kus figureerivad kehatud aistingud, tõestada, et vaimsed omadused ei ole identsed füüsiliste omadustega, kuna füüsikalised omadused saavad realiseeruda ainult füüsilistes objektides; sellise argumendi esitab Dealer (1994). Kuid ka säärasest mitteidentsusest ei tarvitse materialismi kummutamiseks piisata​: ta on endiselt ühitatav loogilise superventsusega ja seega ka materialismiga. Tõepoolest, sarnast mitteidentsusargumenti saab kohaldada peaaegu igale kõrgtaseme omadusele. [↑](#footnote-ref-197)
197. ​**AM 4.24**​ Oma hoolikas analüüsis mainib Boyd (1980, lk. 98), et erinevalt kehatute aistingute võimalikkusest viib sombide võimalikkus loogiliselt välja materialismi vääruseni. Seepärast esitab ta sombide võimalikkuse vastu eraldi argumendi, mis jääb aga skemaatiliseks ja ebaveenvaks. Boyd vaatleb analoogiat arvutiga, mis arvutab mingit kindlat funktsiooni, ja väidab, et 1) kuigi meile võib intuitiivselt tunduda, et arvuti võiks täpselt samasuguste ühendustega oma elektrooniliste komponentide vahel teha ka hoopis midagi muud, on see tegelikkuses ikkagi võimatu​, ja et 2) sombide näiline võimalikkus on midagi analoogilist. Analoogia ei pea aga paika. Olukord arvutiga on analoogne (väga nõrga) “näilise võimalikkusega”, et minu füüsiline koopia ei õpiks seda, mida õpin mina, või ei eristaks seda, mida eristan mina. Selles analoogias pole midagi, mis seletaks, miks on teadvuselamusteta koopia näiline võimalikkus palju veenvam. [↑](#footnote-ref-198)
198. See koht jäi mulle originaalis arusaamatuks; õnneks on Boydi originaaltekst Google Booksis olemas, ja Google soostus just 98. lehekülge ka [näitama​](https://books.google.ee/books?id=5Nyh9goAKgsC&pg=PA98). Ma vist ei oskagi sõnastada seda osalauset nii, et tänapäeva lugeja sellest aru saaks. Asi on selles, et Boydi tolleaegne ettekujutus arvutiprogrammidest ja arvuti tööpõhimõttest mõjub tänapäeval anakronistlikuna: ta kujutab arvutiprogrammi nähtavasti ette füüsiliste ühenduste konfiguratsioonina, kus lülituspaneelil on hulk mingeid kindlaid punkte juhtmete abil kindlal viisil ühendatud. Kui need ühendused on korraldatud nii, et arvuti tõstaks arve ruutu (nagu Boydi näites), siis ei ole juhtmepuntrale pilku peale heites tõepoolest ilmne, miks peaks arvuti tegema just seda, mida ta teeb. Mõttekäik jääb arusaadavaks ka siis, kui ühenduste (“circuits”) asemele panna aukude read perfokaartidel või -lindil; ei tea, kuidas mujal maailmas, aga Eestis oli see programmeerimisviis 1980. aastal minu mäletamist mööda veel tavaline. Enamik siinse raamatu lugejaid aga ei tea sellest kõigest vist midagi, mistõttu neil pole kuidagi võimalik kogu sellest mõttekäigust aru saada — kui just mitte panna “ühenduste” asemele “programm”, aga siis tuleks põhimõtteliselt välja midagi seesugust, nagu “võiks intuitiivselt tunduda, et programm arvude ruututõstmiseks võiks teha ka midagi muud kui arve ruutu tõsta” — mis kõlaks juba lausa tobedalt ja viiks lugeja mõte parimal juhul ehk programmeerimisveale, aga kindlasti mitte sellele, mida mõtles Boyd. [↑](#footnote-ref-199)
199. [Siin](#PARA.84-11) ja [siin](#PARA.84-12). [↑](#footnote-ref-200)
200. ​**AM 4.25**​ Kripke märkus, et materialist peab näitama, et “asjad, mida me suudame ette kujutada, ei ole tegelikult asjad, mida me suudame ette kujutada” (Kripke 1971, eelviimane lõik), viib samuti mõtte nõrgale tõlgendusele. Kripke ei välista võimalust, et näilist võimalikkust saab olematuks seletada mingil muul viisil, mis on täiesti erinev standardsetest juhtumitest nagu vesi​/​H₂O, kuid ütleb, et “selleks on tarvis sügavamat ja peenemat mõttekäiku, kui mina aimatagi oskan, ja peenemat kui kõik need, mida on iial toodud kogu materialistlikus kirjanduses, niipalju kui mina seda lugenud olen.” [↑](#footnote-ref-201)
201. ​**AM 4.26**​ Kuigi kogu Kripke argumenti peetakse sageli tema rigiidsete designaatorite teooria rakenduseks, oleks midagi taolist võidud käiku lasta põhimõtteliselt juba kümme aastat varem, kui kogu seda teooriat veel ei olnudki. Esimeste identsusteoreetikute käest oleks võidud küsida, miks ei tulene füüsikalistest faktidest aju seisundite kohta paratamatult valu olemasolu, kuigi füüsikalistest faktidest H₂O kohta tuleneb, et H₂O on paratamatult vesi (või vesijas ollus). [↑](#footnote-ref-202)
202. ​**AM 4.27**​ ​**[lk.375]**​ Horgan (1987) ja ka Byrne (1993) kõnelevad selles kontekstis “metafüüsilisest” superventsusest. Ent kui minu seisukoht, et (maailmade) metafüüsiline võimalikkus ja loogiline võimalikkus ühtivad, on õige, siis tuleneb sellest edasi ka loogiline superventsus. [↑](#footnote-ref-203)
203. [Wikipedia](https://en.wikipedia.org/wiki/Interactionism_%28philosophy_of_mind%29): “Interactionism or interactionist dualism is the theory in the philosophy of mind which holds that matter and mind are two distinct and independent substances that exert causal effects on one another.” [↑](#footnote-ref-204)
204. Originaalis “strong Humean account”, aga siin kirjeldatud arusaama kohta ei sobi “strong” ju sugugi! Võtsin eeskuju venelaselt, kellel on “радикальное” — mingis mõttes võibki ju Hume’i arusaama radikaalseks (s.t. radikaalselt vähenõudlikuks) nimetada. [↑](#footnote-ref-205)
205. Mis neil kahel (“nomic regularity” (Hume’il) ja “uniform regularity”) vahet on? Mul ei õnnestunud Interneti-otsinguga selles selgust saada. “Nomic regularity” kohta [leidsin](https://www.degreesofclarity.com/writing/nomic/) järgmise seletuse: “Nomic regularity is the general principle that the functioning of the world will tend to follow common rules or principles; we will generally find that things behave “regularly,” and hence that we can form certain expectations about them.” (Mille poolest see ei ole “uniform”, vaid midagi “enamat”?) (“Ühetaoline” on Hume’i “uniform” vastena juba [käibel](https://www.ester.ee/record%3Db5406806%2Aest).)  — Kui noomiline regulaarsus ongi midagi enamat kui ühetaoline regulaarsus, siis tuleb “isegi kui” (mille valisin venelase eeskujul, temal on “даже если”) asemele panna “kuigi”. [↑](#footnote-ref-206)
206. Originaalis “P₁”, mis on ilmne näpuviga. [↑](#footnote-ref-207)
207. Originaalis on siin “unknowable causal relation”, jutumärkides, mis jätab mulje, nagu oleksid need sõnad otsene tsitaat Hume’ilt. Aga Interneti-otsinguga ei õnnestunud mul Hume’il niisugust sõnaühendit leida. Seepärast jätsin jutumärgid ära.  — Mõte ise on Hume’i põhjuslikkusekäsitluses küll täiesti läbiv, vt. näit. [Uurimus inimese arust](https://www.ester.ee/record%3Db5406806%2Aest), VII, 2. osa, lk. 81: “… kui objektide vahel on üldse mingit suhet, mida meil on tähtis täielikult teada, siis on selleks kindlasti põhjuse ja tagajärje suhe. … Ometi on meie ideed selle kohta nii ebatäiuslikud, et võimatu on anda mingit õiget põhjuse definitsiooni peale sellise, mis lähtub millestki põhjuslikkusele välisest ja võõrast.” [↑](#footnote-ref-208)
208. Oletus — ma ei ole kindel, et ma originaali segaselt väljendatud mõttest õigesti aru saan: “tells us nothing about what all this causation relates”. Järgnev lause näib minu oletust kinnitavat.  — Ühendeid, kus on niiviisi koos “causality” ja “relate”, tuleb siin jaos allpool veel; olen säärastes ühendites tõlkinud “relate” “kandmiseks” “kandjaks” jne.; nii tundub justkui loogiline. [↑](#footnote-ref-209)
209. ​**AM 4.28**​ Sarnaseid mõtteid leidub ka allikates Black­burn 1990, Feigl 1958, Lock­wood 1989, Max­well 1978 ja Robinson 1982. [↑](#footnote-ref-210)
210. ​**AM 4.29**​ Ühe ettekujutuse maailmast kui puhtast põhjuslikkusvoost on esitanud Shoemaker (1980), kes väidab, et kõik omadused on “jõud” [​*powers*​] ning et pole olemas mingeid muid omadusi, mis oleksid nende jõudude aluseks. Shoemakeri argument selle seisukoha kaitseks on suures osas verifikatsionistlik ja selle pildiga seostuvaid raskusi ta otseselt ei vaatle.

Edasi väidab Shoemaker, et kuna mingi omaduse juurde kuuluvad jõud on sellele omadusele olemuslikult omased, siis peavad loodusseadused olema *a posteriori* paratamatud. (Sarnaselt arutleb Swoyer (1982), ja sellise järeldusega flirdib ka Kripke (1980).) Aposterioorse paratamatuse kahemõõtmeline analüüs osutab, et selles väites peab olema midagi valesti, või et ta ei ole vähemasti nii laiahaardeline, kui pealtnäha tundub. Parimal juhul võib asi olla nii, et maailmade kohta, kus valitsevad teistsugused seadused, ei ole õige öelda, et nad sisaldavad (näiteks) elektrone; kuid see ei luba niisuguseid maailmu veel välistada. Ka ei tundu usutav, et elektronsuse konstitueerivad viimseni kõik elektronidega seostuvad jõud. Tõenäolisem on, et selleks, et mingit entiteeti saaks pidada elektroniks, on tarvilik ainult mingi osa neist jõududest, ning et võimalikud on ka pisut kontranoomilised maailmad, mis sisaldavad samuti elektrone. Shoemaker väidab, et konstitutiivseid jõude ei ole mitte mingil viisil võimalik eristada mittekonstitutiivsetest. Kuid kahemõõtmeline analüüs osutab, et nende erinevus tuleneb otseselt elektronsuse mõistest.

Siin tuleks eristada mitut küsimust: 1) Kas osutused füüsikalistele omadustele on relatsioonilised​? (Shoemaker, Chalmers: jah.) 2) Kas füüsikalised omadused on (sekundaarse intensiooni alusel) identsed relatsiooniliste omadustega? (S: jah; C: tõenäoliselt jah, kuid semantilised intuitsioonid võivad olla erinevad​.) 3) Kas füüsikalisele omadusele on olemuslikult omased kõik tema noomilised suhted? (S: jah; C: ei.) 4) Kas nende relatsiooniliste omaduste aluseks on mingid seesmised omadused? (S: ei; C: jah). [↑](#footnote-ref-211)
211. “semantic intuitions may differ” — mida see tähendab? Ei saa aru! [↑](#footnote-ref-212)
212. ​**AM 4.30**​ Seda seisukohta on viimastel aastatel toetanud Lockwood (1989) ja Maxwell (1978), kes mõlemad pidasid seda üheks identsuseteooria mitteortodoksseks versiooniks. Seda vaadet on mulle halastamatult propageerinud Gregg Rosenberg. [↑](#footnote-ref-213)
213. Millele õieti? [↑](#footnote-ref-214)
214. “second”, aga “teise” ei oleks üheselt arusaadav. (Loodan, et link on õige.) [↑](#footnote-ref-215)
215. ​**AM 4.31**​ Kuigi on esitatud ka vastupidist seisukohta: vt. Lahav ja Shanks 1992. [↑](#footnote-ref-216)
216. ​**AM 4.32**​ Lewis (1990) jõuab sarnasele järeldusele hoopis teisel teel. [↑](#footnote-ref-217)
217. “change the subject”; venelasel on hästi: “подменяют предмет обсуждения”, kas eesti keeles ka midagi analoogilist leidub? [↑](#footnote-ref-218)
218. Lisasin omalt poolt juurde viited ülaltootud loetelule — lugejal ei tarvitse kõigi nende numbrite tähendused meeles seista. [↑](#footnote-ref-219)
219. Lisasin “mingi kujuteldav” ja (allpool) “praegune” omalt poolt, sest ilma nende lisandusteta ei tundu autori jutt mulle arusaadav olevat. Aga nad kajastavad minu isiklikku arusaamist; võib-olla on autoril siin mingi muu mõte, mida ma ei taba. [↑](#footnote-ref-220)
220. “change the subject”; venelasel on hästi: “подменяют предмет обсуждения”, kas eesti keeles ka midagi analoogilist leidub? [↑](#footnote-ref-221)
221. Ma ei oska kuidagi teisiti aru saada originaali (vähemalt minu jaoks) hämarast sõnastusest: “semantic determinacy of attributions of consciousness”. [↑](#footnote-ref-222)
222. Täpselt sellist kohta ei leidnud, autor mõtleb vist [seda](#PARA.96-3). [↑](#footnote-ref-223)
223. “identifies”; tundub, et see on liiga kitsas, võib-olla on autori mõte “peab fenomenaalseid omadusi füüsikaliste entiteetide seesmisteks omadusteks”? [↑](#footnote-ref-224)
224. ​**AM 4.33**​ Teadvuse kohta ettekandeid pidades ja muudel puhkudel olen ma korraldanud selle kohta mõnikord küsitluse. Tulemused on järjekindlalt 2:1 või 3:1 selle kasuks, et teadvuse juures on veel midagi, mis vajab seletamist. Loomulikult ei ole hääletamine parim viis filosoofiliste küsimuste lahendamiseks, ent sedaliiki küsimustes, mida ei ole võimalik lahendada argumentidega, on algsete intuitsioonide vahekorral siiski mõnevõrra kaalu. [↑](#footnote-ref-225)
225. ​**AM 4.34**​ *Bioloogiline materialism.* Ühe levinud seisukohta järgi (Hill 1991; Searle 1992) on teadvus kindlalt bioloogiline nähtus. Materialism on tõene, kuid teadvuseta süsteemid, mille funktsionaalne organisatsioon oleks samasugune nagu teadvusega süsteemidel, on loogiliselt ja arvatavasti isegi empiiriliselt võimalikud. Ent niipea kui me oleme tunnistanud minu teadvuseta funktsionaalse isomorfi loogilist võimalikkust, peame vältimatult tunnistama, et loogiliselt on võimalik ka minu bioloogiline teadvuseta isomorf, sest neurofüsioloogial ei ole teadvuselamustega kontseptuaalselt rohkem ühist kui ränil. Seepärast on seda seisukohta arvatavasti kõige parem vaadelda kui üht omadusdualismi erikuju, kus teadvus on lisafakt, mis jääb füüsikaliste faktide valdkonna piiridest välja. Kui sellega mitte nõustuda, siis saaks seda positsiooni — kuna ta rõhutab biokeemia ja teadvuse vahelist seost — ​**[lk.376]**​ hädakorral rühmitada kokku vaadetega, mis apelleerivad tugevale metafüüsilisele paratamatusele, misläbi talle kanduvad üle ka kõik nende vaadete hädad.

(Searle (1992) tunnistab sombide loogilist võimalikkust ja on tegelikult seisukohal, et seos mikrofüüsika ja teadvuselamuse vahel on vaid põhjuslik, ja seega on teda ehk kõige õigem pidada omadusdualistiks. (Hill (1991) püüab sombide võimalikkust vältida apelleerimisega rigiidsetele designaatoritele, kuid see strateegia, nagu nägime​, ei tööta). [↑](#footnote-ref-226)
226. ​**AM 4.35**​ *Füsikalistlik funktsionalism.* Selle populaarse seisukoha (nt. Shoemaker 1982) järgi on teadvuselamuste olemasolu funktsionaalne omadus, kuid iga konkreetne teadvuselamus (näiteks punase värvuse aisting) on neurofüsioloogiline omadus. Selle vaate järgi on invertspektritega funktsionaalsed isomorfid loogiliselt ja võib-olla ka empiiriliselt võimalikud, kuid funktsionaalsed isomorfid, kellel teadvus täielikult puuduks, seda ei ole. Ent jällegi, kui me tunnistame invertspektriga funktsionaalse isomorfi loogilist võimalikkust, siis peame tunnistama ka invertspektriga füüsilise isomorfi loogilist võimalikkust, sest neurofüsioloogial ei ole konkreetsete elamustega kontspetuaalselt rohkem ühist kui ränikiipidel. Seega tundub taas, et füüsikalised faktid ei determineeri kõiki fakte, ja siit tuleneb mingis vormis omadusdualism. Ja ka siin tuleb füsikalismi positsioonil püsimiseks omaks võtta see kahtlase väärtusega mõiste — tugev metafüüsiline paratamatus.

Seda seisukohta kujutatakse sageli kui aposterioorsete fenomenaalsete omaduste samastamist neurofüsioloogiliste omadustega. Sellisena tuleb tal seista silmitsi samade tuttavate raskustega, mis on omased nii säärasele aposterioorsele identifitseerimisele (mis on siin primaarne intensioon?) kui ka ülaltoodud argumendile. Nagu märgib White (1986), kes kritiseerib seda vaadet sarnases vaimus, oleks selle seisukoha pooldajatel parem jääda üldlevinud funktsionalismi juurde. [↑](#footnote-ref-227)
227. ​**AM 4.36**​ ​*Psühhofunktsionalism.*​ See vaade identifitseerib mentaalsed omadused aposterioorselt funktsionaalsete omadustega nende rollide alusel küpses empiirilises psühholoogias​ (vt. Block 1980). Kui kohaldada seda seisukohta fenomenaalsetele omadustele, tuleb välja, et fenomenaalsetel mõistetel ühtivad sekundaarsed intensioonid funktsionaalsete mõistete sekundaarsete intensioonidega, hoolimata primaarsete intensioonide erinevusest. Selle seisukoha raskusi on kõige parem analüüsida [2. jao](#ORG.4.2) meetodil, s.t. koondades tähelepanu primaarsetele intensioonidele. Kui fenomenaalsetel mõistetel on primaarne intensioon ise funktsionaalne, siis viib see seisukoht ikkagi välja mingitlaadi analüütilise funktsionalismini; kui aga mitte, siis viib tähelepanu koondamine omadusele, millele see intensioon vastab, meid vältimatult mingitlaadi dualismini. Nii või teisiti ei aita see seisukoht materialismi päästmisele kuidagi kaasa.

Selle vaate poolehoidjad on sageli mööda vaadanud mõistete rollist viidete fikseerimisel primaarsete intensioonide kaudu. Isegi kui meil oleks teaduslik teooria, milles figureerib teoreetilise terminina “uskumus”, tuleb ikkagi jutustada kontseptuaalne lugu, mis ütleb, miks kvalifitseerub sedaliiki seisund uskumuseks, mitte aga sooviks või millekski hoopis muuks. Tõenäoliselt oleks see viidet fikseeriv intensioon ise funktsionaalne, osutades seisundile või millelegi selletaolisele, mille roll teoorias sarnaneks kõige rohkem sarnaneb uskumuse rollile, kusjuures uskumuse rolli määratleb meie senine mõiste. Missugune ka ei oleks fenomenaalsete omaduste primaarsete intensioonide olemus, raskustest me ei pääse. Kõnelda siin sekundaarsetest intensioonidest tähendab lihtsalt raskused kalevi alla lükata.

Lisaks käib psühhofunktsionalismiga kaasas veel üks raskus: ta viib välja omamoodi šovinismini, kuna siin antakse inimese psühholoogiale eriline kaal otsustamisel, mida lugeda (näiteks) uskumuseks. Selle kohta on esitanud suurepärase kriitilika Shoemaker 1981, kuigi vt. ka Clarki vastust (1986). Tundub olevat usutavam, et enamikul mentaalsetest mõistetest primaarne ja sekundaarne ​**[lk.377]**​ intensioon ühtivad. Vastasel juhul satume olukordadesse, kus meie ja meie kolleegid Teisikmaal mõtlevad “uskumuse” all erinevaid asju, hoolimata sellest, et meie algsed mõisted oleksid identsed. [↑](#footnote-ref-228)
228. ​**AM 4.37**​ *Anomaalne monism.* Selle vaate järgi on iga vaimne seisund individuaalselt identne mingi füüsilise seisundiga, kuid mingeid rangeid psühhofüüsikaseadusi ei ole. Anomaal-monistliku positsiooni pakkus Davidson (1970) välja intentsionaalsete, mitte fenomenaalsete seisundite käsitlemise eesmärgil, kuid seda vaadet võib siiski pidada meie teemasse puutuvaks kahel põhjusel: esiteks annab ta füsikalismi kasuks aprioorse argumendi, mis toetub ainult füüsiliste ja mentaalsete seisundite põhjuslikele (olgugi ühesuunalistele) vastasmõjudele, ja teiseks eitab ta psühhofüüsika seadusi, mida eeldab minu positsioon.

Nägemaks, et Davidsoni argumendid minu positsiooni ei ohusta, paneme tähele, et neis ei ole midagi, mis eitaks üksikjuhulisi​ seadusi kujul “kui mingi süsteemi maksimaalselt spetsiifiline füüsiline seisund on P, siis tema (maksimaalselt spetsiifiline) vaimne seisund on M”. Tõepoolest, Davidson tunnistab vaimsete omaduste superventsust füüsikaliste omaduste suhtes, millest näib loomuliku tõlgenduse korral tulenevat selliste seaduste olemasolu (selleteemalist arutelu vt. Kim 1985). Davidsoni vaadete kõige heatahtlikum tõlgendus oleks, et ta ei eita mitte üksikjuhuseadusi, vaid huvitavamaid *tüübi*​seadusi, mis ühendavad vaimseid seisundeid füüsikaliste seisundite avarate tüüpidega, näiteks nendega, mida võib kohata rahvalikus psühholoogias. Tundub küll, et tema argumentidest, mis toetuvad vaimse elu holismile, midagi enamat tuletada ei saa. Ja sellisel juhul ei ähvarda looduslikku superventsust miski. Sellest järeldub ka, et ei saa paika pidada argument individuaalse identsuse kasuks, mis toetub eeldusele, et pole olemas rangeid seadusi, mis kehtestaksid põhjusliku seose füüsiliste ja vaimsete omaduste vahel (mistõttu selle asemel nõutakse identsust). Kuid isegi rangetest üksikjuhuseadustest piisab, põhistamaks seesugust füüsiliste ja fenomenaalsete seisundite vahelist seost, mida pooldan mina. Seega ei ähvarda miski ka dualismi. [↑](#footnote-ref-229)
229. “pointwise laws” — saan vist enam-vähem aru, mida autor silmas peab, aga kuidas seda eesti keeles öelda? [↑](#footnote-ref-230)
230. ​**AM 4.38**​ *Representatsionalism.*​ Viimasel ajal on populaarsust võitnud seisukoht (nt. Dretske 1995; Harman 1990; Lycan 1996; Tye 1995), mis ütleb, et fenomenaalsed omadused on vaid *representatsioonilised* omadused: kollased kvaalid on lihtsalt tajuseisundid, mis representeerivad kollaseid asju, või veel midagi samas vaimus. Loomulikult sõltub selle seisukoha tõlgendus omakorda sellest, kuidas tema esitaja mõistab representatsioonilisi omadusi. Enamasti kaasneb selle väitega representatsiooni reduktiivne (tavaliselt funktsionaalne või teleofunktsionaalne) käsitlus; sellisel juhul osutub ta reduktiivse funktsionalismi erikujuks ja peab seisma silmitsi selle tavaliste raskustega. Representatsiooni mittereduktiivse käsitluse korral saaks neid raskusi vältida (kuigi asemele võivad tulla teised), ent see viiks elamuse mittereduktiivse käsitluseni.

Mõnede representatsionalistlike seisukohtade näiline usutavus tuleb võib-olla sellest, et nad mõistavad terminit “representatsioon” läbisegi inflatiivselt ja deflatiivselt; selles viimases mõttes on representatsiooni mõiste puhtalt funktsionaalne (või teleofunktsionaalne), esimeses mõttes aga mitte. Seos fenomenoloogia ja representatsiooni vahel on usutav esimeses tõlgenduses, representatsiooni reduktsioon teises. Veel üks võimalus on võtta representatsiooniliste ja fenomenaalsete seisundite seostamiseks appi tugev metafüüsiline paratamatus ja leppida raskustega, mida viimane kaasa toob. (Meieaegsetest representatsionalistidest näivad Dretske [1995] ja Harman [1990] pooldavat tugevalt reduktiivset A-tüüpi seisukohta, Lycan [1996] ja Tye [1995] aga B-tüüpi seisukohta, mis toetub aposterioorsele paratamatusele).

Veel võib vastuseks representatsionalismile ka tähendada, et peaaegu kõik on nõus, et mitte kõik representatsioonilised seisundid ei ole fenomenaalsed seisundid (need aga, kes nii ei arva, vaatlevad mõlemaid peaaegu kindlasti mittereduktsionistlikult), ja et seega võib küsida: mis ​**[lk.378]**​ teeb mõnedest representatsioonilistest seisunditest fenomenaalsed seisundid? Just see lisakriteerium kannab teadvuse representatsionalistlikus teoorias tegelikku koormust. Sageli võetakse selleks kriteeriumiks nõue, et tsentraalsetel protsessidel peab olema representatsioonilisele seisundile mingi sobivat laadi ligipääs; sellisel juhul saame selgelt reduktiivfunktsionalistliku vaate, mis seisab silmitsi kõigi niisuguste vaadete tavaliste raskustega (miks peaks just see asjaolu tegema representatsioonilise seisundi fenomenaalseks?). Teine võimalus on liigitada relevantseteks just need representatsioonilised seisundid, mis on fenomenaalsed, kuid see tee viib meid otsejoones tagasi omadusdualismi juurde. [↑](#footnote-ref-231)
231. ​**AM 4.39**​ *Teadvus kui kõrgemat järku mõte.* Sarnaselt võib käsitleda ka mõtet, et üks või teine seisund on teadvuslik juhul, kui ta on mingi kõrgemat järku mõtte objekt (vt. nt. Rosenthal 1996 jt.). Kui kombineerida seda mõtet reduktiivse ettekujutusega kõrgemat järku mõtetest, tuleb sisuliselt välja reduktiivfunktsionalistlik vaade koos oma tavaliste hädadega. Kui seda mitte teha, viib see positsioon välja mittereduktiivse (B- või C-tüüpi) vaateni elamusele ja on seega ühildatav omadusdualismiga, mida pooldan mina, kuigi ta võib põrkuda teistele raskustele (mis tulevad arutuse alla [6. peatükis](#PARA.126-4)). [↑](#footnote-ref-232)
232. ​**AM 4.40**​ *Reduktiivne teleofunktsionalism.* Siinkohal tasub mainida Dretske (1995) seisukohta, et üheks elamuse olemasolu kriteeriumiks on ka teleoloogiline komponent: selleks, et süsteemil oleksid elamused, ei piisa sellest, et ta talitleks teataval viisil — lisaks sellele pidi ta tegema relevantsetest protsessidest nende ajaloo käigus ka sobiva valiku. Väidetavalt on see seisukoht suuteline saama üle mõningatest standardse funktsionalismi raskustest, lubades (ja seletades) näiteks funktsionaalselt identsete sombide võimalikkust: need on lihtsalt ebaõige ajalooga süsteemid. Kuid sellel vaateviisil on peamistest raskustest samuti omad versioonid. Näiteks ei tundu loogiliselt vähem võimalikuna, et funktsionaalselt identsel ja sobiva ajalooga süsteemil võib teadvus ikkagi puududa; ning teadmised organisatsiooni kohta ei anna teadmisi elamuse kohta ka siis, kui lisada neile teadmised süsteemi ajaloo kohta. Võiks öelda, et see vaade püüab hoiduda reduktiivse funktsionalismi raskustest ebaõigel viisil. Kokkuvõttes läheneb kõnealune positsioon oma ilmelt pigem A-tüüpi reduktiivfunktsionalistlikule vaatele kui vaatele, mis võtab teadvust tõsiselt. [↑](#footnote-ref-233)
233. ​**AM 4.41**​ *Emergentne põhjuslikkus.*​ Paljud on soovinud loobuda teadvuse reduktiivsest käsitlusest, loobumata teadvuse kesksest põhjuslikust rollist. Sel otstarbel on sageli toodud argumente *emergentse põhjuslikkuse* kasuks — väidetud, et keerukates füüsikalistes süsteemides tekib uusi põhjuslikkuse liike. Näiteks Sperry (1969, 1992) on väitnud, et teadvus on keerukate süsteemide emergentne omadus, mis mängib ise põhjuslikku rolli; sarnasel seisukohal olid briti emergentistid, nagu näiteks Alexander (1920, arutelu selle üle võib leida artiklist McLaughlin 1992). Sarnase oletuse on esitanud Sellars (1981; vt. ka Meehl ja Sellars (1956)): tema järgi võib teatud süsteemides, näiteks protoplasmas või süsteemides, mis on substraadiks mõistusega olenditele, tulla mängu uusi füüsikalise põhjuslikkuse seadusi. (Ta andis sellele seisukohale nimeks “füsikalism₁”, et eristada seda “füsikalism₂-st”, mis ütleb, et anorgaanilises mateerias leitavad füüsikalised põhiprintsiibid kehtivad universaalselt). Neid vaateid ei tohiks ajada segi keerukate süsteemide teoorias kohatava “süütu” vaatega emergentsele põhjuslikkusele, kus madaltaseme seadustest kerkib vastasmõjuefektide kaudu esile kvalitatiivselt uudne käitumine. Radikaalsema vaate järgi astuvad siin mängu uued fundamentaalsed printsiibid, mis ei tulene madaltaseme seadustest.

Sellel seisukohal on kaks viga. Esiteks pole sääraste emergentsete põhjuslikkusprintsiipide kohta mingeid tõendeid. Niipalju kui meie teame, tuleneb igasugune põhjuslikkus madaltasemelisest füüsikalisest põhjuslikkusest, ja põhjuslikkus “ülalt alla” ei sekku kunagi madaltaseme protsessidesse. Teiseks, mis võib-olla veelgi olulisem: hoolikamal analüüsil selgub, et ka see vaateviis jätab ​**[lk.379]**​ teadvuse sama üleliigseks nagu enne. See saab ilmseks, kui panna tähele, et loos emergentsest põhjuslikkusest pole midagi, mis nõuaks viitamist fenomenaalsetele omadustele. Kogu põhjusliku jutu võib pajatada seoste kaudu füüsikaliste omaduste konfiguratsioonide vahel. Endiselt jääb võimalikuks maailm, mis on füüsikaliselt meie maailmaga identne, ent kus teadvus täielikult puudub. Sellest järeldub, et fenomenaalsed omadused on põhjuslikult toimivate konfiguratsioonidega parimal juhul vaid korrelatsioonis. Kui selle positsiooni puhul on siiski mingil moel võimalik omistada fenomenaalsetele omadustele põhjuslik toimivus, siis saab sama manöövrit rakendada ka minu kontseptsiooni järgides. Tegelikult ongi niisugust vaadet kõige parem vaadelda kui minu vaate erikuju, mille järgi teadvus on füüsika suhtes superventne kontingentse noomilise ühenduse kaudu. Erinevus on selles, et siin figureerivad uued, emergentsed füüsikalised põhjusseadused, mis aga teevad asja vaid keerulisemaks, muutmata midagi sisuliselt. [↑](#footnote-ref-234)
234. ​**AM 4.42**​ *Müsterianism.* Need, kellele teadvuse reduktiivsed käsitlused ei meeldi, on sageli seisukohal, et teadvus jääbki võib-olla igavesti mõistatuseks. Sellist seisukohta on väljendanud Nagel (1974) ja Jackson (1982) ning edasi arendanud McGinn (1991). Selle vaate kohaselt võib teadvus käia üle meie mõistuse samavõrd, nagu astronoomia käib üle meritigude mõistuse.

See seisukoht on võib-olla ahvatlev, kuid siiski ennatlik. Öelda, et teadvusele ei ole võimalik anda reduktiivset seletust, ei tähenda öelda, et teadvusele ei ole üldse võimalik seletust anda. Eelkõige võib printsiipide selgitamine, millede tõttu teadvus on füüsika suhtes looduslikult superventne, viia selgusttoova teadvuseteooriani, isegi kui jääda mittereduktiivse vaate juurde.

McGinn (1989) väidab, et ajuseisundite ja teadvusseisundite vahel on paratamatu seos (vastasel juhul oleks teadvuse ilmumine ime), kuid et me ei saa kunagi teada, milles see seos seisneb. Tema mõttekäikudest paistab, et ta peab silmas loogilist või metafüüsilist paratamatust; kuid argument lubab kõnelda parimal juhul vaid looduslikust paratamatusest. Kindlasti ei ole kontingentne noomiline seos teadvuse ja füüsilise maailma vahel suurem ime kui iga teinegi kontingentne seadus, ja tõepoolest tundub selline seos palju vähem salapärasena kui loogiliselt või metafüüsiliselt paratamatu seos, milleni meie arusaamine ei küüni. Ja pole selge, miks me ei võiks tuletada selliseid seadusi oma teadmiste põhjal füüsikalisi protsesse ja elamusi ühendavate reeglipärade kohta. Järgmistes peatükkides püüan ma asjaomaste seaduste iseloomustamiseks üht-teist öelda. Nii näeme, et mittereduktiivne vaade teadvusele ei too tingimata kaasa pessimismi. [↑](#footnote-ref-235)
235. Nii ongi: “the argument”, ilma täpsustuseta. Ma ei saa aru, mis argumenti autor silmas peab. ([Originaalartikkel](https://philosophy.tamucc.edu/graphics/berkich/texts/mcginn-can-we-solve-the-mind-body-problem.pdf) on Internetis kättesaadav.) [↑](#footnote-ref-236)
236. Kus? [Siin](#ORG.2.1.4)? [↑](#footnote-ref-237)
237. Autor teeb Newtonile ülekohut. Newton oli nii vinge tegija, et ta ei vajanud oma teooriate ebakohtade nägemiseks oponente. Kaugmõju idee ei meeldinud talle mitte üks raas ja ta on öelnud selle kohta võib-olla sandimaid sõnu kui ükski tema oponent. Näiteks ütleb ta kirjas [Richard Bentley](https://en.wikipedia.org/wiki/Richard_Bentley)​le (1692): “That gravity should be innate, inherent, and essential to matter, so that one body may act upon another at a distance through a vacuum, without the mediation of anything else, by and through which their action and force may be conveyed from one to another, is to me so great an absurdity that I believe no man who has in philosophical matters a competent faculty of thinking can ever fall into it.” ([Wikiquote](https://en.wikiquote.org/wiki/Isaac_Newton)) Ta sai väga hästi aru, et ta ei oska seletada, miks tema gravitatsioonivalem (pöördruutude seadus) töötab — et selle taga peab olema “veel midagi”. Füüsika pidi (kusjuures just Newtoni rajatud vundamendilt) terve paar-kolm sajandit edasi arenema, enne kui see “veel midagi” hakkas selgemat kuju omandama. [↑](#footnote-ref-238)
238. “in a deflationary way”; vt. [Stanford Encyclopedia of Philosophy](https://plato.stanford.edu/entries/truth-deflationary/). [↑](#footnote-ref-239)
239. Ongi autoril nii selles kui järgnevas punktis “sensation”, kuigi loomulikum oleks “experience”. [↑](#footnote-ref-240)
240. ​**AM 5.1**​ Elitzur ütleb olevat saanud selle idee Penrose’ilt (1987). [↑](#footnote-ref-241)
241. ​**AM 5.2**​ Ma ei vaatle siinkohal religioosseid elamusi. Seletust nõuavad siin arvatavasti sügavad vaimulikud elamused ja hardus. [↑](#footnote-ref-242)
242. viljaringid? [↑](#footnote-ref-243)
243. ​**AM 5.3**​ Ma ei ole kindel, kas täpselt seda liini võib kirjanduses kohata, kuid sarnaseid argumente leidub. Näiteks Foss (1989) ütleb vastuseks Jacksoni (1982) teadmisargumendile, et kui Mary ka teabki kõike, mida värvusi nägev subjekt erinevate värvuste kohta ütleks, ja isegi kõike, mida ta öelda võiks, ei piisa sellest veel kaugeltki, et teada kõike, mida on tarvis teada. [↑](#footnote-ref-244)
244. Originaalis õieti “I am left defending” — mida peab siin tähendama see “left”, seda ma ei oska arvata ilma laiemat konteksti tundmata, aga seda artiklit saab Internetist kätte ainult kalli raha eest. [↑](#footnote-ref-245)
245. ​**AM 5.4**​ Muidugi ei annaks ka see tegelikult midagi. [↑](#footnote-ref-246)
246. ​**AM 5.5**​ Analoogiliselt arutleb Dretske (1995), väites, et tema teooria seletab, kuidas asjad näivad, ja seega seletab seda, mis vajab seletamist. Temagi kasutab sõna “näima” ​**[lk.380]**​ psühholoogilises ja fenomenaalses tähenduses läbisegi. Just selle mitmetimõistetavuse tõttu on “näima” üldse halb sõna teadvuseteooria *explanandum*​ite iseloomustamiseks. Huvitaval kombel kasutavad seda peaaegu eranditult just reduktiivsete käsitluste pooldajad. [↑](#footnote-ref-247)
247. Ei saa sellest kohast aru: “when talking about a zombie’s beliefs and judgments, **a** deflationary notion is being stipulated” — umbmäärane artikkel ütleks justkui, et jutt on mingist kindlast mõistest (või mida see “notion” siin tähendab), aga missugusest? [↑](#footnote-ref-248)
248. *Reliabilismi* (vt. [artikkel ingliskeelses Wikipedias](https://en.wikipedia.org/wiki/Reliabilism)​) on meil *teadmine*​, et ​*P*​, juhul, kui 1) *P* on tõene, 2) me usume, et ​*P*​, ja 3) me oleme jõudnud selle uskumuseni usaldusväärse protsessi teel. (Kas “reliabilism” on hea vaste?  — On olemas ka [sellepealkirjaline artikkel eestikeelses Wikipedias](https://et.wikipedia.org/wiki/Reliabilism).) [↑](#footnote-ref-249)
249. Ei saa loogikast aru: kui usaldusväärne seos on olemas, kuidas siis ei saa välistada võimalust, et seda ei ole? [↑](#footnote-ref-250)
250. Nagu ka ka sombide otsesed andmed! [↑](#footnote-ref-251)
251. ​**AM 5.6**​ Muidugi on kirjanduses otsese teadmise ideed ja “meeleandmete” ideed mitmeti rünnatud. Kuid minu arvates ei ole mitte ükski ründaja suutnud kummutada mõtet, et kinnitus meie uskumustele elamuse kohta tulenebki elamuse olemasolust. On toodud küll kaalukaid põhjendusi, miks tuleks eitada mitmesuguseid tugevamaid väiteid — näiteks et igasugune teadmine tuleneb elamuse teadmisest, või et me tajume maailma meeleandmete tajumise läbi, või et elamuse olemasolu paigutab ta automaatselt mingi mõiste alla. Kuid mina ei väida siin midagi niisugust.

Sellars (1956) kritiseerib veenvalt ideed, et “meelelise sisu *s* aistimine” toob kaasa mittetuletuslikke​[fn:113noninferential] teadmisi sisu *s* kohta. Ta märgib, et igasugune teadmine on alati kontseptuaalne seisund, mistõttu kõnealune teadmine ei ole tõenäoliselt elementaarteadmine; elamus tundub aga olevat elementaarsem. Läbielamine on vaieldamatult mittekontseptuaalne seisund, ja meie tutvus elamusega on mittekontseptuaalne suhe (kuigi see punkt oleneb sellest, kuidas määratleda kontseptuaalsust). Jääb seega küsimus, kuidas saab mittekontseptuaalne seisund olla tõenduseks kontseptuaalse seisundi kohta. See on raske küsimus, kuid niiisugune, mis ei puuduta ainult mittereduktsionistlikku positsiooni teadvuse küsimuses. Sama küsimus kerkib ka juba tavalise tajulise teadmise puhul maailmast, kus isegi reduktsionist peab tunnistama, et uskumuste põhjendused toetuvad osalt ka mittekontseptuaalsele alusele — kui ta ei taha võtta omaks alternatiive, mis tooksid kaasa vist veelgi suuremaid raskusi. Ma arvan, et niisugust põhjenduste alust oleks võimalik leida, kuid see oleks väga mahukas omaette projekt. Siinkohal märgin lihtsalt, et ükski raskus, mis ses vallas meie ette kerkib, ei ole eriomane mittereduktsionistlikule teadvusekäsitlusele. [↑](#footnote-ref-252)
252. ​**AM 5.7**​ Selle üldise liini on valinud Hill (1991) oma üksikasjalikus vastuses “ersatsvalu” võimalikkusele toetuvatele skeptilistele argumentidele elamuse kohta, eriti Shoemakeri (1975a) argumendile. Kuigi Hill ei poolda omadusdualismi, vaid B-tüüpi bioloogilist materialismi, esitab ta mitu punkti sarnases vaimus sellega, kuidas minagi neid küsimusi käsitlen. Nende punktide seas on elamust puudutavate skeptiliste argumentide võrdlus välismaailma puudutavate skeptiliste argumentidega ning argumendid “äratuntavusnõude” vastu, mille kohaselt usul ​*P*​-sse on alust ainult juhul, kui igas olukorras, kus puuduvad tõendid ​*P* kasuks, on võimalik ära tunda, et need tõendid puuduvad. [↑](#footnote-ref-253)
253. Segane sõnastus originaalis (ja venelasel samuti), tõlgin aimamisi. [↑](#footnote-ref-254)
254. Autoril lihtsalt “memory”, aga ta peab vist siiski silmas just elamuste mäletamist, mitte mälu üldiselt. [↑](#footnote-ref-255)
255. Originaalis “or”, aga see ei anna mõtet kokku — keeleliselt annaks tõlkida ka “ehk”, aga see ei tundu sisuliselt sobivat, need kaks ei ole üks ja sama asi. [↑](#footnote-ref-256)
256. “impossible to refer to causally irrelevant experiences” — puht keeleliselt kahemõtteline: kas “causally irrelevant” on kirjeldav või piiritlev täiend? Otsustasin, et kirjeldav. [↑](#footnote-ref-257)
257. “nothing in its world” — tundub, et see on autori näpuviga ja et siin oleks pidanud olema “nothing in the world”. (Nii on nähtavasti mõelnud ka venelane: temal on “ничего из существующего в мире”, mitte “ничего из существующего в его мире”.) [↑](#footnote-ref-258)
258. Oletus; originaalis on siinkohal lihtsalt “it”. [↑](#footnote-ref-259)
259. ​**AM 5.8**​ Selle idee ütles ühe diskussiooni käigus välja John O’Leary-Hawthorne. [↑](#footnote-ref-260)
260. ​**AM 5.9**​ Võiks arvata, et vastupidine mõtteviis — mille järgi sombi uskumus “mul on teadvus” on tõene, kuna tema mõiste osutab mingile funktsionaalsele omadusele — võiks omadusdualistidele nende epistemoloogilistes raskustes kasulik olla, sest nüüd ei pea enam paika järeldus, et minu uskumused on legitiimsed seal, kus sombi uskumused seda ei ole. Ent ükskõik, kumma seisukoha me siin valiksime, jääb sombile ikka mingi hulk ekslikke uskumusi — näiteks et tal on omadusi, mis väljuvad tema füüsiliste ja funktsionaalsete omaduste piirest — ja seega kerkiks probleem, kuidas põhjendada minu vastavaid uskumusi, teisel kujul uuesti. [↑](#footnote-ref-261)
261. “public-language”; vt. [Wikipedia, Private language argument](https://en.wikipedia.org/wiki/Private%20language%20argument). Sõnaühendil “public language” on juba ammust ajast (mitu sajandit) ka tavapärane tähendus: “keel, mida kõneldakse avalikus elus ja mida mõistab enamik kõnelejaid”. [↑](#footnote-ref-262)
262. ​**AM 5.10**​ Näiteks on seda meie omavahelises jutuajamises öelnud Bill Lycan. Siin on võimalik minna ka tagasi mõiste “elamus, mille põhjustavad (enamikus meist) tavaliselt punased asjad” juurde — ehkki siin on relativismist pääsemise hinnaks oht süstemaatiliselt eksida iseenda elamuste kategooriates. [↑](#footnote-ref-263)
263. ​**AM 5.11**​ Seesugusest relativismist on vaba väline värvusemõiste, näiteks ​**[lk.381]**​ punasuse mõiste, kui pidada selle all silmas objektide, mitte elamuste omadust. Esimeses lähenduses võib öelda, et osutus punastele asjadele fikseerib punased asjad kui asjad, mis tüüpiliselt põhjustavad (enamikus meist) samasuguseid värvuselamusi nagu mõned etalonnäidised. Selline osutus on “avalikum” kahes suhtes: kuna ta toetub avalikele etalonnäidistele ja kuna ta kõneleb elamustest, mis on kogukonnale ühised. Seega viitab ka invertspektriga indiviid “punastest esemetest” kõneldes samadele esemetele nagu mina, kuigi “punased elamused” osutab tema puhul millelegi muule.

Veel võib tekkida küsimus: kuidas jääb indiviididega, kelle värvusruum liigendub teisiti, näiteks sellega, kelle silmis porgandid on sama värvi nagu roosid ja tomatid. Kõige loomulikum on arvatavasti öelda, et kuna mõistes “punane” sisaldub kogukondlik element, siis on niisuguse inimese ütlus “porgand on punane” väär; kuid samas on võib-olla võimalik pidada seda mingis vähem kogukondlikus mõttes (tähenduses “punane minu silmis”) ka tõeseks. Kuid isegi sellisel juhul ei jää relativismile väga palju ruumi, kuna iga indiviidi jaoks jääb mõiste ikkagi ühendusse väliste etalonidega. Isegi selles relativistlikus tähenduses peab mõiste “punane” osutama tublile hulgale punastele esemetele, ja nii on see ka teiste mõistetega.

Võimalik oleks ka jätta väliste värvusmõistete kirjeldustest välja üldse igasugune sõltuvus elamusest, määratledes neid hoopis otsustuste kaudu: nii osutuksid punasteks esemed, mida tavaliselt otsustatakse olevat vastavate etalonnäidistega sama värvi. Sellise valiku eelis on, et nüüd on ka sombidel võimalik kõnelda rohelistest objektidest tõeselt, mis ongi võib-olla mõistlik. Tundub ju, et selleks, et värvusmõisted töötaksid, ei ole tarvis, et indiviidide elamused oleksid sarnased — piisab sellest, et otsustused on sarnased. [↑](#footnote-ref-264)
264. ​**AM 5.12**​ Siinne vahetegemine “punase elamuse” kvalitatiivse mõiste ja relatsioonilise mõiste vahel on väga sarnane Nida-Rümelini (1995) vahetegemisega “fenomenaalse” ja “mittefenomenaalse” arusaama vahel säärastest uskumuseomistustest, nagu näiteks “Marianna usub, et taevas paistab Peetrile sinisena”. Nida-Rümelini “fenomenaalse” käsituse järgi puudutab selle lausega väljendatud omistus vastavat kvalitatiivset mõistet, mittefenomenaalse käsituse järgi aga relatsioonilist. (Nida-Rümelin paistab oma näidetes mõtlevat relatsioonilise mõiste all midagi [10. märkuses](#fn-5-10) mainitud kogukondliku mõiste taolist.) [↑](#footnote-ref-265)
265. Vt. [Wikipedia:Private\_language\_argument#The\_beetle-in-a-box](https://en.wikipedia.org/wiki/Private_language_argument%22%20%5Cl%20%22The_beetle-in-a-box). [↑](#footnote-ref-266)
266. ​**AM 5.13**​ Muidugi võtsin ma selle mõiste tähiseks siin meelega ​*R*​, et see meenutaks lugejale sümbolit​ *E*​ Wittgensteini privaatkeele argumendis​. Ma ei proovigi seda argumenti siinkohal analüüsida; ülesande teeks eriti raskeks asjaolu, et puudub üldtunnustatud tõlgendus selle kohta, milles see argument õigupoolest seisneb. Piisab, kui öelda, et kõik selle argumendi versioonid, mida ma näinud olen, kas toetuvad väga kahtlastele eeldustele või on ühevõrra rakendatavad nii igapäevasetele mõistetele kui ka privaatsetele elamuslikele mõistetele, või siis on neil mõlemad puudused korraga. [↑](#footnote-ref-267)
267. Vt.[Wikipedia:Private\_\_language\_argument](https://en.wikipedia.org/wiki/Private_language_argument). [↑](#footnote-ref-268)
268. ​**AM 5.15**​ Mingil määral peegeldab see mõttekäik üht Shoemakeri (1975) arutlust: kui invertspekter on võimalik, siis ei ole võimalik funktsionaalselt defineerida ei kvalitatiivseid seisundeid ega kvalitatiivseid uskumusi. Kuid Shoemaker väljendab oma mõtet rigiidsete designaatorite kaudu, mille referent fikseerub relatsiooniliselt, koondades seega tähelepanu tegelikult sekundaarsetele intensioonidele. [↑](#footnote-ref-269)
269. ​**AM 5.16**​ ​**[lk.382]**​ See tähelepanek on paralleelne Nida-Rümelini tähelepanekuga (1995), et fenomenaalsete ja mittefenomenaalsete uskumuste erinevus ei ole tavapärase *de re*​/​*de dicto* erinevuse erijuht. [↑](#footnote-ref-270)
270. ​**AM 5.17**​ Pange tähele, et “see” ei ole siin lihtne indeksikaalne “see” — mille primaarne intensioon on alati sama, olenemata sellest, kas elamus on ​*R* või ​*S*​, ja mille puhul oleks sisutu ja triviaalne öelda elamuse kohta, et see on “seda” liiki elamus —, vaid sisukas “see”, mille primaarne intensioon osutab igas keskpunktiga maailmas *S*​-elamustele. [↑](#footnote-ref-271)
271. ​**AM 5.18**​ Vastuses Shoemakeri (1975a) epistemoloogilisele argumendile toetubki Conee (1985b) selletaolisele konstitutiivsele seosele kvaalide ja kvalitatiivsete uskumuste vahel. [↑](#footnote-ref-272)
272. “it will systematize this connection …” — ei saa aru! See “systematize” peaks ju olema sama sõna paarik järgmisest lausest, aga järgmises lauses on jutt tähelepanekute (mitmus!) süstematiseerimisest; kuidas saaks samas mõttes aga süstematiseerida seost (ainsus!), seda mina välja mõelda ei suutnud.  — Ja see “connection” ajab jutu siin üldse minu jaoks segaseks. [↑](#footnote-ref-273)
273. ​**AM 6.1**​ Paneme tähele, et selleks, et neist printsiipidest saaks tuletada psühhofüüsika seadusi, tuleb mõista teist järku otsustusi, näiteks nagu “mul on praegu punane elamus”, [5. peatüki viimases jaos](#PARA.117-8) käsitletud relatsioonilises mõttes: umbes nagu “mul on praegu sedaliiki elamus, mida tavaliselt põhjustavad punased objektid”. Erinevalt “punase elamuse” mõiste seesmistest kvalitatiivsetest elementidest kajastuvad tema relatsioonilised elemendid füüsikalistes protsessides: vastav uskumus, milles figureerib kvalitatiivne mõiste “punane elamus”, ei ole füüsika suhtes loogiliselt superventne ning seega ei saa sääraste uskumuste õigsusest tuletada ainsatki psühhofüüsikaseadust. Ma ei pea seda punkti liiga tähtsaks, kuna minu käsitluse tulipunktis on esimest järku otsustused, mille puhul neid küsimusi ei teki. [↑](#footnote-ref-274)
274. “blindness denial” [↑](#footnote-ref-275)
275. Shakespeare, “Macbeth”, 2. vaatus, 1. pilt. [↑](#footnote-ref-276)
276. Tõlkisin “occurrent” venelase eeskujul parema puudumisel “aktuaalseks”, aga see on halb hädalahendus, sest sõna “aktuaalne” on siinses tõlkes juba vaste ka sõnale “actual”; kas eesti keeles leidub paremat vastet? [↑](#footnote-ref-277)
277. ​**AM 6.2**​ Sisuka analüüsi aktuaalsete mõtetega seonduvast fenomenoloogiast on esitanud Siewert (1994). [↑](#footnote-ref-278)
278. ​**AM 6.3**​ Vrd. ka Nageli (1974) tähelepanekut, et “tajude struktuursed omadused võivad olla objektiivselt hõlpsamini kirjeldatavad, ehkki midagi jääb ka kirjeldamata”. [↑](#footnote-ref-279)
279. ​**AM 6.4**​ Sellega väga sarnane on Jackendoffi arvutusliku piisavuse hüpotees: “Iga fenomenoloogilist erinevust põhjustab​/​toetab​/​projitseerib vastav arvutuslik erinevus” (Jackendoff 1987, lk. 24). [↑](#footnote-ref-280)
280. ​**AM 6.5**​ Samas vaimus on pimenägemise üle arutlenud ka Tye (1993), Block (1995) ja eriti Dennett (1991). [↑](#footnote-ref-281)
281. vt. [PDF](https://web.archive.org/web/20120926104736id_/http%3A//www.nyu.edu/gsas/dept/philo/faculty/block/papers/1995_Function.pdf) [↑](#footnote-ref-282)
282. Sõnademäng: “higher order” — “kõrgemat järku”, “order” — “tellima”. [↑](#footnote-ref-283)
283. ​**AM 6.6**​ Arutelu ohtudest ajada teadvuselamus segi teadvusse jõudvate elamustega, samuti suurepärase kriitilise käsitluse teooriate kohta, mis opereerivad kõrgemat järku mõtetega, võib leida allikast Siewert 1994. Paralleelset kriitikat reduktsionistlikult positsioonilt vt. ka Dretske 1995. [↑](#footnote-ref-284)
284. ​**AM 6.7**​ Esimest ja teist järku käsitluste erinevus on paralleelne erinevusega kahe funktsionaalse teadvusemõiste C1 ja C2 vahel (Nelkin 1989). [↑](#footnote-ref-285)
285. ​**AM 6.8**​ Sellist oletust vaatleb Carruthers (1992), kes väidab, et kvalitatiivse tunde looduslikult tarvilikuks ja piisavaks tingimuseks on kättesaadavus refleksiivsele mõtlemisele. (Tundub siiski, et Carruthers peab silmas mingit tugevamat liiki kättesaadavust, kuivõrd Armstrongi autopiloodirežiimis veoautojuht tema kriteeriumile ei vasta.) Kuivõrd ta väidab selge sõnaga, et see seos on ainult looduslikult paratamatu, tundub, et tema idee on mittereduktiivne, kuigi Carruthers nimetab oma vaadet ka füsikalistlikuks. Alvin Goldman pakkus ühes jutuajamises välja sarnase käsitluse, mõeldes seda vähem reduktiivse kontseptsioonina kui teadvusseisundite iseloomustusena meile tuttavates süsteemides. [↑](#footnote-ref-286)
286. Armstrong 1981: “If you have driven for a very long distance without a break, you may have had experience of a curious state of automatism, which can occur in these conditions. One can suddenly ‘come to’ and realize that one has driven for long distances without being aware of what one was doing, or, indeed, without being aware of anything. One has kept the car on the road, used the brake and the clutch perhaps, yet all without any awareness of what one was doing.” [↑](#footnote-ref-287)
287. ​**AM 6.9**​ Sellele vahetegemisele esimest järku registreeringute ja esimest järku otsustuste vahel leidub paralleel Dretskel (1995), kes eristab kognitiivsete seisundite ​**[lk.383]**​ *fenomenaalseid* ja *doksastilisi* vorme. Viimased vastavad süsteemi arvamusele olukorrast, esimene olukorra representatsioonile süsteemis. Loomulikult on Dretske arusaam minu omast oluliselt erinev: oma vaadete põhituumalt on Dretske reduktiivne funktsionalist (täpsemini, reduktiivne teleofunktsionalist), kes samastab elamusi esimest järku registreeringutega nende (teleo)funktsionaalses määratluses. Mina seda samastamist juba tuttavatel põhjustel ei poolda, kuid sellegipoolest on usutav, et elamuste ja esimest järku registreeringute vahel on mingi vastavus. Minu mudelis on lihtsalt üks eristus juures: ma tunnistan tervelt kolme erinevat seisundite liiki — otsustusi, registreeringuid ja fenomenaalseid seisundeid —, ning ütlen, et registreeringud on fenomenaalsete seisunditega korrelatsioonis, kuid ei ühti nendega. Dretskel aga on seisundeid ainult kaht liiki — otsustused ja fenomenaalsed seisundid —, ja ta ei tee isegi kontseptuaalselt vahet fenomenaalsete seisundite ja neile vastavate registreeringute vahel. [↑](#footnote-ref-288)
288. ​**AM 6.10**​ Teadvust konstitueerivate esimest järku registratsioonide representatsioonilise sisu iseloomu ning nendega paralleelsete elamuste sisu kohta kerkib tohutu hulk huvitavaid küsimusi. Kuna sisu puudutavad küsimused ei ole minu käsitluses kesksed, siis puudutan neid üsna lühidalt, kuigi need kuuluvad kõige sügavamate ja peenemate küsimuste hulka elamuse kohta ja väärivad palju üksikasjalikumat käsitlust teisal.

Nii teadlikkuse kui ka elamuse sisulise külje üks peamisi omadusi on, et see sisu on enamasti mittekontseptuaalne — ta ei eelda agenti, kes valdaks mõisteid, millega seda sisu kirjeldada. Näiteks tundub olevat usutav, et lihtsatel süsteemidel — näiteks koeral või hiirel — võib olla väga peeni värvuselamusi, millele vastavatel representatsioonidel kognitiivsüsteemis on värvuseristused sama peened; kuid nende värvusmõistete süsteem on väga lihtne. Ka inimeste teadvus- ja teadlikkusseisundite sisu võib muusikat tajudes olla nii külluslik, et ületab kaugelt subjekti muusikaliste mõistete varamu.

(Mittekontseptuaalse sisu kohta vt. Crane 1992, Cussins 1990, Evans 1982, Pea­cocke 1992. Kirjanduses näib valitsevat üksmeelne seisukoht, et elamuse sisu on mittekontseptuaalne. Üheks erandiks on McDowell (1994), kes asjaolust, et me võime elamusi tagantjärele iseloomustada, nimetades neid “selleks värvitooniks” vms., järeldab, et igasugune elamussisu on kontseptuaalne. Ei ole aga ilmne, et elamused sellist võimet eeldavad: näiteks tundub kaunis usutav, et mõni muusikaelamuse peen aspekt (nt. väikesed muutused helistikus) võib jääda mõnele subjektile nii tabamatuks, et ta ei suuda seda ei kontseptualiseerida ega tagantjärele identifitseerida. Veel võib näiteks tuua loomade elamused. McDowell paistab olevat väga nõus järeldusega, et loomadel ei ole elamusi, kuid mõne teise arvates on *modus tollens*​ siin võib-olla vähemalt sama veenev kui *modus ponens*​.​**(\*)** Ent isegi kui McDowelli seisukohaga tuleks nõustuda, võiks midagi sellele eristusele vastavat siiski tunnistada, näiteks mõistelise sisu kahe eri astme kujul.)

Muidugi võib võib mõistete ja teadvuse vahel olla ka mingi põhjuslik seos; mõisted võivad küllalt sageli mõjutada elamuse iseloomu. Kuid tundub, et säärased kontseptuaalsed ressursid ei ole teadvuselamuseks tarvilikud. Sama võib öelda ka teadlikkuse kohta, kuivõrd see on paralleelne teadvusega. Sisu, mida representeerivad teadvuselamustele vastavad esimest järku registreeringud, näiteks nägemistajus, samavõrd rikkalikke kontseptuaalseid ressursse ei nõua. Seega on elamuse ja teadvuse sisu üldiselt primitiivsem kui otsustuste sisu, mida on loomulik pidada kontseptuaalseks.

​**[lk.384]**​ Üks huvitavamaid küsimusi sisu kohta on, kas representatsiooniline sisu on seesmiselt omane elamusele endale või tuleneb ta mingil moel elamust kandvast kognitiivsest seisundist. Viimane seisukoht võib tunduda ahvatlevana, kuid ei ole vist päris õige: näiteks tundub, et minu visuaalne elamus representeerib praegusel hetkel maailma, kus minu ees paikneb suur ruudukujuline objekt, ning selleks pole tarvis midagi muud, kui et elamus oleks niisugune, nagu ta ongi. Elamuse representatsiooniline sisu oleks sarnane isegi hüpoteetilisel kehatul vaimul, kui tal oleks sarnane elamus. Siewert (1994) argumenteerib veenvalt, et elamusele on olemuslikult omane anda meile informatsiooni maailma seisundi kohta: näiteks visuaalse elamuse puhul on võimalik hinnata tema täpsust (kas ta kujutab maailma õigesti või vääriti?), ja seda ainuüksi tänu oma olemusele visuaalse elamusena. On võib-olla mõistlik öelda, et representatsiooniline sisu on elamusele olemuslikult omane.

Mõnel võib olla kiusatus asuda vastupidisele seisukohale ja kinnitada, et tõeline sisu on ainult elamusel ning et elamust kandva esimest järku registreeringu sisu sõltub ise elamuse sisust. Selles võib midagi olla, kuid ka see ei ole päriselt rahuldav; me võiksime ju tahta öelda, mingis mõttes representeerivad ka sombi esimest järku registreeringud maailma. Meil on ju kahtlemata sisukaid seisundeid ilma vastavate elamusteta ja on raske uskuda, et kõik meie sisud sõltuksid nii või teisiti elamuste sisudest. Võiksime asuda ka vahepealsele seisukohale: (1) esialgne sisu oli mingis mõttes see, mis leidus elamuses, kuid (2) meil on evolutsiooni käigus kujunenud sisu atribueerimiseks kognitiivsetele seisunditele süsteem, mis mingis osas toetub koherentsusele nendega seostuvate elamuste sisuga, ning (3) olles välja kujunenud, muutus see süsteem autonoomseks, nii et me võime kõnelda kognitiivsete seisundite sisust isegi elamuse puududes. See tähendaks, et nii elamustel kui ka neile vastavatel registreeringutel võib olla omaenda sisu, ja sellisel juhul poleks tarvis imestada kummalise liigdeterminatsiooni üle, mille varal üks ja sama sisu konstitueerub kaks korda järgemööda. Need on üpris peened küsimused, mida tasuks võib-olla põhjalikult analüüsida.

Teine huvitav küsimus on, kas meid huvitav sisu on “avarat” või “kitsast” liiki, s.t. kas ta sõltub ka keskkonnas leiduvatest objektidest või ainult seesmistest protsessidest. Kuivõrd elamusele on olemuslikult omane, et tal on sisu, ja kuivõrd elamus on subjekti organisatsiooni suhtes superventne, peab see sisu olema kitsast liiki. (Elamustel võib endistviisi olla ka avarat liiki sisu, kuid seda sisu ei fikseeri siis elamus üksi.) Mõnikord on kinnitatud, et ainus tõeline representatsiooniline sisu on avar sisu, kuid minu arvates on võimalik loomulikul viisil mõista ka kitsast representatsioonilist sisu (vt. Chalmers 1994c). Sellist arusaama oleks võimalik edasi arendada kirjelduseks elamuse kitsast, mittekontseptuaalsest sisust kui sisust, mis seab kitsendusi niisuguste keskpunktiga maailmade hulgale, mida on võimalik pidada subjekti aktuaalseks maailmaks.

Veel üks küsimus: kas representatsiooniline sisu on kõigil elamustel? Paljudel või enamikul neist arvatavasti küll: on ilmne, et tajuelamustele on enamasti seesmiselt omane anda informatsiooni maailma kohta. Kuid on ka juhtumeid, kus asi pole nii selge: kuidas on lugu näiteks orgasmi või iivelduse või mõnede tunde-elamustega (vt. Block 1995 ja Tye 1995)? Kuid isegi nendel juhtudel võib vastavates elamustes leida mingisugusegi representatsioonilise sisu, kuna elamuse sisu ütleb sageli midagi asukoha (“siin”, “seal all”) või kvaliteedi (“hea”, “halb”) kohta. Pole kerge leida elamusi, millel representatsiooniline sisu täielikult puuduks; samas pole ka ilmne, et niisuguseid elamusi ei saa olla.

​**[lk.385]**​ Mõned filosoofid on oletanud, et fenomenaalsed omadused ongi vaid representatsioonilised omadused ning et seega ammenduvadki elamused oma representatsioonilise sisuga (nt. Dretske 1995; Harman 1990; Lycan 1996; Tye 1992). Enamasti väljendatakse seda oletust koos reduktiivse vaatega representatsioonilisele sisule, mis taandub seega mingil kujul reduktiivsele funktsionalismile ja on väheveenev põhjustel, mida me juba teame. Kuid seda mõtet saaks esitada ka ühenduses mittereduktiivse vaatega representatsioonilisele sisule, näiteks niisugusega, mille järgi tõelist representatsiooni leidubki ainult elamuses. See sobiks paremini kokku tõsise huviga teadvuse vastu, kuid ka nii jääkisid raskused püsima. Näiteks tundub, et representatsiooniline sisu võiks olla samasugune ka invertspektriga funktsionaalsel isomorfil, ja säärasel juhul oleks fenomenoloogias siiski midagi rohkemat kui representatsiooniline sisu. Ka eelmises lõigus loetletud juhtumid kipuvad viima mõttele, et kui elamustel ongi alati representatsiooniline sisu, on neil siiski ka omadusi, mis sellesse sisusse ei mahu. Seega pole kindel, et see oletus isegi oma mittereduktiivsel kujul paika peab. [↑](#footnote-ref-289)
289. **(\*)** Mul kulus tükk aega, enne kui ma selle lause lõpuosast (loodetavasti) aru sain: kus on siin *modus ponens*​, kus *modus tollens*​? Viimaks taipasin: *modus ponens*​ peab vististi olema kolm lauset tagasi öeldud “all experiential content is conceptual”. Kui nii, siis tundub, et autor on, püüdes oma mõtet vigurlikult sõnastada, öelnud midagi, millest võib aru saada kahte moodi, olenevalt sellest, mida mõelda siin *modus tollens*​i all: kas “no conceptual content is experiential” või “no experiential content is conceptual”; ma ei näe, mis alusel peaks lugeja eelistama üht teisele, aga samas need kaks välistavad teineteist, ja kuna esimene variant väljendab täpselt sama, mida ütles juba *modus ponens*​, siis väljendab autori mõtet vististi teine. Ma ei näe tarvidust nii keerulise sõnastuse järele: autor oleks võinud öelda lihtsalt “one might find the contrary position at least as compelling”, ilma loogikamõistatusteta.  — Ma siiski ei julge tõlkes vastavat muudatust teha, äkki leiab mõni targem minu siinses mõttekäigus vea.

(McDowelli raamatu “Mind and World” [1996. aasta väljaanne](https://teoriadelconocimientocontemporanea.files.wordpress.com/2014/06/mcdowell-mind-and-world.pdf) on Internetis vabalt saadaval.) [↑](#footnote-ref-290)
290. ​**AM 6.11**​ Isiklikus jutuajamises ja seni veel ilmumata töödes. [↑](#footnote-ref-291)
291. ​**AM 6.12**​ Printsiipe, mida ma siin mainisin, on mõnikord otsesõnu formuleeritud ühe osana vaimu käsitlevate empiiriliste uurimistööde metodoloogiast. Pole üllatav, et nii on see kõige sagedamini olnud selles peavoolupsühholoogia valdkonnas, millel on teadvuselamusega kõige rohkem kokkupuudet, nimelt psühhofüüsikas. Selle valdkonna ülesandeks on sageli peetud meie aistingute omaduste seostamist vastavate füüsiliste stiimulite omadustega. Tüüpilisteks saavutusteks on siin näiteks Weberi–Fechneri seadus ja Stevensi eksponentsiaalseadus (Stevens 1975), mis annavad kumbki ühe võimaliku seose stiimuli intensiivsuse ja vastava aistingu intensiivsuse vahel. Kuigi mõnikord arvatakse, et psühhofüüsika peamiseks *explanandum*​iks on kolmanda isiku andmed, näiteks subjektiivsed ütlused, tundub olevat vaieldamatu, et kesksete fenomenide hulka, mida see uurimisala selgitada püüab, kuuluvad ka esimese isiku elamuse omadused — näiteks elamused, mis tekivad teatavate optiliste illusioonide puhul.

(Horst [1995] näitab veenvalt, et mitmesuguste fenomenide, näiteks illusioonide puhul on selles valdkonnas lähteandmeteks sageli esimese isiku elamused. Teadlased panevad konverentsidel suurt rõhku võimalusele mitmesuguseid efekte “oma silmaga näha”. Samuti on võimalik, et poleemikal selle üle, kumb aistingute intensiivsuse seadus on õige, kas Fechneri või Stevensi oma [vt. Stevens 1975], on mõtet ainult eeldusel, et nad püüavad mõõta sama asja, nimelt fenomenaalset elamust; vastasel korral pole nende vahel mingit vastuolu, kuna nad mõõdavad erinevaid funktsionaalseid fenomene.)

Psühhofüüsikas on aeg-ajalt diskuteeritud küsimuse üle, kuidas võivad empiirilised vaatlused aidata kaasa subjektiivsete aistingute seletamisele. Mõned uurijad on püüdnud formuleerida selgeid printsiipe, millele see töö toetub, nimetades neid “psühhofüüsikalisteks sidehüpoteesideks” [​*psychophysical linking hypotheses*​] (Brindley 1960) või “üldisteks sidepropositsioonideks” [​*general linking propositions*​] (Teller 1984). Heaks näiteks on Mülleri (1896; tsitaat võetud allikast Boring 1942, lk. 89) “psühhofüüsikalise vastavuse aksioomid” [​*Axioms of Psychophysical Correspondence*​]:

	1. Iga teadvusseisundi aluseks on materiaalne protsess, nn. psühhofüüsikaline protsess, mille kulgemisega kaasneb vastav teadvusseisund.
	2. Igale võrdsusele, sarnasusele või erinevusele aistingute koosseisus … vastab võrdsus, sarnasus või erinevus psühhofüüsikalise protsessi koosseisus ja ümberpöördult. Veel enam, suuremale või väiksemale ​**[lk.386]**​ sarnasusele aistingus vastab suurem või väiksem sarnasus ka psühhofüüsikalises protsessis ja ümberpöördult.
	3. Kui aistingu muutumisel on mingi kindel suund või kui erinevused aistingujadade vahel on samasuunalised, siis on sarnane suund ka vastaval psühhofüüsikalisel protsessil või vastavate psühhofüüsikaliste protsesside erinevustel. Veel enam, kui aistingu muutumise suundi on mingi arv ​*n*​, siis on ka aistingut kandval psühhofüüsikalisel potsessil ​*n* võimalikku muutumissuunda ja ümberpöördult.On selge, et need printsiibid on lähedases suguluses struktuurse koherentsuse printsiibiga. Kõik ülaltoodud väited on mõningatest sõnastuslikest erinevustest hoolimata koherentsusprintsiibi otsesed järelmid, mis üheskoos annavad annavad talle suure osa tema jõust. Seega näeme veel kord, et struktuurse koherentsuse printsiip ja selle variandid mängivad keskset rolli elamuste mitmesuguste omaduste seletamisel empiiriliste uurimistulemuste abiga.

Pole üllatav, et diskussioonidel selliste printsiipide staatuse üle psühhofüüsikas on paralleele analoogiliste diskussioonidega vaimufilosoofias (vt. nt. Brindley 1960; Marks 1978; D. Teller 1984, 1990). Mõned autorid on pidanud neid empiirilisteks hüpoteesideks, kuid nähtavasti ei ole neid empiiriliselt ei kinnitatud ega ümber lükatud, vähemasti mitte testidega kolmanda isiku positioonilt. Teised uurijad, eriti niisugused, kelle mõttelaad kaldub operatsionalismi poole, on pidanud neid definitiivseteks väideteks; filosoofias vastab sellele seisukohale reduktiivfunktsionalistlik positsioon. Sageli on neid võetud lihtsalt kui tausteeldusi või postulaate keha–vaimu seose olemuse kohta. Nii või teisiti on teadus suutnud üsna kenasti teha oma tööd ka ilma neid küsimusi reaalselt lahendamata. Seletamisel on sildprintsiibi kuju tähtsam kui tema metafüüsiline staatus.

Üldiselt pulbitsevad paljudes teoreetilistes psühhofüüsika-alastes diskussioonides füüsikaliste protsesside ja elamuse vahelisi suhteid puudutavad “filosoofilised” küsimused otse pealispinna all. Niipalju kui ma tean, ei ole seda küsimust filosoofilises kirjanduses kuigi palju arutatud (kuigi vt. aistingute mõõtmise metoodika filosoofilist kriitikat Savage’il (1970)). See teema vääriks nähtavasti põhjalikumat uurimist. [↑](#footnote-ref-292)
292. “Some may find an appeal to preexperimental bridging principles disturbing in an experimental science” — tõlgin siin venelase eeskujul; **juhul, kui** venelase arusaamine on õige, on “rikub puhtust” palju parem vaste kui näiteks “mõjub häirivalt” vms. [↑](#footnote-ref-293)
293. Segane lause, millest ma ei saagi päriselt sotti: “It is striking that each of these candidates is itself a plausible candidate to play a role in facilitating awareness—direct availability for global control.” Panin siia lihtsalt midagi, mis tundus sobivat, aga see ei tarvitse olla see, mida mõtles autor. [↑](#footnote-ref-294)
294. “pre-existing” (selles ja järgmises lauses); venelane on tõlkinud “арпиорных”. [↑](#footnote-ref-295)
295. ​**AM 7.1**​ Näiteks Lycan 1987. [↑](#footnote-ref-296)
296. ​**AM 7.2**​ Churchland ja Churchland (1981) on “Hiina elanikkonna” argumentidele vastu vaielnud, öeldes, et selline süsteem peaks suutma töödelda umbes 1030 000 000 erinevat sisendit võrkkestalt ja veel suuremat arvu aju seesmisi seisundeid. Simulatsioon, mis vajab üht inimest iga sisendsignaali kohta ja üht inimest iga seisundi kohta​, nõuaks seega tohutult rohkem inimesi, kui ühegi maa elanikkonnal pakkuda oleks.

See vastuväide jätab tähelepanuta asjaolu, et nii sisendsignaalidel kui ka seesmistel seisunditel on kindel kombinatoorne struktuur. Me ei vaja üht inimest iga sisendmustri elemendi kohta rohkem kui 10⁸ võrkkesta rakul, seega 210⁸ inimest, vaid võime kujutada sisendit kindlastruktuurilise mustrina, mis nõuaks ainult 10⁸ inimest. Sama kehtib ka seesmiste seisundite kohta. Seega ei ole meil vaja rohkem inimesi, kui ajus on rakke.​ [↑](#footnote-ref-297)
297. Siinkohal jäin lugedes nõutuks: kõigepealt, kas tõesti Churchlandid arvavad, et inimesel on nii palju neuroneid? ja edasi: kust, mis loogika järgi tulevad järsku jutuks sisendid ja seisundid? Hakkasin asja uurima ja selgus, et Chalmers ajab siin lugejat segadusse: kolm lõiku tagasi ütles ta, et iga Hiina elanik pannakse simuleerima üht neuronit, aga nüüd jätab ütlemata, et mudel, mille üle arutlevad Churchlandid ([versioon P. Churchlandi kodulehel](https://patriciachurchland.com/wp-content/uploads/2020/05/1981-Functionalism-Qualia-and-Intentionality.pdf), lk. 133–137), on hoopis midagi muud: siin ei simuleeri elanikud neuroneid, vaid Turingi masina seisunditabeli ridu, mida on teadagi üüratult palju rohkem! [↑](#footnote-ref-298)
298. Nii ütleb autor: “… one person per input and one person per state …”; aga kas õige poleks hoopis “üht inimest iga sisendsignaali ja seisundi kombinatsiooni kohta”? [↑](#footnote-ref-299)
299. ​**AM 7.3**​ Bogen (1981) ja Lycan (1987) arvavad, et sellistes “juhuslikes” olukordades kvaale ei oleks, kuna kvaalid eeldavad teleoloogiat. Niisugusest loogikast tuleneks ​**[lk.387]**​ kummaline järeldus: see, kas süsteemil on kvaalid või mitte, oleneb tema ajaloost. Minu arvates on parem tunnistada, et niisuguses süsteemis oleksid kvaalid, ning ühtlasi rõhutada, et sellise süsteemi juhusliku tekkimise tõenäosus on äärmiselt väike. [↑](#footnote-ref-300)
300. ​**AM 7.4**​ Olen selle üle põhjalikumalt arutlenud artiklis Chalmers 1994a. [↑](#footnote-ref-301)
301. ​**AM 7.5**​ Selle arvu saame, kui paneme tähele, et süsteemi järgmist globaalset seisundit kujutava konditsionaali tagaliikme valikuks on 1010⁹ võimalust. Seega saame tõenäosuseks, et mingile ühele globaalsele seisundile järgneb õige seisund, ${1}/{10^{10⁹}}$. Tegelikult on see arv veelgi väiksem, sest füüsikalise süsteemi iga globaalset seisundit võib realiseerida suur hulk erinevaid “maksimaalseid” seisundeid, ja igast sellisest seisundist peab süsteem siirduma õigesse järgmisse seisundisse. Selliseid konditsionaale, mis peavad kõik kehtima, on 1010⁹ — nii tulebki välja põhitekstis toodud arv. [↑](#footnote-ref-302)
302. ​**AM 7.6**​ Mõned sarnaseid stsenaariumid on välja pakkunud ka Harrison (1981a, 1981b). [↑](#footnote-ref-303)
303. “… for the more powerful **second** argument” — kui vaadata argumentide esitamise järjekorda (“absent qualia”, “fading qualia”, “inverted qualia”, “dancing qualia”), selgub, et **teine** neist on “hääbuvate kvaalide” argument ise, seega on see “second” nähtavasti autori näpuviga. Aga mida ta siis mõtleb? Kuna jao 7.3 järel tuleb jagu 7.4, siis nähtavasti [invertkvaalide argumenti](#ORG.7.4). [↑](#footnote-ref-304)
304. “without loss of generality” — kas selle kohta eesti keeles paremat ütlemist ei ole? [↑](#footnote-ref-305)
305. “uniforms” — kas seda tuleks tõlkida teisiti? “Munder”, “univorm”, “vormiriietus” vist ei sobi, aga mis sobiks? [↑](#footnote-ref-306)
306. “… but its antecedent plausibility is low” — mis on “antecedent plausibility”? [↑](#footnote-ref-307)
307. Ei ole tõsi! (Vt. [Bifurkatsioon](https://et.wikipedia.org/wiki/Bifurkatsioon_%28matemaatika%29).) [↑](#footnote-ref-308)
308. ​**AM 7.7**​ Mingi osa Joe eksitusest saaks ehk olematuks seletada uskumuse konstitueerumisega elamuse alusel, nagu seda on käsitletud [5. peatüki 7. jaos](#ORG.5.7): võib-olla viitab Joe mõiste “punane elamus” nüüd näiteks roosale elamusele, nii et ei saa ka öelda, et ta täielikult eksib. See strateegia ei aitaks aga kindlasti mitte seletada vigu tema otsustustes eristuste kohta. [↑](#footnote-ref-309)
309. Ei leia [5. peatüki 7. jaost](#ORG.5.7) midagi niisugust! [↑](#footnote-ref-310)
310. ​**AM 7.8**​ Cuda (1985) väidab, et selliste ekslike uskumustega süsteemide kirjeldus on mõttetu. Ta ei paku sellele väitele ühtki põhjendust peale selle, et kui see kirjeldus oleks mõttekas, siis oleks sama mõttekas arvata, et ka meie eksime samasugusel viisil, kuid see ei ole (tema enda sõnul) ilmselt mitte nii​. Kuid tundub, et tal ei ole õigus. Oletus, et ma võiksin niiviisi eksida, ei ole mõttetu, kuivõrd selles pole loogilist vastuolu; kogu asi on vaid selles, et oma episteemilise situatsiooni tõttu ma tean, et minu puhul ei ole see oletus tõene, kuna erepunased kvaalid jms. on otseselt minu elamuses. [↑](#footnote-ref-311)
311. “which (he says) it clearly does not.”  — Autor tsiteerib valesti. Vastav koht [originaalartikli](https://philosophy.as.uky.edu/sites/default/files/Against%20Neural%20Chauvinism%20-%20Tom%20Cuda.pdf) 117. leheküljel ütleb: “ If it makes sense to think that Fredm could be mistaken in such a way, then it makes sense to think that we could be mistaken in such a way also. Hence we would have no reason to think that things made of neurons (i.e., ourselves) have red qualia. But clearly it makes no sense to think that we act like we have red qualia, believe that we have them, etc., but that we are all mistaken and really never have any red qualia. Therefore, it makes no sense to think that any Fredi could be mistaken in this way either.” Õige oleks seega mitte “see ei ole nii”, vaid “nii mõelda on mõttetu” vms. [↑](#footnote-ref-312)
312. ​**AM 7.9**​ See ei tähenda, et kuhjaparadoksi meie uurimisvallas üldse ette ei tuleks, vt. näit. Tienson 1987. [↑](#footnote-ref-313)
313. ​**AM 7.10**​ Seda positsiooni omistatakse kõige sagedamini Shoemakerile (1982), kuid teda on kaitsnud ka Horgan (1984a), Putnam (1981) ja mitmed teised. [↑](#footnote-ref-314)
314. Segane sõnastus originaalis, vist näpu- või trükiviga: “… so that there could be no could be no real difference …” [↑](#footnote-ref-315)
315. ​**AM 7.11**​ Pole võimatu, et niisuguseid juhtumeid esineb ka aktuaalses maailmas. Nida-Rümelin tähendab oma huvitavas artiklis (1996), et uurimistööd värvusnägemise alal lubaksid oodata värvuste “pseudonormaalse” nägemise juhtumeid, kus (1) R-kolvikeste reaktsioonimuster on niisugune, mis on tavaliselt omane G-kolvikestele, ja (2) G-kolvikeste reaktsioonimuster on niisugune, mis on tavaliselt omane R-kolvikestele. Eraldi võetuna on (1) ja (2) tavalised daltonismi​ juhtumid. Teoreetiliselt võiksid geneetilised hälbed, mis põhjustavad juhtumeid (1) ja (2), esineda ka koos, andes tulemuseks subjekti, kes oleks käitumise poolest väga sarnane normaalsele subjektile, kuid kelle värvuselamused võiksid väga hästi olla kõigi teiste inimeste omadega võrreldes invertsed. [↑](#footnote-ref-316)
316. “red-green color blindness” — on ju “daltonism” õige sõna? [↑](#footnote-ref-317)
317. “Deemon” (“demon”) niimoodi järsku, ette hoiatamata?! [↑](#footnote-ref-318)
318. ​**AM 7.12**​ Putnam (1981) ja Shoemaker (1982) argumenteerivad selle näite varal kvaalide funktsionalistliku käsitluse vastu, kuid nende argumendid riivavad parimal juhul “jämedateralist” invariantsusprintsiipi, mis ütleb, et seisunditest, mida põhjustavad sinised esemed ja mida subjekt ütleb olevat sinised, sünnivad alati ühesugust tüüpi elamused. Peeneteraline printsiip on endiselt väljaspool ohtu. (Midagi sarnast väidab ka Levine (1988).)

Ühes sama stsenaariumi variandis läbivad ümberkorraldatud ühendustega subjektid kohanemis-, õppimis- ja lõpuks (unustamaks, et nad on asju iial teisiti näinud) amneesiakuuri, nii et pärast kõike seda ei erine nende käitumine senisest. Pole aga kuigi palju põhjust uskuda, et nad on samasugused ka organisatsiooniliselt, eriti kui arvestada, et ümberlülitused kajastuvad endiselt nende aju seisundis. Kui nendega toimuks nüüd mingi eriline protsess, mille tagajärjel nende organisatsioon muutuks taas endiseks, siis võiks vist oletada, et ka ​**[lk.388]**​ elamused peaksid muutuma jälle endisteks. (​Mingil kujul pooldavad hüpoteesi, et elamused selles olukorras ennistuksid,​ Cole (1990) ja Rey (1992).) [↑](#footnote-ref-319)
319. ​**AM 7.13**​ Sarnase argumendi on esitanud Seager (1991, lk. 39–41), kes kirjeldab juhtumit, kus võrkkesta rakud “häälestatakse” optilise spektri kõrgemale sagedusvahemikule. Sellele argumendile võib vastata samamoodi nagu Blocki omale. [↑](#footnote-ref-320)
320. Jätsin ära lause algusest “after some time” — ei ole vaja ju mingit aega, nii hakkab olema kohe. Või? [↑](#footnote-ref-321)
321. Ei saa aru! [↑](#footnote-ref-322)
322. ​**AM 7.14**​ Ühele sarnasele vastuväitele vastab Block, et me võime viia “läätsed” süsteemis sügavamale, lülitades ühendusi ümber näiteks nägemisnärvis või nägemiskorteksis. Ent ka see ei anna meile ühtki juhtu, kus organisatsioonilistel isomorfidel oleksid erinevad elamused. Tõepoolest, Blocki kirjeldus sellistest juhtumitest tundub olevat otseselt ühildatav seisukohaga, et kvaalid sõltuvad “tsentraalsete” süsteemide organisatsioonist. Meid palutakse uskuda elamuste samasust mõningatel juhtudel just sellepärast, et tsentraalsed protsessid jäävad puutumata, ning eeldatakse, et me usume elamused olevat erinevad juhtudel, kus tsentraalsed protsessid on erinevad. Niisugused argumendid ei saa invariantsusprintsiipi mitte kuidagi ümber lükata. [↑](#footnote-ref-323)
323. ​**AM 7.15**​ Käesoleva jao argumendiks on mingil määral innustust andnud Dennetti looke “Kus ma olen?” (1978d). Allpool kirjeldamisele tulevaga mõnevõrra sarnast olukorda on vaadelnud Shoemaker (1982). Veelgi tihedamalt haakuvaid mõtteid võib leida Seageri raamatust (1991, lk. 43), kuigi Seager invariantsusprintsiipi ei poolda. Käesolev raamat oli juba trükki minemas, kui ma avastasin huvitava artikli Arnold Zuboffilt (1994), milles autor, soovides põhjendada üht reduktiivse funktsionalismi versiooni, väidab põhimõtteliselt hüplevate kvaalide argumendile toetudes, et kvaalide hüplemine on ​*a priori* võimatu. [↑](#footnote-ref-324)
324. Omavolitsesin, parandasin autori näpuvea: originaalis on selles reas ainult 10 %, …, 90 % (ilma otspunktideta 0 % ja 100 %), ja niisugusel kujul mõttekäik muidugi vett ei pea. Ühtlasi tegin esitust selgemaks. [↑](#footnote-ref-325)
325. Originaalis: “unless it is in an **accompanying** Cartesian disembodied mind”; ma ei saa aru, mida peab siin tähendama see “accompanying”. [↑](#footnote-ref-326)
326. ​**AM 7.16**​ Midagi sarnast ütleb ka White (1986): kui mittefunktsionaalsed füüsilised erinevused on kvaalidele relevantsed, siis võivad kvaale mõjutada isegi tibatillukesed erinevused DNAs. [↑](#footnote-ref-327)
327. ​**AM 7.17**​ Shoemaker (1982) pakub välja keeruka kriteeriumi otsustamaks, kui spetsiifiline peab olema füsioloogiline omadus, et kvaale fikseerida või, nagu ta ütleb, “realiseerida”. Mulle tundub aga, et kui minu siinne mõttekäik oli õige, siis osutab tema kriteerium tegelikult väga peeneteralisele funktsionaalsele omadusele. [↑](#footnote-ref-328)
328. “The very **immediacy** of the switch seems to make a significant difference” — ei oska sellest teisiti aru saada, aga kas see on õige? [↑](#footnote-ref-329)
329. “antecedent plausibility” — mis see on? [↑](#footnote-ref-330)
330. Jälle “antecedent plausibility”! [↑](#footnote-ref-331)
331. ​**AM 7.18**​ Selle idee pakkus ühes jutuajamises välja Terry Horgan. [↑](#footnote-ref-332)
332. “… for every physical system that gives rise to conscious experience, there is some functional organization ​*F* realized by the system, such that it is naturally necessary that any system that realizes ​*F* will have identical conscious experiences. To pick out the relevant ​*F*​, we need to go to a fine enough grain to fix cognitive states such as judgments. This in turn can be achieved by requiring that ​*F* is fine-grained enough to fix the mechanisms responsible for the production of behavior, and to fix behavioral dispositions.” — omavolitsesin, jätsin tolle F-tähe ära, sai selgem. Variant ilma omavolitsemiseta: “igas füüsikalises süsteemis, milles sünnib teadvuselamus, realiseerub mingi niisugune funktsionaalne organisatsioon ​*F*​, et igal süsteemil, milles realiseerub​ *F*​, on loodusliku paratamatusega identsed teadvuselamused”, ja järgmises lauses “Selleks, et leida õige organisatsioon ​*F*​ …” ja ülejärgmises: “… kui nõuda, et ​*F* oleks piisavalt peeneteraline …” [↑](#footnote-ref-333)
333. Kuidas on tingimuslause poolte nimed (siin “premise” ja “conclusion”) eesti keeles? [Eestikeelse Wikipedia artiklis “Tingimuslause”](https://et.wikipedia.org/wiki/Tingimuslause) kasutatakse neis tähendustes sõnu “tingimus” ja “väide”. [↑](#footnote-ref-334)
334. “… for every physical system that gives rise to conscious experience, there is some functional organization ​*F* realized by the system, such that it is naturally necessary that any system that realizes ​*F* will have identical conscious experiences. To pick out the relevant ​*F*​, we need to go to a fine enough grain to fix cognitive states such as judgments. This in turn can be achieved by requiring that ​*F* is fine-grained enough to fix the mechanisms responsible for the production of behavior, and to fix behavioral dispositions.” — omavolitsesin, jätsin tolle F-tähe ära, sai selgem. Variant ilma omavolitsemiseta: “igas füüsikalises süsteemis, milles sünnib teadvuselamus, realiseerub mingi niisugune funktsionaalne organisatsioon ​*F*​, et igal süsteemil, milles realiseerub​ *F*​, on loodusliku paratamatusega identsed teadvuselamused”, ja järgmises lauses “Selleks, et leida õige organisatsioon ​*F*​ …” ja ülejärgmises: “… kui nõuda, et ​*F* oleks piisavalt peeneteraline …” [↑](#footnote-ref-335)
335. Omavolitsesin nagu enne, jätsin F-tähe ära: “… a law, for certain functional organizations *F*​, that realization of *F* will be accompanied by a specific kind of conscious experience.” Omavolitsemiseta variant: “et mingi kindla funktsionaalse organisatsiooni ​*F* puhul kaasneb ​*F* realiseeringuga mingit kindlat liiki teadvuselamus.” [↑](#footnote-ref-336)
336. Originaalis “state”; proovisin “seisundit” ja “olekut”, kumbki ei kõla hästi. Panin “võimalus”, “võimalik väärtus” ja “valitud väärtus” ja sain (mõningase ümbersõnastamise hinnaga) enda arvates parema tulemuse. [↑](#footnote-ref-337)
337. Kohmakas sõnastus originaalis, kirjutasin selgemaks. [↑](#footnote-ref-338)
338. “message”; termini võttis tarvitusele Shannon oma artiklis [“A Mathematical Theory of Communication”](https://people.math.harvard.edu/~ctm/home/text/others/shannon/entropy/entropy.pdf) (kohe alguses, teises lõigus); usun, et on olemas käibiv eestikeelne vaste, aga mul pole õnnestunud seda leida. [↑](#footnote-ref-339)
339. “proximity” (vt. [artikkel “Proximity space” ingliskeelses Wikipedias](https://en.wikipedia.org/wiki/Proximity_space)); mul ei õnnestunud Internetist käibivat eestikeelset vastet leida, kuigi usun, et see on olemas. [↑](#footnote-ref-340)
340. Võtsin terminid “lohk” ja “põhipind” [eestikeelsest Wikipediast](https://et.wikipedia.org/wiki/CD%22%20%5Cl%20%22Tehnoloogia)​; selle kohta, kuidas lohkudest ja põhipinna lõikudest lohkude vahel moodustuvad bitid, on hea seletus [siin](https://www.cs.tlu.ee/~rinde/mm_materjal/pdf/mm_cd.pdf). [↑](#footnote-ref-341)
341. Originaalis järgneb “, one for every pit or land on the disk”. Autori ettekujutus asjast on lihtsustav: [tegelikult](https://www.cs.tlu.ee/~rinde/mm_materjal/pdf/mm_cd.pdf) vastab igale 1-bitile üleminek lohult põhipinnale või vastupidi, ja vahepealsed alad (jätkuv lohk või jätkuv põhipind) vastavad 0-bittidele, millede arv on võrdeline vahepealse ala pikkusega. Seega on originaali lauselõpp vigane, aga kui see ära jätta, läheb lause õigeks. [↑](#footnote-ref-342)
342. Jätsin siitkohalt tähele panemata sõnakese “different” (“teisel”), mis ei tundu siia loogiliselt sobivat.  — Tagant järele mõeldes: võib-olla on autor mõelnud “distinct roles on different causal pathways” — “erinevaid rolle erinevatel põhjuslikel trajektooridel”? [↑](#footnote-ref-343)
343. Huh? Mis mõttes? Ei saa aru! [↑](#footnote-ref-344)
344. Omavolitsesin: originaalis on “we can respond separately to specific queries about the color at a given location”, millest ma ei oska aru saada teisiti kui “me võime anda küsimustele värvuse kohta **mingis kindlas nägemisvälja punktis** erinevaid vastuseid” (või mida veel tähendab see “separately”?) — jääb mulje, et me võime anda erinevaid vastuseid värvuse kohta ühes ja samas punktis; autor on vist tahtnud öelda siiski nii, nagu mina siin tõlkisin. [↑](#footnote-ref-345)
345. Segane sõnastus: “this separate space of effects for each location yields a separate subspace for each location”; püüdsin selgemaks teha, loodetavasti ei maganud ma sellega mõnd olulist peensust maha. [↑](#footnote-ref-346)
346. Miks “mõnede vormide” (“certain holographic forms of information storage”)? Kas holograafia ei ole juba iseenesest “holistlik” (s.t. kujutise iga punkt on osa interferentsipildist, millesse annab oma panuse kujutatava objekti iga punkt)? Kas poleks õigem öelda lihtsalt “nagu holograafia puhul”? [↑](#footnote-ref-347)
347. ​**AM 8.1**​ Seoseid teadvuse ja informatsiooni vahel on välja pakkunud veel Bohm (1980), Sayre (1976) ja Velmans (1991), kuigi nende ideed erinevad üksikasjus minu omadest oluliselt. Siiski väärib tähelepanu ka Sayre’i idee informatsiooni “neutraalsest monismist”. (Aitäh Steve Horstile, kes minu tähelepanu sellele juhtis.) Topeltaspektiprintsiibiga üsna sarnasest ideest on juttu raamatus Lockwood 1989 (11. ptk.), kuigi tema ei formuleeri seda informatsioonilistes mõistetes. [↑](#footnote-ref-348)
348. Kus? Ei leia! [↑](#footnote-ref-349)
349. Siin on originaalis vahel sõnake “here”, millega ma ei oska midagi peale hakata: “the physical information space here is …” [↑](#footnote-ref-350)
350. “roots”; “juured” kõlab mööndavasti narripoolselt, aga mis oleks parem? [↑](#footnote-ref-351)
351. Improviseerisin; originaalis on lihtsalt “some of the feeling of outrageous coincidence will be removed”; määrav artikkel vihjaks otsekui millelegi ülal öeldule, aga ma ei suutnud leida tekstist midagi niisugust, mis siia hästi sobiks. [↑](#footnote-ref-352)
352. ​**AM 8.2**​ Üks minu avaldamata töö (Chalmers 1990) vaatlebki seda teadvuse ja teadvust puudutavate otsustuste vahelise suhte mõistmise strateegiat ning kasutab seda teadvuse (mustritele ja informatsiooni mõistele toetuva) “alusteooria” väljatöötamiseks; siit ongi saanud alguse mõned siinse peatüki ideed. Tolles artiklis nimetasin ma seletusliku koherentsuse nõuet *koherentsustestiks*​; selle testi edukat läbimist tuleb nõuda igalt teadvuseteoorialt. [↑](#footnote-ref-353)
353. Miks “võib-olla”? Kas see ei ole kindel teadmine? [↑](#footnote-ref-354)
354. “This is one of those” — ei saa originaalist aru: kuidas see saab olla vastuseks küsimusele “What is the difference between these states?”? Kui just mõte ei ole niisugune, nagu mina seda siin tõlkisin. [↑](#footnote-ref-355)
355. “This is one of those” — ei saa originaalist aru: kuidas see saab olla vastuseks küsimusele “What is the difference between these states?”? Kui just mõte ei ole niisugune, nagu mina seda siin tõlkisin. [↑](#footnote-ref-356)
356. “this information” — miks “this”? Siinkohal on mul kiusatus jätta see sõnake ära ja teha lause palju lühemaks. [↑](#footnote-ref-357)
357. ​**AM 8.3**​ Kuid väite kohta, et termostaatidel on uskumused ja soovid, vt. McCarthy 1979. [↑](#footnote-ref-358)
358. ​**AM 8.4**​ Wheeleri erilise tähelepanu all on mõõtmistulemused ehk “vastused jah-ei-küsimustele” kui kõige alus, ja sellisena võib tema vaateviis olla lähedasem idealismile kui seisukohale, mida mina siin pakun. [↑](#footnote-ref-359)
359. FIXME leheküljenumbrid õigeks! [↑](#footnote-ref-360)
360. ​**AM 8.5**​ Wright’i raamatus (1988) leidub Fredkini ideede huvitav kirjeldus ning on rohkem juttu ka nende alusmetafüüsikast. [↑](#footnote-ref-361)
361. Autori tekstist ei selgu mulle, mille “under” see “underlying quality” on (ja seega millele viitab minu “nende”). Lõigu kolm esimest lauset on originaalis järgnevad: “The second problem is that it is not obvious that the notion of pure informational flux is coherent. One may feel that on this view the world is too lacking in substance to *be* a world. Could there be differences that are *primitive* differences, not grounded in differences in any underlying quality?”. Puhtalt teksti vaadates tuleb pähe kaks varianti: “pure informational flux” ja “world”; ma ei oska otsustada kummagi kasuks. [↑](#footnote-ref-362)
362. Leheküljenumbrid õigeks! [↑](#footnote-ref-363)
363. ​**AM 8.6**​ ​**[lk.389]**​ Viljakaid arutelusid russelliaanliku vaate raskuste üle leidub allikates Foster (1991, lk. 119-30) ja Lockwood (1992). [↑](#footnote-ref-364)
364. ​**AM 8.7**​ Lockwood (1992) oletab, et russelliaanlik positsioon võib seesugustel eesmärkidel viidata elementaarseadustele. Ta ei vaatle vastuväidet, et uute seaduste tarvituselevõtt sellisel kujul näib ohustavat russelliaanliku vaate algset atraktiivsust. Uute seadustega tulevad kaasa needsamad epifenomenalismi raskused, mida russelllik vaade tõotas vältida, ja nad nõuavad ka ontoloogia olulist laiendamist füüsika põhjendamiseks vajalike seesmiste omaduste ringist väljapoole. (Tuleb märkida, et Lockwood ei toetu nendele seadustele teralisusprobleemi lahendamiseks; selleks on tal teine idee — omaette põnev hüpotees, mis toetub kvantmehhaanikale). [↑](#footnote-ref-365)
365. Ma ei saa selle sulgudes märkuse mõttest aru, tõlgin venelase eeskujul. [↑](#footnote-ref-366)
366. Omavolitsesin kogu selles jaos kõvasti, tegin teksti selgemaks ja voolujoonelisemaks. [↑](#footnote-ref-367)
367. ​**AM 9.1**​ Käesoleva osa materjal on suures osas võetud artiklist Chalmers (1994a). [↑](#footnote-ref-368)
368. Omavolitsesin siin kõvasti. Ma no ausalt püüdsin sõnastada seda definitsiooni originaalilähedasemalt, aga sellest ei tahtnud midagi välja tulla. Variant, mille juures ma viimaks käega lõin (kuna ma ei suutnud juba enam omaenda kirjutatud teksti jälgida), oli järgmine: “Füüsiline süsteem ​*P* on CSA ​*M* *implementatsioon*​, kui tema seisundid saab lahutada komponentideks [​*s*​¹, …, ​*s*​ⁿ], korraldada vastavuse ​*f*​ tema osaseisundite ​*s*​ⁱ hulgast CSA ​*M* osaseisundite $S^{j}$ hulka, ning kui ka sisend- ja väljundandmetega saab sooritada analoogilise lahutuse ja korraldada vastavuse, nii et CSA ​*M* iga siirdereegli $$ puhul on tõsi, et kui ​*P* on seisundis [​*s*​¹, …, *s*​ⁿ], mille vastavad äsjakirjeldatud vastavuse mõttes formaalsele seisundile [​*S*​¹, …, ​*S*​ⁿ], ning kui tema sisendisse saabuvad andmed $\left[i^{1},…,i^{k}\right]$, mis vastavad samas mõttes andmetele $\left[I^{1},…,I^{k}\right]$, siis siirdub ​*P* alati seisundisse, mis vastab seisundile [​*S*​′¹, …, ​*S*​′ⁿ], ning väljastab andmed, mis vastavad väljundandmetele $\left[O^{1},…,O^{l}\right]$.”  — Loodan, et nii, nagu mul praegu põhitekstis on, sai selgem ja et mõte jäi endiseks, aga sõnastus läks palju lihtsamaks. (Muu hulgas ei olnud mul üldse tarvis mainida siirdetabelit, sest igale võimalikule seisundite ja sisendandmete kombinatsioonile vastab nagunii ka omaette kirje siirdetabelis.)  — Originaalis on ka näpu- või trükiviga: “in” *pro*​ “$i^{k}$” (mille venelane on muide parandamata jätnud). [↑](#footnote-ref-369)
369. Ei saa aru: “… the value of each element of the vector must supervene on a separate region of the physical system, to ensure that the causal organization relates distinct components of the system.”  — Võib-olla tahab autor öelda, et süsteemi osad, millele vastavad vektori komponendid, ei tohi omavahel kattuda? [↑](#footnote-ref-370)
370. ​**AM 9.2**​ Putnamil (1988, lk. 120–25) on terve omaette mõttekäik põhjendamaks väidet, et iga tavaline lahtine süsteem on iga lõpliku seisundiautomaadi implementatsioon. Ma analüüsin seda argumenti põhjalikult artiklis Chalmers 1995a. Analüüs näitab, et Putnami argumendile annab jõu eeldus, et füüsikaliste seisundite siirdekonditsionaalideel, mis figureerivad implementatsiooni definitsioonis, puudub igasuguse tugev modaalne jõud​. [↑](#footnote-ref-371)
371. Sellest ei ole vist võimalik aru saada lugejal, kes pole [seda Chalmersi artiklit](https://consc.net/papers/rock.html) lugenud. [↑](#footnote-ref-372)
372. ​**AM 9.3**​ Seda arvutuste seletusliku rolli mõistmise viisi kognitiivteaduses kirjeldan ma põhjalikumalt artiklis Chalmers 1994b. [↑](#footnote-ref-373)
373. ​**AM 9.4**​ Sarnaseid seisukohti on esitanud Korb (1991) ja Newton (1989); mõlemad väidavad, et hiina tuba on hea argument kui mitte masinintentsionaalsuse, siis vähemalt masinteadvuse vastu. [↑](#footnote-ref-374)
374. ​**AM 9.5**​ Hofstadter (1981) vaatleb sarnast spektrit, mille ühes otsas on aju ja teises hiina tuba. [↑](#footnote-ref-375)
375. ​**AM 9.6**​ Mõte, et hiina toa homunkulus on analoogiline kolju sisemuses ringi sibava deemoniga, kuulub Haugelandile (1980). [↑](#footnote-ref-376)
376. “… the *prima facie* case against external objections to artificial intelligence is strong …”: ei oska sellest osalausest kuidagi aru saada, tõlgin venelase eeskujul, s.t. loen selle “​*prima facie*​” mitte sõna “case”, vaid sõna “strong” juurde kuuluvaks. [↑](#footnote-ref-377)
377. Tõlkimatu sõnademäng: “sum” tähendab ladina keeles “olen” (ning viitab Descartes’i ütlusele “cogito, ergo sum”), inglise keeles aga “summeerin”.  — Kas jätta tõlkesse “​*sum*​” (kursiivis, tsitaatsõnana) või panna “olen”, ja (mõlemal juhul) selgitada sõnademängu tõlkemärkuses? [↑](#footnote-ref-378)
378. ​**AM 9.7**​ On tähelepanuväärne, et kuigi Dreyfus (1972) pani oma raamatule, milles ta taolisi vastuväiteid esitab, pealkirjaks “Mida arvutid ei suuda” [“What Computers Can’t Do”], möönis ta hiljem, et õiget tüüpi arvutussüsteemide (näiteks konnektsionistlike süsteemide) kohta need vastuväited ei käi. Tegelikkuses on seda, “mida arvutid suudavad”, sageli samastatud sellega, mida suudab teha mingi väga kitsas arvutussüsteemide klass. [↑](#footnote-ref-379)
379. ​**AM 9.8**​ Selle otsese vastulause göödellikele argumentidele esitas esmakordselt vist Putnam (1960) ja minu teada ei ole seda kunagi ümber lükatud, hoolimata Lucase ja Penrose'i meeleheitlikest jõupingutustest. Penrose (1994, ptk 3.3) väidab, et kuna ta suudab otsustada aksioomide tõesuse ja järeldusreeglite õigsuse üle, siis suudab ta järelikult ka mõista tema enda arutlusi kujutava formaalse süsteemi vastuolutust. Nähtavasti toetub ta seda öeldes eeldusele, et arvutussüsteem on ennekõike aksioomide ja reeglite süsteem. Kuid üldjuhul ei pea ta seda olema. (Seda ei ole näiteks aju simulatsioon neurontasemel.) Ja isegi aksioomidest ja reeglisüsteemidest kõneldes ei ole minule hoopiski ilmne, et me suudaksime kindlaks teha iga reegli õigsust meie süsteemis, eriti nende, mille rakendamine nõuab väljumist tavapäraste arvutustehete piiridest ja nõuavad korduvaid gödeliseerimisi​; sellepärast ongi göödellikel argumentidel inimese juhul reaalne kaal. [↑](#footnote-ref-380)
380. “especially those applying to the outer limits of ordinal counting in iterated Gödelization” — sellest kohast ei saanud ma mõhkugi aru, tõlgin venelase eeskujul: tema on tõlkinud “ordinal” “tavapäraseks” (tal on “за пределы обычного вычисления”) — nähtavasti oletades, et see “ordinal” on autori näpuviga ja õige on “ordinary”. Mul on raske seda talle ette heita, sest “[ordinal](https://en.wikipedia.org/wiki/Ordinal_number)” (“[ordinaalarv](https://et.wikipedia.org/wiki/Ordinaalarv)”) ei tundu tõepoolest tahtvat siia kuidagi sobida, kuna vähemasti see aritmeetika, mida süsteem peaks võimaldama, et rahuldada Gödeli teoreemi eeldusi, ei vaja ordinaalarve sugugi — tavalisest naturaalarvude aritmeetikast aitab juba küll ja küll.) [↑](#footnote-ref-381)
381. Omavolitsesin veidi, vabanesin tarbetust tähistusest. [↑](#footnote-ref-382)
382. Originaalis “absolute continuity”, aga see saab olla ainult autori näpuviga, sest see termin tähendab matemaatikas [hoopis midagi muud](https://en.wikipedia.org/wiki/Absolute_continuity). [↑](#footnote-ref-383)
383. ​**AM 9.9**​ Seda tasub teha juhul, kui konkreetne ümardustase 10-10 toob käitumise jaotusse süstemaatilisi nihkeid. Kindluse mõttes võiksime — eeldusel, et müra tase on 10-10 — aproksimeerida süsteemi täpsusega 10-20, ning sama täpsusega aproksimeerida ka müra jaotust. [↑](#footnote-ref-384)
384. Ei osanud seda “encode” ülejäänud tekstiga siduda teisiti, kui lausestust põhjalikult ümber tehes.  — Ja veel, tekstis on öeldud “one might encode” — otsekui oleks jutt mingist põhimõtteliselt võimalikust asjast; aga peatumisprobleem on teatavasti põhimõtteliselt lahendamatu. [↑](#footnote-ref-385)
385. ​**AM 9.10**​ ​**[lk.390]**​Siin-seal — üldiselt aga siiski ainult vaimufilosoofias — tähistatakse terminiga “arvutus” ainult sümboolsete arvutuste klassi või arvutusi, millede objektiks on representatsioonid (s.t. süsteemid, mille süntaktilised algelemendid on ühtlasi ka semantilised algelemendid). Muidugi ei olene sellest terminoloogilisest nüansist kuigi palju: tehisintellekti seisukohalt on oluline see, et leiduks formaalne süsteem, mille implementatsioonist piisab vaimuks, ükskõik kas ta on selle kriteeriumi järgi arvutus või mitte. Kuid nii või teisiti tuleb märkida, et selle termini kasutamine säärasel viisil katkestab ühenduse tema informaatiliste allikatega. Selles mõttes ei saa isegi enamikku Turingi masinaid pidada arvutusteks, kuna vaid mõnede kohta nendest võib öelda, et nad sooritavad tehteid kontseptuaalsete representatsioonidega. Sarnastel põhjustel viib arvutuste klassi seesugune piiramine just nende (Churchi–Turingi mõttes) universaalsuse kaoni​, mis annaks muidu võib-olla parima aluse (funktsionaalse) tehisintellekti teesidesse uskumiseks. [↑](#footnote-ref-386)
386. Mida see tähendab? On olemas [Churchi–Turingi tees](https://en.wikipedia.org/wiki/Church%E2%80%93Turing_thesis), kas jutt on mingil moel sellest? [↑](#footnote-ref-387)
387. “basic value”; “baasväärtus” ei ole päris kindlasti hea vaste, aga ma ei oska siia midagi paremat panna. Kolm lauset edasi tuleb “combination of basic values”, kus “kvantmehhaanilisem” oleks ehk öelda “as … combination of eigenstates” ehk “omaseisundite kombinatsioonina”. [↑](#footnote-ref-388)
388. ​**AM 10.1**​ Vähemalt sellise spinn-osakese puhul nagu elektron. Jätan kõrvale juhud, kus spinnil on veel muid baasväärtusi. [↑](#footnote-ref-389)
389. “nonseparable” — kindlasti on eestikeelne termin olemas, aga ma pole seda leidnud. [↑](#footnote-ref-390)
390. “meat and potatoes” — kuidas seda eesti keeles kujukamalt edasi anda? [↑](#footnote-ref-391)
391. ​**AM 10.2**​ See on, nagu kogu siinne käsitlus üldse, lihtsustus. Kuna ükski mõõtmine ei ole absoluutselt täpne, siis ei ole tulemuseks kunagi absoluutselt ühene asukoht; täpsem on öelda, et lainefunktsioon kollabeerub seisundiks, kus kogu amplituud koondub väga kitsasse asukohtade vahemikku. Kuid esituse mõttes on lihtsam kõnelda nii, nagu oleksid asukohad pärast kollabeerumist tõepoolest ühesed. [↑](#footnote-ref-392)
392. ​**AM 10.3**​ Või pideva suuruse korral tõenäosustihedust. [↑](#footnote-ref-393)
393. Omavolitsesin, originaali lause kõlab minu kõrvus imelikult: “What could be happening in the real world to make the predictions of the calculus so accurate?” [↑](#footnote-ref-394)
394. Ei saa aru, ei oska tõlkida: “the very different measurement dynamics have an independent source”. Mis mõttes oleks see “measurement dynamics” “very different”? (Ja üldse, mida õieti mõtleb autor siin “dünaamika” all?) [↑](#footnote-ref-395)
395. ​**AM 10.4**​ Albert (1992) juhib tähelepanu sellele, et “teadvus” on sama ebamäärane mõiste nagu “mõõtmine” ja “makroskoopiline”. Kuid mulle tundub, et mingi osa selle kriteeriumi atraktiivsusest tuleb sellest, et küsimust, kas mingil süsteemil on teadvus või mitte, on võimalik pidada faktiküsimuseks. [↑](#footnote-ref-396)
396. Ei saa aru, ei oska tõlkida: “the very different measurement dynamics have an independent source”. Mis mõttes oleks see “measurement dynamics” “very different”? (Ja üldse, mida õieti mõtleb autor siin “dünaamika” all?) [↑](#footnote-ref-397)
397. Omavolitsesin: autoril on “what sort of effect on consciousness brings about collapse”, millest ma ei oska sotti saada ja mis tundub pigem autori näpuveana. [↑](#footnote-ref-398)
398. “effective collapse” — mida see tähendab? [↑](#footnote-ref-399)
399. Lisasin selle sõna omalt poolt — mulle tundub, et see toob autori tegeliku mõtte paremini välja, arvestades järgnevat lõiku, mille sõnadega “ei olegi fakte” see “faktiliselt” riimuma peaks. [↑](#footnote-ref-400)
400. Lisasin need sulgudes sõnad omalt poolt, et lugeja ei peaks tagasi minema ja järele vaatama. [↑](#footnote-ref-401)
401. “last paragraph” — kas see, millest eelmises lõigus juttu oli, on operatsionalism?  — See siin on ainus lause kogu raamatus, kus esineb sõna “operatsionalism”! [↑](#footnote-ref-402)
402. [Ghirardi–Rimini–Weber](https://en.wikipedia.org/wiki/Ghirardi%E2%80%93Rimini%E2%80%93Weber_theory) [↑](#footnote-ref-403)
403. ​**AM 10.5**​ Võlgnen GRW interpretatsiooni arutamisel tänu allikatele Albert ja Loewer (1990) ja Albert (1992). [↑](#footnote-ref-404)
404. “nonseparability” — kuidas see eesti keeles oleks? Ma tean, et kvantfüüsika matemaatilises aparaadis on “separability” mingi tähtis asi, aga ei tunne asja nii põhjalikult, et sellest aru saada ja eestikeelseid termineid teada. [↑](#footnote-ref-405)
405. ​**AM 10.6**​ Pangem tähele, et see siin on käesolevas peatükis ainus koht, kus ma puudutan Belli teoreemi ja Einstein-Podolsky-Roseni (EPR) tulemusi. Mõnikord peetakse neid tulemusi kvantmehhaanika filosoofiliste probleemide peamiseks allikaks, kuid minu arvates kerkivad need probleemid juba enne EPR järeldusi. Isegi ilma EPR mõttekäikudeta seisaks meie ees raske valik kollapsi, peitmuutujate ja Everetti vahel. EPR lisab peitmuutujate teooriate raskustele vaid uusi, näidates, et need muutujad​ (nagu ka kollaps) peavad olema mittelokaalsed, ja võib arvata, et see annab Everetti interpretatsioonile kui ainsale selle tulemusega ühildatavale lokaalsele interpretatsioonile atraktiivsust juurde. [↑](#footnote-ref-406)
406. “they” — kahemõtteline — võib puht keeleliselt võttes viidata nii muutujatele kui teooriatele. [↑](#footnote-ref-407)
407. ​**AM 10.7**​ “Ühe suure maailma” vaade näib olevat füüsikute seas (eriti kvantkosmoloogide seas, kes seda mudelit alatasa kasutavad) kõige levinum arusaam Everetti interpretatsioonist. Hargmaailmade idee on suuresti populariseerimise käigus tekkinud artefakt. Mõnikord räägivad isegi ühe suure maailma vaate pooldajad “hargnemisest”, kuid see on lihtsalt kujukas viis kõnelda lainefunktsiooni ajalisest muutumisest superpositsiooniliseks. Mingit omaette protsessi, mille käigus maailmad mitmeks hargneksid, ei ole; parimal juhul toimub omamoodi lokaalne hargnemine lainefunktsioonis. Kuid nii või teisiti arvan ma, et parem on “hargnemisest” mitte rääkida, sest sellest tuleb ainult segadust. [↑](#footnote-ref-408)
408. Omavolitsesin, tegin lühemaks. [↑](#footnote-ref-409)
409. Omavolitsesin: originaalis on öeldud: “Why do **the only minds associated with a superposed brain state** correspond to its decomposition along the preferred basis?”, s.t. “Miks peaksid **ainult superpositsioonilise ajuseisundiga seonduvad vaimud** vastama selle seisundi lahutusele eelisbaasi järgi?”, millel ei tundu mõtet olevat ja mis ei sobi kokku ka järgneva lausega. (Ka venelane on tõlkinud nagu mina: “почему сознания, связанные с суперпозиционным состоянием мозга, соответствуют **только его декомпозиции сообразно некоему предпочтительному базису**​?”) [↑](#footnote-ref-410)
410. ​**AM 10.8**​ Seda vastuväidet on esitanud teiste hulgas Bell (1981), Bohm ja Hiley (1993) ning Hodgson (1988). [↑](#footnote-ref-411)
411. Ei saa loogikast aru: kuidas ei peaks olema tegemist tõenäosustega, kui see 90:10 on **tõenäosuste** suhe? [↑](#footnote-ref-412)
412. Lähtun siin oletusest (mis tundub tekstiga sobivat), et autor mõtleb selle “measure” all siin [seda, mida nimetatakse eesti keeles mõõduks](https://en.wikipedia.org/wiki/Measure_%28mathematics%29) (nagu seda on näiteks [Lebesgue’i mõõt](https://et.wikipedia.org/wiki/Lebesgue%27i_m%C3%B5%C3%B5t)). [↑](#footnote-ref-413)
413. “canonical measure” — kas see on termin? [↑](#footnote-ref-414)
414. Ei saa aru — mis “components”? [↑](#footnote-ref-415)
415. ​**AM 10.9**​ Kõige mõistlikum oleks võib-olla teha oma panus kvantmehhanismi abil, mis annab tõenäosusega 0.999 eitava ja tõenäosusega 0.001 jaatava vastuse. Kui Everetti seisukoht on ekslik, jääb mulle peaaegu kindlasti õigus, kui aga tõene, jääb vähemalt üks minu vaimu harudest elama. [↑](#footnote-ref-416)
416. Panin kursiivi tekstiosad, mis on kursiivis originaalraamatu bibliograafias. Meil on arvatavasti teistsugused konventsioonid (näiteks on meil vist kombeks panna ajakirja aastakäigunumbrid rasvasesse kirja); kuna ma neid täpselt ei tunne, siis ei hakanud ma siin omaloomingut harrastama. [↑](#footnote-ref-417)
417. Puudulikud ilmumisandmed. [Täpsed andmed on](https://link.springer.com/article/10.1007/BF00413692): ​*Synthese*​, Volume 108, pp. 309–333 (1996). [↑](#footnote-ref-418)